

---

# **STRENGTHENING TEACHER ACCOUNTABILITY TO REACH ALL STUDENTS (STARS)**

## **World Bank SIEF | Milestone 3: Observational Survey Field Report**

**Sabrin Beg, Assistant Professor of Economics, Lerner College of Business and Economics,  
University of Delaware**

**Anne Fitzpatrick, Assistant Professor of Economics, College of Liberal Arts, University of  
Massachusetts Boston**

**Adrienne Lucas, Associate Professor of Economics, Lerner College of Business and  
Economics, University of Delaware**

**Edward Tsinigo and Henry Atimone, Innovations for Poverty Action**

April, 2019

IPA Ghana



## EXECUTIVE SUMMARY

Previous research in Ghana and India demonstrated the effectiveness of "targeted instruction"—teaching students at their level of knowledge, not their grade level. The Teacher Community Assistant Initiative (TCAI) in Ghana found that this model increased learning by students despite limited teacher take-up of the program. Low adherence to this intervention, and other interventions that have been proven effective more broadly, raises challenges as governments attempt to scale-up effective interventions. One potential factor impacting low adoption rates among teachers is the lack of managerial support, which may prevent teachers from thoroughly implementing the program. The Strengthening Teaching Accountability to Reach All Students (STARS) program is designed to train teachers on targeted instruction and improve the support that head teachers and circuit supervisors provide for those teachers. Through a randomized controlled trial, this project will (1) test the effect of training teachers on targeted instruction and (2) test whether additional management training of head teachers (i.e. school principals) and circuit supervisors (i.e. middle-level management responsible for a subset of schools within a district) increases the quality of implementation of targeted instruction and student outcomes. This study works within the system to improve educational outcomes. Ghana Education Services (GES), National Council for Curriculum and Assessment (NaCCA), and the National Inspectorate Board (NIB) have designed the materials and trained the teachers, head teachers, and circuit supervisors.

In May and June 2018, prior to the end of the 2017-2018 academic year, Innovation for Poverty Action (IPA) enumeration teams visited 210 schools, interviewing 209 HT and 671 P4 through P6 teachers and interviewing and testing 5,893 P4 and P5 pupils in both English and math. The Milestone 2: Baseline Report summarizes those findings.

In August 2018, existing GES personnel trained teachers, head teachers, and circuit supervisors. In November 2018 during Term 1 of the 2018-2019 school year, IPA enumeration teams conducted an unannounced observational study, i.e. a spot check, in the study schools. This report summarizes the findings.

The intervention did not change the likelihood that students, teachers, or head teachers were present at school prior to the team's arrival. It did increase the likelihood (by approximately 20 percentage points, 44 percent) that the head teachers were present by the start of the school day. The intervention both increased the likelihood that teachers were in their classrooms for the duration of the class and engaged with students by about 10 percentage points. Specific to the pedagogical recommendations of targeted instruction, students were divided by levels instead of grade in over 57 percent of intervention schools. The intervention changed the frequency and nature of classroom observations. Teachers reported that their circuit supervisors were more likely to have conducted a classroom observation, that conditional on being observed the teachers were more likely to receive feedback from both the head teacher and circuit supervisors, and that conditional on receiving feedback it was more likely to be useful. Head teachers also reported more visits by circuit supervisors.

Across all of the outcomes presented, in many cases the point values for the intervention that included management training were larger, but never statistically significantly so. The final analysis will include an additional observational study, improving our statistical power, potentially allowing us to distinguish statistically between the two intervention arms.

Nothing in the observational study data collection or analysis portends issues in the research design or future data collection.

# TABLE OF CONTENTS

<b>1</b>	<b>INTRODUCTION .....</b>	<b>1</b>
<b>2</b>	<b>SURVEY METHODOLOGY .....</b>	<b>1</b>
2.1	Sample Design .....	1
2.2	Observation Instruments .....	2
2.3	Ethical Protocol .....	2
2.4	Data Collection Method .....	3
2.4.1	Arrival Survey .....	3
2.4.2	Classroom Observation Instrument .....	3
2.4.3	Roll Call Attendance .....	4
2.4.4	Counting Pupils' Sheet .....	4
2.4.5	Learning Progress Sheet .....	4
2.4.6	Surveys .....	4
2.5	Field staff recruitment and training .....	5
2.5.1	Recruitment of fieldwork staff .....	5
2.5.2	Training of field staff .....	5
2.6	Fieldwork and Logistics .....	5
2.7	Fieldwork Quality Control Measures .....	5
2.8	Data Management, Editing, and Analysis .....	6
<b>3</b>	<b>SAMPLE COVERAGE AND OBSERVATIONAL SURVEY FINDINGS .....</b>	<b>6</b>
3.1	Results of Interviews and Observations .....	6
3.2	Findings from the Observational Survey .....	7
3.2.1	Effects on Attendance .....	7
3.2.2	Effects on Classroom Activities .....	8
3.2.3	Teaching Support and Management .....	8
<b>4</b>	<b>CHALLENGES AND LESSONS LEARNED .....</b>	<b>9</b>
<b>5</b>	<b>NEXT STEPS .....</b>	<b>11</b>
	<b>References .....</b>	<b>12</b>

## LIST OF TABLES

Table 1. Contents of the Observational Survey Instruments .....	2
Table 2. Coverage Rates for the Observational Survey.....	6
Table 3. Effects of STARS Intervention on Attendance .....	7
Table 4. Classroom Observation .....	8
Table 5. Program Impact on Teaching Support and Management .....	10
Table 6. Timing of Future Activities.....	11

---

# 1 INTRODUCTION

Previous small-scale studies in Ghana and India demonstrated that targeted instruction - dividing students by learning level, as opposed to by grade for an hour or more per day - is a promising and cost-effective way to improve student learning outcomes (Banerjee et al. 2007; Banerjee et al. 2017; Duflo, Kiessel, and Lucas 2019; Muralidharan et al. 2019).<sup>1</sup> Yet, once an intervention has been tested and known to be effective, how can governments effectively implement it, especially at scale? With any public sector reform, the challenge is to train, support, motivate, and monitor front-line civil servants (i.e. teachers) to embrace necessary changes. With this specific reform—targeted instruction, previous research in Ghana found that only 15 percent of teachers trained in the method correctly implemented it during unannounced spot-checks (Duflo, Kiessel, and Lucas 2019).

STARS is a randomized intervention aimed at evaluating the scale-up of targeted instruction in primary schools in 20 districts of Ghana. STARS has three primary objectives: (a) testing whether targeted instruction improves student learning in upper primary grades; (b) evaluating whether a management training program increases the adoption of targeted instruction in schools; (c) assessing if a management training program in addition to targeted instruction additionally increases student learning gains. To achieve these objectives, this study randomly divided 210 schools in Ghana into three groups: targeted instruction (TI); TI plus enhanced management training; and business-as-usual (Control). Students were assessed during a baseline exam in May-June 2018, and the follow-up round of data collection that includes achievement exams will be conducted in May-June 2020. The STARS project is a partnership between the Ghanaian Ministry of Education (including its National Inspectorate Board, National Teaching Council, and National Council for Curriculum and Assessment), Ghana Education Service, United National International Children’s Emergency Fund (UNICEF), and Innovations for Poverty Action.

Prior to the data collection round described in this report, the research team collected baseline data in May and June of 2018. Teacher, head teacher, and circuit supervisor training occurred in August of 2018, prior to the start of the 2018-2019 academic year in September.

This field report is based on the Observational Study conducted during Term 1 of the academic year 2018-2019 by Innovations for Poverty Action (IPA). The Observational Study was designed and conducted to observe the implementation of the targeted instruction pedagogy, measure student, teacher, and head teacher attendance, and collect data on intermediate outcomes relevant to the study objectives. To accurately measure attendance and program implementation, respondents were not told which days IPA would be visiting their school. Data from this data collection effort provides evidence essential for evaluating the success of the targeted instruction scale-up and developing evidence-based policies toward improving the quality of basic school education in Ghana and beyond.

---

## 2 SURVEY METHODOLOGY

### 2.1 SAMPLE DESIGN

The study sample is 210 randomly selected schools within all circuits across the 20 UNICEF districts (N=140). Schools were randomly assigned to be Control; TI-Only; TI+Management. Within all randomly

---

<sup>1</sup> Duflo et al. 2011 in Kenya was a variation on this in which students remained in their same class level, but divided by baseline learning achievement for the entire academic year. Teachers taught as they wished without additional targeted instruction support.

selected schools, we randomly selected 15 P4 and 15 P5 students during the baseline survey in May-June 2018. These students are expected to be in P5 and P6 at the time of this observational study. For a complete description of the sampling strategy, please see the baseline report (Beg et al., 2018).

## 2.2 OBSERVATION INSTRUMENTS

We used eight questionnaires for the Observational Study: Arrival Survey, Classroom Observational Tool, Roll Call Tool, Pupil Counting Tool, Learning Progress Tool, Teacher Survey, Head Teacher Survey, and Circuit Supervisor Survey. The contents of these instruments are shown in Table 1. All questionnaires were administered in the English language. The instruments were piloted in two schools<sup>2</sup> in the Asikuma-Odoben-Brakwa (AOB) district of the Central Region on the 19<sup>th</sup> of September 2018. The pilot exercise involved interviewing five teachers and two headmasters. Classroom observations were carried out in five different classrooms. Based on the results of the piloting, modifications were made to the wording and content of the questionnaires to ensure appropriate content and length.

**Table 1. Contents of the Observational Survey Instruments**

Instrument	Modules	Administration Mode
Arrival Survey	Background information and teacher roster.	In-person interview
Classroom Observation	Classroom processes and practices, teacher-child interaction, and student behavior.	Observation
Roll Call Tool	Teacher and pupil information, pupil grouping (by class or level), and pupil-level attendance.	Roll call
Pupil Counting Tool	Count of pupils in levels and class	Roll call
Learning Progress Tool	Pupil information, pupil levels in English and maths	Documentary review
Teacher Survey	Background characteristics, teacher supervision, support, teacher satisfaction, work stress/burnout, and implementation of targeted instruction pedagogy.	In-person interview
Head Teacher Survey	Background characteristics, school characteristics, and activities, participation in professional development activities, perceptions about their role, work stress/burnout, technology use, and implementation of targeted instruction pedagogy,	In-person interview
Circuit Supervisor Survey	Background characteristics; management and supervision; participation in professional development activities; perceptions about the role as a circuit supervisor; work stress and burnout; and technology use.	Phone or in-person interview

## 2.3 ETHICAL PROTOCOL

IPA institutional Review Board approved the protocol and questionnaires in October 2018 for the Observational Study. Verbal consent was obtained for each respondent participating. All respondents were informed of the voluntary nature of participation and the confidentiality and anonymity of

<sup>2</sup> These schools were included in the piloting of the STARS' implementation activities.

information. Each participating teacher received GH¢ 5 airtime while the head teachers and circuit supervisors received GH¢ 10 each.

## 2.4 DATA COLLECTION METHOD

The Observational Study used Computer-Assisted Personal Interviewing (CAPI) and Samsung tablets. The data collection application was based on the SurveyCTO software, incorporating IPA's data management system. IPA Ghana's research quality team programmed the questionnaires. The programmed instruments included constraints, skip patterns, relevance commands to automate the administration process and automatically check inconsistencies or errors associated with the administration of the instruments on the field. The CAPI application was bench-tested during the training field practice. Modifications of the questionnaires based on the pilot and field practice were incorporated into the electronic versions of the questionnaires.

The data collection procedures used during the Observational Study are described below.

### 2.4.1 ARRIVAL SURVEY

Upon arriving at the school, enumerators administered informed consent to a responsible party, most often the head teacher, to conduct the day's activities. Upon receiving consent, enumerators then conducted the Arrival Survey. The Arrival Survey contained a roster of all teachers surveyed at the baseline and also new teachers currently working in the school. In addition, information was recorded for each teacher on whether the teacher was currently present, whether the teacher was P4, P5, or P6 teacher, and other information needed for the enumerator to conduct the classroom observations. The head teacher (or alternate responsible party, if the head teacher was unable to talk to the enumerator) then guided enumerators to the teachers and their respective classrooms (P4 to P6) to be observed.

### 2.4.2 CLASSROOM OBSERVATION INSTRUMENT

Immediately after the Arrival Survey, the enumerators observed each classroom (P4, P5, and P6) in each school, during the first two hours of the school day. The classroom observation was administered twice, once during the first hour of school and once during the second hour. During these observations, the enumerator peeked into the classroom and took 5-10 minutes to complete a form on what behavior was occurring in the classroom. These data can be considered as recording a "snapshot" of what is happening in the classroom.<sup>3</sup> We use this data to measure teacher attendance and behavior (subject being covered, teacher and learning materials being used, activities or tasks being carried out), classroom interactions, and students' behavior. It also included a series of questions aimed at measuring targeted instruction implementation.

We conducted all school observations during the first 2 hours when targeted instruction was most likely to take place. During the targeted instruction training, schools were advised to dedicate the first hour of class time (i.e., periods 1 and 2) to targeted instruction lessons. Schools were given leeway to change the time of day for targeted instruction to fit their specific needs and conditions. These data will, therefore, allow us to measure whether targeted instruction lessons were occurring during the first two

---

<sup>3</sup> Classroom observations were done for the stream that was randomly selected at baseline. For schools that changed or added a stream between baseline and observational study, the evaluation team randomly selected a stream to observe.

hours, and accurately compare teacher attendance and pedagogy at the same time of day at all schools. Our approach also allows us to estimate whether teacher pedagogy is different between treatment and control schools during non-targeted instruction periods. By only observing schools during the first two hours, we have data for all schools are approximately the same time of day, but we might underestimate program implementation. Some head teachers indicated that they implemented TI at a different time of day.

### 2.4.3 ROLL CALL ATTENDANCE

At the end of the first round of classroom observations, the enumerators entered each classroom and did a roll call of students who were enrolled in the study at baseline. To complete the Roll Call instrument, the enumerator entered the classroom and called out the name of each student from the STARS student roster after obtaining the teacher's permission. Enumerators recorded whether that student was present in the given teacher's classroom and then proceeded to the next classroom. In treatment schools, the enumerators also asked the teacher if the pupils were grouped by class or level. After all roll calls were completed, enumerators then proceeded with the next round of classroom observations. This roll call would be our primary measure of pupil attendance. Also, we can use it to determine whether students were actually divided by levels or grade level since we know the students' grade level at baseline and the grade or level that the teacher is teaching during the observation study visit.

### 2.4.4 COUNTING PUPILS' SHEET

At the end of the second set of classroom observations, the enumerators counted the number of boys and girls present in each classroom.

### 2.4.5 LEARNING PROGRESS SHEET

TI requires that teachers administer a short leveling exam to each student to determine her learning level. The school's Learning Progress Sheet should record the learning levels, one through three, for all P4-P6 students separately for math and English. The enumerators copied the records of all sampled students from the Learning Progress Sheet into SurveyCTO in TI schools. We will additionally use these data as a measure of TI implementation.

### 2.4.6 SURVEYS

After the completion of the above activities, enumerators administered Teacher Survey and Head Teacher Survey to the observed teachers and head teachers, respectively. Team leaders then contacted the circuit supervisor of the schools and administered a Circuit Supervisor Survey. The Circuit Supervisor Survey was conducted either via phone or in person. Absent teachers and head teachers were revisited during a "mop up" visit. Only one mop-up attempt was made per school.

## 2.5 FIELD STAFF RECRUITMENT AND TRAINING

### 2.5.1 RECRUITMENT OF FIELDWORK STAFF

Field staff was recruited based on IPA's short-term recruitment policy, which followed a competitive process. The recruitment of field staff for the Observational Survey was based on the following selection criteria: a minimum of a diploma certificate in any related field; at least a one-year experience in school-based data collection; experience in the use of CAPI; experience managing fieldwork teams [team leaders and supervisors only]; and working knowledge about the 20 UNICEF districts. Thirty percent of the field staff was shortlisted for training to account for attrition and ensure the hiring of qualified candidates from the pool of trainees.

### 2.5.2 TRAINING OF FIELD STAFF

Field staff training centered on the questionnaires, observation protocols, and use of CAPI for data collection. The training was non-residential and was conducted in the IPA Office, Abelemkpe from 10<sup>th</sup> to 17<sup>th</sup> October 2018. The Research Associate and Field Manager led the training. The training sessions included presentations, questions and answers, group discussions, role-play, and field practice.

The evaluation team conducted field practice in seven implementation pilot schools in Asikuma-Odoben-Brakwa District to help the trainees apply what they learned during the training, and for the trainers to provide them with specific and constructive feedback during a debriefing session. The field practice was conducted on 16<sup>th</sup> October 2018. Twenty-six field staff participated in the field practice. The field practice involved observing 7 classrooms and interviewing 21 teachers and 7 head teachers.

Trainees' performance was evaluated to gauge their progress; provide performance feedback to both the trainers and the trainees; as well as to select the best candidates for the Observational Study. Trainees' performance was assessed using quizzes, role-plays, and field practices or school visits.

## 2.6 FIELDWORK AND LOGISTICS

Four survey teams [made up of 15 enumerators, 4 team leaders, and 1 auditor] conducted the Observational Study. Data were collected during the first term of the 2018-2019 academic year, from 29<sup>th</sup> October 2018 to 30<sup>th</sup> November 2018. Data collection activities were conducted concurrently in both the treatment and control schools. Logistics for data collection activities included Samsung tablets, netbooks for team leaders, answer aids, and backpacks for the field staff. The rest were life jackets, first aid kits, hard copies of instruments and survey manual.

## 2.7 FIELDWORK QUALITY CONTROL MEASURES

Data collection activities were monitored to assess the (a) performance of the fieldwork teams in administering the various instruments and (b) quality of the data being collected. Field teams were monitored using IPA's standardized monitoring tool, hosted on SurveyCTO. On average, at least two different monitors monitored each field staff during the data collection period. The results from the monitoring largely showed the field staff's strict adherence to the established protocols. This was partly due to the use of experienced enumerators, feedback-based training, and the provision of timely feedback to the fieldwork team.

## 2.8 DATA MANAGEMENT, EDITING, AND ANALYSIS

A standardized and coordinated system of checks and systems were developed and implemented in managing the data flow, collection, cleaning, and storage. These coordinated systems helped to ensure the accessibility, quality, reliability, and timeliness of the data. Data were encrypted using BoxCryptor from the point of collection to storage. Using IPA's Data Management System, high-frequency checks were run daily to identify inconsistencies, electronic programming errors, and enumerator errors. The high-frequency checks indicated a minimal violation of the data quality checks such as duplicate IDs, missing values, constraints, skip patterns and survey logic or inconsistencies. Also, 10% each of the completed Head Teacher Survey, Teacher Survey, and Circuit Supervisor Survey were audited to establish whether there were variations in key outcome variables. The audit checks showed that discrepancies were largely within the acceptable range. During data collection and following the completion of fieldwork, data were edited and cleaned using STATA do-files.

---

## 3 SAMPLE COVERAGE AND OBSERVATIONAL SURVEY FINDINGS

### 3.1 RESULTS OF INTERVIEWS AND OBSERVATIONS

Table 2 presents the results of the sample implementation, including response rates. Of the 210 study schools, 208 head teachers [representing 99%] were successfully interviewed. Within the 210 schools, 65% of baseline teachers (n = 436 were successfully interviewed). Due to additional teachers currently working in study schools, we conducted interviews in total with 598 teachers out of a total of 630 teachers sampled and interviewed at baseline. Relatedly, 95% of the circuit supervisors were successfully interviewed during the Observational Study. Less than 100% coverage was due to refusals and the non-availability of the target respondents for surveys due to transfers, ill-health, and maternity or study leave. Only 436 of teachers from baseline were still working in study schools. In addition, the classroom observations and roll call were only conducted in 209 schools because classes were not being held the day of our spot check.

**Table 2. Coverage Rates for the Observational Survey**

Instruments	Target	Completed	Not Completed	% Coverage
School Arrival Survey	210	210	0	100%
Circuit Supervisor Survey	145	137	8	95%
Head Teacher Survey	210	208	2	99%
Teacher Survey	630	598 <sup>4</sup>	32	94.9%
Roll Call Survey	5893	5863	30	99.5%
Learning Progress Sheet <sup>5</sup>	3895	3895	0	100%
Classroom Observation <sup>6</sup>	1260	1228	32	97%

---

<sup>4</sup> Of the 671 teachers surveyed at baseline, 630 were sampled for the study. From this, 436 remained in the study during the Observational Study. Additionally, 162 teachers were recruited and interviewed as replacement for those who GES transferred to other schools.

<sup>5</sup> Learning progress sheets were only obtained for students in the intervention schools.

<sup>6</sup> Expected and actual teacher-level frequencies reported.

## 3.2 FINDINGS FROM THE OBSERVATIONAL SURVEY

### 3.2.1 EFFECTS ON ATTENDANCE

In Table 3 we present the effects of the intervention on the student, teacher, and head-teacher attendance. We find that there is no difference in student (Column 1) and teacher attendance (Column 2). We further find no difference by treatment status of whether the baseline teacher is still employed at the school (Column 3).

Head teachers across all treatment groups were equally likely to be present when the enumeration team arrived prior to the start of the school day. Note that this was only approximately 21 percent of head teachers (Column 4). By the official start time of school, an additional 20 percentage points of the control group head teachers had arrived (41 percent present) with even more of the treatment head teachers arriving (Column 5). We find that the treatment increased the likelihood that head teachers were present at the start of school by 18 percentage points for TI only and 22 percentage points for TI + Management. While numerically different, these values are statistically indistinguishable.

**Table 3. Effects of STARS Intervention on Attendance**

	Student Attendance	All Teachers: Present on School Grounds?	Baseline Teachers: Still Working At School?	Current Head Teacher Present at Arrival	Current Head Teacher Present at Official Start Time
	(1)	(2)	(3)	(4)	(5)
TI Only	-0.009	0.035	-0.002	0.027	0.180**
	(0.014)	(0.044)	(0.037)	(0.064)	(0.078)
TI+Management	-0.007	0.033	0.008	0.098	0.215**
	(0.016)	(0.051)	(0.034)	(0.072)	(0.084)
Observations	5863	752	671	210	210
R-squared	0.02	0.11	0.07	0.24	0.20
Pvalue of TI-Only =TI+Management	0.90	0.95	0.78	0.31	0.69
Pvalue Treatment Jointly = 0	0.81	0.71	0.95	0.38	0.02
Mean of Dep. Control	0.85	0.69	0.81	0.21	0.41

Notes: Data from the first observational study. Columns 1, 4, 5: Linear Probability Models. Column 1: includes controls for student age, student age-squared, a dummy variable for whether the student is female, baseline ability, baseline grade, and a district fixed effect. Columns 2-3: include a control for teacher's class, age, age-squared, gender, years of experience, and years of experience-squared. Column 2: all teachers currently employed by the school. Column 3: all teachers employed by the schools during the baseline survey. Columns 4-5: include a control for age, age-squared, gender, years of experience, years of experience-squared. Column 4 also includes a control for whether the school was open at the time of arrival. Robust standard errors in parentheses, clustered at the school level. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

### 3.2.2 EFFECTS ON CLASSROOM ACTIVITIES

In Table 4 we explore ways in which STARS changed the classroom environment. STARS increased at least two desirable teaching characteristics. First, teachers were about 11 percentage points more likely to be in the classroom for the duration of the lesson (Column 1). This is a 15 percent increase over the control group mean of 72 percent. Second, teachers were about 10 percentage points more likely to be engaged with students (Column 2), a 16 percent increase over the control group mean of 64 percent.<sup>7</sup> In our final column of Table 4, we show that most treatment schools were implementing targeted instruction—about 57 percent of schools had divided students by learning levels instead of grade level (Column 3). For all of these outcomes, the TI and TI+Management interventions result in statistically indistinguishable differences.

**Table 4. Classroom Observation**

	Teacher in classroom for duration of class	Teacher engaged with students	Students divided into levels
	(1)	(2)	(3)
TI Only	0.124***	0.112***	0.573***
	(0.037)	(0.037)	(0.058)
TI+Management	0.109***	0.097**	0.599***
	(0.040)	(0.041)	(0.057)
Observations	1,228	1,224	209
R-squared	0.08	0.08	0.48
Pvalue of TI-Only =TI+Management	0.66	0.67	0.72
Pvalue Treatment Jointly = 0	0.00	0.01	0.00
Mean of Dep Control	0.72	0.64	0.00

Notes: Data from the first observational study. Includes all teachers. Binary outcomes estimated using a linear probability model. All regressions include a district fixed effect, an indicator for whether or not it was the first or second observation of the day, and control for baseline teacher attendance at the school level. Robust standard errors in parentheses, clustered at the school level. One school was not holding classes during the unannounced spot check and so there is no classroom observation or roll call data available for that school. Column 2: There are four fewer observations than in Column 1 due to special situations as described in the text. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

### 3.2.3 TEACHING SUPPORT AND MANAGEMENT

In Table 5 we present evidence that the intervention improved both the number and quality of interactions between teachers, head teachers, and circuit supervisors. On average teachers reported being visited by teachers about 5 times in this term, with no difference by treatment status (Column 1). In contrast, both treatment groups report increases of 0.77 (TI only) and 0.93 (TI plus management) additional classroom observations by their Circuit Supervisor. Compared to the Control group average

<sup>7</sup> This variable is missing four observations because in addition to the expected teacher a second teacher was in the classroom. This second teacher was acting as the classroom teacher and the enumerator left the question blank.

of 1 visit, this difference translates into a 78-93 percent increase. While the point value is larger for the management arm, we cannot reject their statistical equality.

Conditional on having any classroom observations conducted, teacher respondents also report that both their head teachers and their circuit supervisors are more likely to give feedback on their teaching. Head teachers are 10-18 percentage points more likely to give feedback while circuit supervisors are 16.8-23.1 percentage points more likely to give feedback with the larger point values for the intervention that included management training. Conditional on receiving feedback, teachers in the TI only arm are also 7.5 percentage points more likely to report that their head teacher's feedback was useful.

Head teachers also report a statistically significant increase in visits by their circuit supervisor - 0.9 for the TI only group and 1.2 for the TI+Management group. These increases are not statistically different from one another. They are increases of over 30 percent relative to the control group mean of 2.7.

---

## 4 CHALLENGES AND LESSONS LEARNED

The Observational Study was conducted amidst several challenges. These challenges have provided useful lessons to the project management team in planning for future data collection activities. The major challenges and lessons learned were:

- a. *Transportation challenges.* Lack of reliable means of transportation to the schools affected the data collection activities. This increased the cost of transporting field staff to remote areas to conduct observation activities. Incentivizing field staff with an additional amount for transportation would help to motivate them to travel to even remote areas for surveys.
- b. *Accommodation cost.* Considering that only one field staff was assigned to a district, it was difficult to fully pay for the high cost of accommodation in some of the areas; accommodation costs can be shared when there are multiple field staff in the area. Future data collection activities should pay the additional cost of accommodation for individuals who work in a district.
- c. *Poor network connectivity.* Due to the poor network connectivity, the field team could not easily reach some of the respondents for a phone interview, and communicating about the study to teams was difficult. The use of text messages was very helpful in providing information as the recipients can receive text messages when they reach a coverage area.
- d. *Large class size in some schools.* The large class size affected the ability of the observers to quickly observe some of the variables (such as a number of students) in the class in the observation instruments.
- e. *Poor classroom setting affects classroom observation and counting pupils.* For classrooms that lack basic facilities such as tables and chairs, the disorganized nature of the sitting arrangement tends to affect students' concentration. This disorganization may affect the observer's assessment of the teacher's classroom management performance and makes counting the number of pupils in a classroom more difficult. To ensure accurate reportage in such cases, the observer needs to document not only his/her assessment of the teacher but also highlight the particular condition under which such an observation was made.
- f. *Teacher absenteeism.* The unannounced nature of the Observational Study meant that respondents were not pre-informed and often absent. This affected the survey team in getting the necessary materials such as the leveling sheets.

**Table 5. Program Impact on Teaching Support and Management**

Sample:	All Teachers		Teachers Who Were Observed by Supervisor		Teachers Observed with Feedback Given by Supervisor		Head Teachers
	Number of Classroom Observations by		Was Feedback Given by		Feedback Given was Useful		Number of CS Visits
	Head Teacher	Circuit Supervisor	Head Teacher	Circuit Supervisor	Head Teacher	Circuit Supervisor	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
TI Only	1.099	0.772***	0.098*	0.166***	0.075*	0.057	0.867***
	(0.872)	(0.218)	(0.057)	(0.062)	(0.041)	(0.056)	(0.322)
TI+Management	0.257	0.930***	0.177***	0.229***	0.013	0.013	1.203***
	(0.700)	(0.258)	(0.053)	(0.062)	(0.045)	(0.053)	(0.350)
Observations	425	435	386	312	295	248	208
R-squared	0.17	0.20	0.11	0.14	0.16	0.19	0.32
Pvalue of TI-Only =TI+Management	0.28	0.58	0.15	0.17	0.15	0.41	0.37
Pvalue Treatment Jointly = 0	0.43	0.00	0.00	0.00	0.14	0.57	0.00
Mean of Dep Control	4.66	1.00	0.65	0.64	0.89	0.85	2.73
<p>Notes: Data from teacher and head teacher respondents during the first observational study. Sample restricted to all teachers from the baseline survey. All regressions include a district fixed effects. Observations differ according to the sample listed and exclude missing values (don't know/refusal). Columns 1-6: include a control for teacher's class, age, age-squared, sex, years of experience, years of experience-squared, and a dummy indicating the teacher was present at baseline. Columns 3-6: estimated using a linear probability model. Column 7: include controls for age, age-squared, sex, years of experience, and years of experience squared. Robust standard errors in parentheses, clustered at the school level. *** p&lt;0.01, ** p&lt;0.05, * p&lt;0.1</p>							

- g. *Language barrier.* The difficulty of some of the survey teams to understand the local language of instruction during the observational periods affected their documentation of some issues in the observational instrument. Future observations should recruit field staff with diverse local languages that pertain to the project catchment areas.

---

## 5 NEXT STEPS

Table 6 detailed the major activities to be undertaken in 2019. The evaluation team has conducted a second Observational Study and is preparing to launch a Follow-up. Data analysis will be done by December 2019.

**Table 6. Timing of Future Activities**

Activities	Timing	Status
Observational Study II	February - March 2019	Completed
First Follow Up Study	May - July 2019	Preparing to launch
Qualitative Research	October - November 2019	Scheduled
Data cleaning, analysis, and reporting	July - December 2019	Scheduled

---

## REFERENCES

- Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *Quarterly Journal of Economics* 122 (3), 1235–1264.
- Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M., and Walton, M. (2017). From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application. *The Journal of Economic Perspectives*, 31(4):73-102.
- Beg, S., Fitzpatrick, A., Luca, A., Comba, R., Atimone, H., Gyamfi, B. K., Jumpah, J., and Tsinigo, E. (2018). Strengthening Teachers Accountability to Reach all Students (STARS). World Bank SIEF: Milestone 2: Baseline Report.
- Duflo, E., Dupas, P., & Kremer, M. R. (2011). Peer effects, teacher incentives, and the impact of tracking: evidence from a randomized evaluation in Kenya. *American Economic Review* 101 (5), 1739–1774.
- Muralidharan, Karthik, Abhijeet Singh and Alejandro J. Ganimian. 2019. "Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India." *American Economic Review* 109 (4): 1426-60.