Pre-analysis Plan for "Classroom Management in Secondary Schools: Reinventing Our Classroom"

Sule Alan[*], Kumar Biswas[†], Christina Sarah Hauser[‡], Shwetlena Sabarwal [§]

March 2023

1 Introduction

We evaluate an educational program, Reinventing our Classroom (RoC), that aims to provide teachers with simple pedagogical tools to manage their classrooms for more effective teaching in overcrowded secondary school classrooms. The core idea is to give students the responsibility of establishing good behavioral norms in their classroom. To this end, the program suggests creating groups within classrooms with rotating group leaders taking charge to ensure a peaceful learning environment for all. With the minimal interference from teachers, students first collectively establish the types of acceptable and unacceptable behaviors in the classroom and put forward a classroom code of conduct outlining good behavioral norms expected (and desired) in their classroom. Then, they offer ways to make these good norms take hold and stick. The most critical component of the program is that students self-assess and self-monitor their progress weekly and announce it to the classroom through charts and evolving graphics.

The program targets secondary schools in Narsingdi district, Dhaka division, Bangladesh. A total of 129 schools were enlisted for the project involving 517 teachers and about 7,000 students.[1] We enlisted one class teacher and one assistant teacher from each school to help implement the program. Baseline data were collected in July and August 2022 when our students were in 6th grade. When they started grade 7 in January 2023, we collected baseline data from their new teachers, and from new students. 64 schools were randomly assigned to treatment. Sixth-grade teachers in 63 treatment schools were trained in August 2022, and seventh-grade teachers in January 2023. The program implementation has been ongoing since September 2022 and will end in March 2023.

[*]European University Institute, Department of Economics, Bilkent University and J-PAL.
[†]World Bank Group.
[‡]European University Institute, Department of Economics.
[§]World Bank Group.

[1] 2 schools dropped out of the program after baseline, 1 from treatment, 1 control group.

Our central hypothesis is that the program will increase teacher and student engagement and well-being by improving behavioral norms and making the classroom amenable to teaching, free from disruptive behavior. In addition, we expect a more cooperative and cohesive classroom environment in treated schools.

2 Outcomes

We will test the effectiveness of the program with respect to a wide range of student and teacher outcomes. To construct our outcomes, we use the following data collection tools: Incentivized experiments, cognitive tests, social networks, administrative records and surveys.

Outcomes of interest for the teachers:

- Mental well-being: motivation, frustration, self worth
- Competency: teaching and classroom management
- Teaching and feedback giving styles
- Absenteeism
- Autonomy, locus of control

Outcomes of interest for students:

- An incentivized public good game to measure cooperation
- Cognitive tests: math and verbal tests.
- Administrative records of absenteeism.
- Social Networks (density, isolation, segregation)
- Sense of belonging and trust
- Behavioral norms
- Perceived support from the teacher
- Mental well-being and self worth

3 Empirical Model

Below, we lay out the empirical specification that we will use to estimate the effect of the treatment.

To test the null hypothesis that the program had no impact on the outcome y, we estimate the average treatment effect conditioning on baseline covariates that are predictive of the outcome of interest:

$$y_{is} = \alpha_0 + \alpha_1 T_s + X_{is}^{'}\gamma + Other_{is} + \delta_b + \varepsilon_{is}$$

where T_s is a dummy variable which equals 1 if school s is in the treatment group and zero otherwise, and X_{is}' is a vector of observables for student i in school s that are potentially predictive of the outcome y. These include demographics, baseline cognitive and sociocognitive skills, IQ (measured by Raven's Progressive Matrices), cognitive empathy (Theory-of-Mind), and outcome variables collected at baseline. $Other_{is}$ captures other variables that might be added for specific regressions, and δ_b are strata (subdistrict) fixed effects. Because we expect some non-compliance, the estimated $\hat{\alpha}_1$ is the intent to treat effect. We will cluster the standard errors at the school level and adjust for multiple hypotheses testing.

We will conduct heterogeneity analysis based on teacher and student gender as well baseline cognitive empathy (Theory of Mind) and IQ. We will also estimate treatment effects on class and assistant teachers separately.

Table 1: Power calculations

Panel I: Students									
Outcome	Alpha	Power	n	k	N	MDE	Mean	SD	ICC
Math score	0.05	0.8	65	74	6670	0.18	0.00	1	0.12
Bangla score	0.05	0.8	65	74	6670	0.20	0.00	1	0.16
Prescriptive norms	0.05	0.8	65	74	6639	0.12	0.00	1	0.04
Descriptive norms	0.05	0.8	65	74	6649	0.13	0.00	1	0.05
Mental health	0.05	0.8	65	74	6634	0.11	0.00	1	0.03
Empathic concern	0.05	0.8	65	74	6656	0.12	0.00	1	0.05
Perspective taking	0.05	0.8	65	74	6650	0.12	0.00	1	0.04
Impulsivity	0.05	0.8	65	74	6644	0.11	0.00	1	0.04
Trust	0.05	0.8	65	74	6630	0.09	0.00	1	0.02
Teacher support	0.05	0.8	65	74	6605	0.13	0.00	1	0.06
Sense of belonging	0.05	0.8	65	74	6631	0.13	0.00	1	0.06
Panel II: Teachers									
Outcome	Alpha	Power	n	k	N	MDE	Mean	SD	ICC
Locus of control	0.05	0.8	65	4	517	0.27	0.00	1	0.06
Frustration	0.05	0.8	65	4	517	0.25	0.00	1	0.02
Motivation	0.05	0.8	65	4	517	0.26	0.00	1	0.04
Mental wellbeing	0.05	0.8	65	4	517	0.26	0.00	1	0.03
Teaching comp.	0.05	0.8	65	4	517	0.27	0.00	1	0.06
Classroom manag. comp.	0.05	0.8	65	4	517	0.25	0.00	1	0.01
Learner-centered teaching	0.05	0.8	65	4	517	0.27	0.00	1	0.08
Authoritative teaching	0.05	0.8	65	4	517	0.27	0.00	1	0.06

4 Power Calculations

We have rich baseline data to conduct extensive power analyses. For all calculations, we assume 95% confidence intervals and 80% power. The table below presents our power calculations for minimum detectable effect sizes for our main outcomes of interest.