



Quantitative assessment of teacher motivation, classroom practices, and student learning

NOVEMBER 2015

Prepared by:

Ronald Abraham
Stuart Shirrell¹
Harlan Downs-Tepper
Varun Chakravarthy

¹ Please direct all correspondence regarding this report to stuart.shirrell@idinsight.org.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	4
EXECUTIVE SUMMARY	5
BACKGROUND	7
STIR Education	7
Evaluation Design	7
Limitations and Threats to the Evaluation	10
BASELINE RESULTS AND DISCUSSION	12
Teacher Motivation	12
Classroom Observation	15
Child-Friendliness of Classrooms	23
Student Activities	24
Linking teacher and student activities	29
Student Learning	31
RECOMMENDATIONS	34
APPENDIX A: TEACHER MOTIVATION SURVEY	36
APPENDIX C: CLASSROOM OBSERVATION TOOL	40
APPENDIX C: DATA QUALITY STANDARDS	56

LIST OF FIGURES

FIGURE 1: STIR'S THEORY OF CHANGE AND DATA COLLECTION INSTRUMENTS	9
FIGURE 2: TEACHER MOTIVATION INDEX BY GEOGRAPHY	13
FIGURE 3: TEACHER MOTIVATION CATEGORIES BY IMPORTANCE AND TEACHER EXPERIENCE	14
FIGURE 4: TEACHING ACTIVITY BY GRADE AND GEOGRAPHY	18
FIGURE 5: TEACHER ACTIVITIES BY DAY OF THE WEEK: DELHI AND U.P.	19
FIGURE 6: TEACHER ACTIVITIES BY CLASSROOM SIZE: DELHI AND U.P.	20
FIGURE 7: TEACHER ACTIVITIES BY TIME OF THE DAY: DELHI AND U.P.	22
FIGURE 8: STUDENT ACTIVITIES BY GRADE AND GEOGRAPHY	26
FIGURE 9: FRACTION OF STUDENTS ATTENTIVE IN THE CLASSROOM BY TIME OF DAY	28
FIGURE 10: FRACTION OF STUDENTS ENGAGED IN DIFFERENT ACTIVITIES CONDITIONAL ON TEACHER ACTIVITIES	30
FIGURE 11: MATH LEVELS BY GRADE AND GEOGRAPHY	32
FIGURE 12: HINDI READING LEVELS BY GRADE AND GEOGRAPHY	34

LIST OF TABLES

TABLE 1: TEACHER CLASSROOM ACTIVITY	16
TABLE 2: ASER CHILD FRIENDLINESS MATRIX SUMMARY STATISTICS	24
TABLE 4: MATHEMATICS LEARNING LEVELS BY GEOGRAPHY	32
TABLE 5: HINDI LEARNING LEVELS BY GEOGRAPHY	33

ACKNOWLEDGMENTS

A large number of people contributed to the creation of this report, directly and indirectly. STIR Education has been very helpful and patient throughout the evaluation design and baseline survey rounds. The World Bank's Strategic Impact Evaluation Fund (SIEF) provided funding for the evaluation. A number of people at the ASER Center have provided useful feedback throughout the process. Morsel provided human-resource and logistical support for the teacher motivation baseline survey in Uttar Pradesh.

EXECUTIVE SUMMARY

This report summarizes the findings from the baseline survey of the impact evaluation conducted by IDinsight² for STIR Education³ in Delhi and Uttar Pradesh in India, funded by a World Bank Strategic Impact Evaluation Fund (SIEF) grant.⁴ STIR works with teachers in low-cost private and government schools in order to improve student learning by empowering teachers to act as change-makers and to innovate to overcome challenges in the classroom.

This study seeks to evaluate the impact of STIR's purely motivational, pedagogically neutral, teacher-focused model on the student learning levels. IDinsight is conducting two three-armed randomized control trials. The study will look at outcomes from 180 Affordable Private Schools (APS) in Delhi and 270 government schools in Rae Bareilly and Varanasi districts of Uttar Pradesh. The study began in early 2015, and is currently planned for two academic years. In addition to measuring STIR's impact in two different contexts, the study will simultaneously test two iterations of STIR's model in these two contexts.

The baseline survey collected information on teacher motivation levels, student learning levels, teachers' and students' activities in the classroom, verbal communication between teachers and students, and the level of teaching content. In addition to validating the randomized evaluation design, the findings in this report are an effort to diagnose problems on the ground and help STIR focus its efforts on key problem areas.

KEY FINDINGS⁵

The key findings highlighted by this report are as follows:

- **Student Learning:** There is steady increase in student learning from one grade to the next. The proportion of students who complete the entire Hindi section of the student learning test increases from less than 1% in grade 1, to 20% in grade 4, and finally to 69% in grade 8 for Delhi. At the same grade levels the proportion of students in U.P who complete the entire Hindi section is less than 1%, 7%, and 48%. A large fraction of students in higher grades do not appear to progress past the early rounds of the test, a trend that is more pronounced in U.P. For example, nearly a quarter of eighth standard students in U.P. government schools cannot even read a short Hindi sentence. Students in Delhi also outperform their U.P.

² IDinsight : <http://www.idinsight.org/>

³ STIR Education: <http://www.stireducation.org/>

⁴ SIEF Grant: The Strategic Impact Evaluation Fund (SIEF) is a trust fund established within the World Bank to carry out and support research evaluating the impact of programs on alleviating poverty and improving people's lives. <http://www.worldbank.org/en/programs/sief-trust-fund>

⁵ This baseline does *not* tell us if STIRs model is or will in the future have an impact. Impact will be measured by taking a difference between treatment and control group teachers and students at the time of the midline or endline survey. *None of the analysis presented here should be mistakenly interpreted as the impact of STIRs model but rather should be thought of as a pre-implementation descriptive and diagnostic analysis.*

counterparts from the same grade in both Hindi reading and mathematics; these differences are highly statistically significant.^{6,7} For example, 76% of first standard students in Delhi APS schools know at least two-digit numbers, compared with 1% of first standard students in U.P. government schools. Similarly, 68% of fourth graders in Delhi APS schools know at least subtraction, while only 16% of fourth graders in U.P. government schools.

- **Teaching Time:** Teachers in Delhi and U.P. spend less than three quarters of their time teaching, which is lower than the best practice standards.⁸ Delhi teachers spend almost 29% of their time on average on classroom management, while teachers in U.P. spend only 7% of time on average. Teachers in U.P. spend around 16% of their time on average off task. The way teachers allocate their time does not vary significantly with day of the week, time of the day or class size.
- **Classroom Environment:** The classroom environment is more interactive in Delhi as compared to U.P. as gauged by the number of students who ask questions and participate in the classroom. At least one question was asked by students in more than 80% of classrooms observed in Delhi and less than 40% classrooms observed in U.P. There was a low incidence of group work among students in both Delhi and U.P. classrooms.
- **Student Participation:** A larger proportion of students spend time off task in U.P. classrooms. Delhi classrooms have a larger proportion of students involved in group discussions and question and answers with teachers in the classroom.

IDinsight has the following recommendations for STIR's programming going forward:

- **Teachers in U.P. should spend more time with remedial students:** A significant fraction of students in U.P. government schools are well below basic reading and math levels. STIR should consider focusing teachers towards this problem and facilitate solution-generation amongst them.
- **Teachers should spend more time teaching in both Delhi and U.P.:** Teachers in both geographies are teaching less than the recommended best practices, according to the Stallings classroom snapshot. STIR should investigate ways to reduce classroom management time in Delhi and time off task in U.P.
- **The emotional tone of the classroom can improve:** Teachers in both Delhi and U.P. scored low on the ASER child-friendliness matrix. To rectify this, STIR could look into micro-innovations that target the emotional tone of the classroom and improve the way students and teachers interact.

⁶ Saying that a result is statistically significant means that there is a low probability that the difference observed between two groups would have happened by chance. This probability is captured by a p -value, which is computed as the result of a statistical test. Statistical significance is determined by specifying a pre-determined cutoff for the probability, and then determining whether the p -value is above or below that cutoff point. Common significance levels are 1%, 5%, and 10%. In this document we report statistically significant differences as those with a significance level below 5% and highly significant differences as those with a significance level below 1%.

⁷ In this case the p -value is less than 0.01 for Hindi and math, accounting for clustering at the school level in the evaluation design.

⁸ The Stallings Snapshot tool defines ideal time allocation as 85% teaching and 15% classroom management.

BACKGROUND

STIR Education

STIR Education (Schools and Teachers Innovating for Results) is a non-governmental organization that seeks to improve student learning in developing countries by empowering teachers and principals to act as changemakers by adopting innovative teaching and management practices. STIR aims to create a movement of teachers who innovate to overcome day-to-day classroom problems, follow effective classroom practices, and influence others around them – all contributing to improved student learning levels. STIR works in Uganda and several states in India.

STIR's theory of change focuses on teacher motivation. By treating teachers as professionals, STIR believes that it can improve teacher motivation. Higher levels of motivation will increase both the quality and quantity of teaching. Students, with access to more and better teaching, will then learn more.

STIR's program begins with a search exercise where they seek both interested teachers and micro-innovations, or small changes teachers can make in their classroom to overcome some challenge. These teachers are then organized into a network, allowing teachers to work with and support their peers. Networks meet on a monthly basis, where teachers discuss challenges they face and receive support and training from an Education Leader (EL), who facilitates the meeting.

STIR currently envisions a three-year teacher changemaker journey. In the first year the network proceeds in two phases – an implementation and an influence cycle. In the implementation cycle teachers are introduced to innovative ways to solve problems. This is meant to increase teacher motivation and make them more confident of their own ability. Teachers then enter the influence stage where they introduce other non-network teachers to the STIR model. By the end of this first year STIR seeks to have a movement of motivated teachers who are confident in their ability to act as problem solvers.

In the second year STIR will implement their classroom challenge program, where they will pose larger, more open-ended classroom management problems to teachers. Teachers will also receive additional support from Education Leaders and external parties to improve their classroom management practices. STIR envisions the third year similarly, with a change in focus. Instead of improving their classroom management practices, teachers should improve reading outcomes in their classrooms.

Evaluation Design

STIR has partnered with IDinsight to undertake an evaluation of the effects of STIR's model as well as the efficiency in its implementation in Delhi and Uttar Pradesh. This report presents findings from the baseline survey of the study.

The study involves two three-arm, stratified cluster-Randomized Controlled Trials (RCT's).⁹ Each study is currently planned for two years. One study will take place in affordable private schools in east Delhi and the other will take place in government schools in Uttar Pradesh. Each study has three arms:

- **STIR 1.0, the “base” model:** In this model, STIR seeks to improve teaching and learning outcomes by making them feel part of a larger movement and by increasing their self-belief.
- **STIR 2.0, the “advanced” model:** In their advanced model, STIR adds to their base model by incorporating a structured system of recognition for the teachers who are part of their network.
- **Control group:** This is a pure control group in Uttar Pradesh and a placebo control group in Delhi.¹⁰ Neither of these groups receives the STIR program.

IDinsight is evaluating the three key components of STIR's theory of change using different evaluation mechanisms and data collection tools (see Figure 1). The levels of teachers' motivation is quantified using a specially-designed teacher motivation questionnaire; the classroom observation tool helps gauge classroom activities, teacher-student communication flow and structure and content levels of teaching; learning levels of students are captured using a student testing tool.

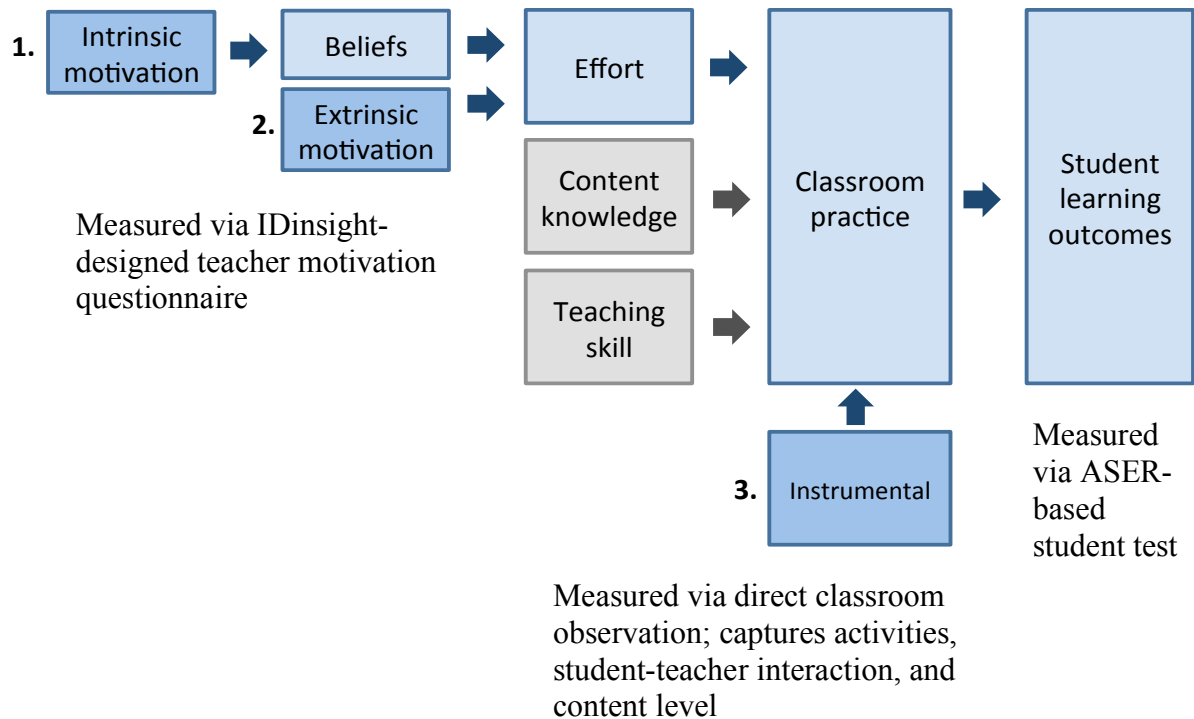
- **Teacher Motivation Survey:** The teacher motivation baseline survey (see sample tool in Appendix A) was conducted between February and April, 2015.¹¹ The teacher motivation survey used a 40-item questionnaire to gauge teachers' motivation and values along several different dimensions defined in the literature.

⁹ A randomized controlled trial is an evaluation design where participants are randomly assigned into a treatment and a control group. The only expected difference between these two groups is that created by the assignment (in the treatment group) and non- assignment of the intervention (in the control group).

¹⁰ The placebo control schools in Delhi are those schools that are part of STIRs network but the services extended to teachers in these schools are such that it does not directly or indirectly influence their teaching practices, *e.g.* a newspaper subscription for the school or yoga classes for teachers.

¹¹ The teacher motivation survey happened much earlier than the classroom observation and student testing survey in order to capture teachers' motivation levels before interacting significantly with STIR.

Figure 1: STIR's Theory of Change and Data Collection Instruments



- Classroom observation:** Between July and November 2015, IDinsight conducted the second round of surveys in Delhi and Uttar Pradesh. One of the two components of this baseline survey was to capture classroom practices of teachers. In order to quantify the classroom practices an activity-based tool was used (see sample tool in Appendix B).^{12,13,14} Additional sections were incorporated to this tool including the ASER child friendliness matrix,¹⁵ a section that

¹² The Stallings Snapshot tool was used loosely as the basis of the tool.

¹³ Stallings, J. (1977). *Learning to look: A handbook on classroom observation and teaching models*. Belmont, Calif.: Wadsworth Pub.

¹⁴ The World Bank Group (2015). *User Guide: Conducting Classroom Observations*. Retrieved November 2, 2015, from the World Bank Group's website: <http://documents.worldbank.org/curated/en/2015/07/24756165/conducting-classroom-observations-analyzing-classrooms-dynamics-instructional-time-using-stallings-classroom-snapshot-observation-system-user-guide.pdf>

¹⁵ Bhattacharjea, S., Wadhwa, W., & Banerji, R. (2011). *Inside Primary Schools: A study of teaching and learning in rural India*. Retrieved November 2, 2015, from the ASER centre's website: http://img.asecentre.org/docs/Publications/Inside_Primary_School/Report/tl_study_print_ready_version_oct_7_2011.pdf

quantifies verbal interaction within the classrooms,^{16,17,18} and a section capturing teaching content level. This helped capture a more holistic view of the activities within a classroom.¹⁹

- **Student testing:** Alongside the classroom observation survey, 10 students from each class were selected at random to assess their learning levels. This was done using an ASER-based student testing tool.^{20,21} The ASER tool has two testing sections – Hindi reading, whose difficulty ranges from letter to a story at roughly a second standard level and mathematics, which ranges in difficulty from 1-digit numbers to division with a 3-digit dividend. In order to avoid ceiling effects,²² particularly in Delhi APS schools, the tool was made more complex by adding a more complex story to the Hindi test and a fractions section to the mathematics test.

Limitations and Threats to the Evaluation

While the research team has tried to design the most rigorous and comprehensive study within STIR's operational constraints, there are limitations to this study as well as potential threats to the evaluation design:

- **Challenging metrics:** It is inherently difficult to quantify teacher motivation and tone of verbal interaction, which are qualitative in nature. While the tools used have been extensively piloted and results from the piloting have been positive, a quantitative tool may not capture the full range of possibilities for some of the outcome indicators.
- **Teacher attrition:** Due to the constant turnover of teachers in both public and private schools, there has been some attrition in both the Uttar Pradesh and Delhi studies. In Delhi, many teachers are younger women, who often leave a school for economic or personal reasons. Some teachers in Delhi also move to administrative roles or take a position at another school. In U.P., however, most attrition results from teachers transferring from one school to another.

¹⁶ The verbal interaction tool recorded the speaker and content of speech for several minutes inside the classroom. This data was then used to compute metrics such as teachers' reactions to students' questions, the relative frequency teachers praising or criticizing students, etc.

¹⁷ Amidon, E. (1966). *Interaction Analysis: Recent Developments*. Retrieved November 2, 2015 from the Education Resources Information Center website: <http://files.eric.ed.gov/fulltext/ED013776.pdf>.

¹⁸ Flanders, N. (1961). Analyzing teacher behavior as part of the teaching-learning process. *Educational Leadership*, 19, 173-180.

¹⁹ English versions of the tools used can be found in Appendix A and Appendix B.

²⁰ The ASER Centre (2015). *Hindi Testing Tool*. Retrieved November 2, 2015, from the ASER Centre's website: http://img.asercentre.org/docs/Asur_survey/Tools_Testing/Reading/indialangtt.pdf.

²¹ The ASER Centre (2015). *Mathematics Testing Tool*. Retrieved November 2, 2015, from the ASER Centre's website: http://img.asercentre.org/docs/Asur_survey/Tools_haryanamathstt.pdf.

²² Ceiling effects occur when a high proportion of students complete the entire student testing tool. In these cases, the testing tool does not capture the full range of student abilities, which may bias estimates of impact. IDinsight will revise and pilot its testing tools before the midline survey in 2016.

- **School Attrition:** Loss of schools also poses a threat to the study. In Uttar Pradesh, schools are unlikely to drop out due to high-level cooperation within the education department. In Delhi, however, there is no administrative structure covering all schools in the study. There has been incidence of attrition in some schools during the baseline survey in Delhi due to lack of interest in STIR's program and administrators' reservations about data collection.²³ This is more of a concern at midline or endline and hence it is important that STIR and IDinsight work together to ensure schools are convinced and willing to be part of both the program and the evaluation.

²³ Affordable private schools in Delhi have come under threat of government ordered closure due to inability to meet some of the quality and infrastructure standards of the Right to Education Act. See, for example, <http://www.thehindu.com/news/national/300-private-schools-under-delhi-govt-scanner-for-flouting-norms/article7670628.ece>. This may cause some schools to be more hesitant in sharing information with external parties.

BASELINE RESULTS AND DISCUSSION

Teacher Motivation

STIR's model rests on the assumption that improving teaching motivation can improve learning outcomes; this makes measuring teacher motivation a crucial indicator in the evaluation. To assess teacher motivation, IDinsight developed a 40-item questionnaire to measure different aspects of teachers' motivation.^{24,25} The responses from this questionnaire were then used to create a teacher motivation "index," summarizing an individual teachers' motivation levels. This portion of the baseline survey was conducted in February through April 2015, before STIR started implementing its program, in order to minimize the chance of STIR influencing teacher motivation before baseline measurements could be taken.

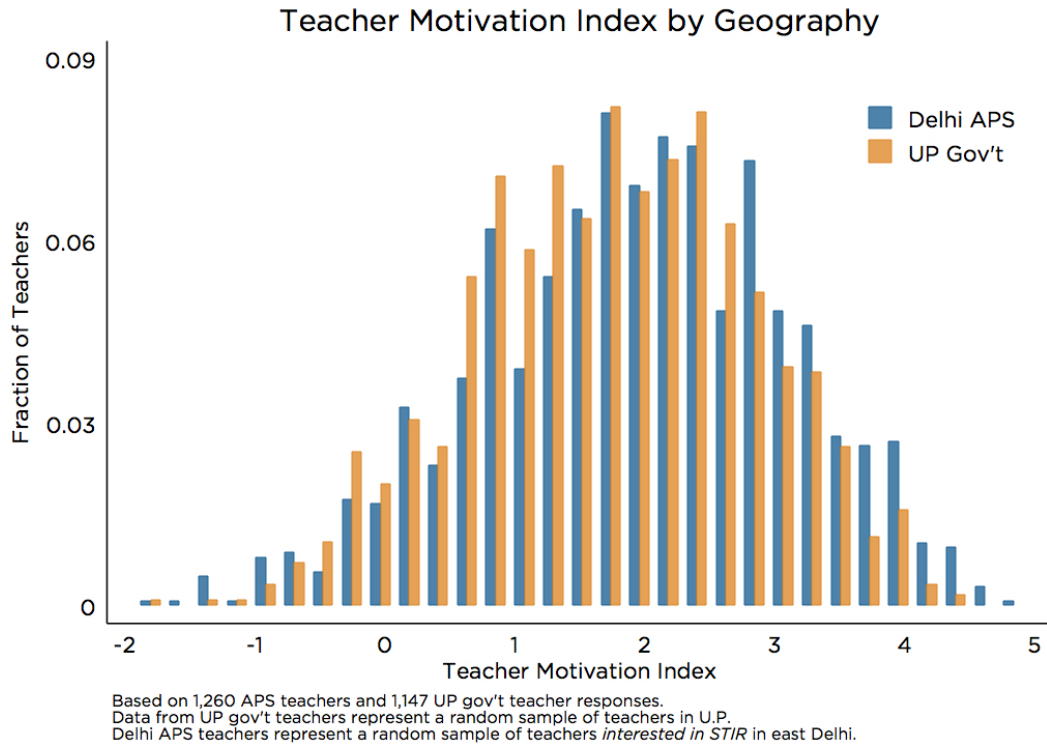
Figure 2 shows a histogram of the teacher motivation index by geography. The teacher motivation index gives a slightly higher average motivation for teachers in Delhi as for teachers in U.P. The reader should note that the teacher motivation index is somewhat difficult to interpret normatively; for example, it is difficult to know how much more motivated "2" is compared with "1."²⁶ As a result, it is likely not appropriate to interpret teachers in Delhi as being "more motivated." Differences in interpretation and culture in the APS versus the government school structure may lead to these differences, and the index is best interpreted by looking at changes within a population.

²⁴ Specifically, the questionnaire asks about ten different potential motivation categories that may influence how teachers feel about their profession. The questionnaire asks both about the teacher's current situation directly as well as hypothetical scenarios to gauge how much a teacher values a particular motivation category. For both the direct question and the hypothetical scenarios, the questionnaire asks both positive and negative questions, to pick up any differences between how the teacher values positive and negative events.

²⁵ The questionnaire was piloted extensively to test for various influences, such as stability of the index over time. The index appears to have relatively high autocorrelation, especially for a metric that is prone to short-term fluctuations.

²⁶ Combining the teacher motivation index with some form of normative motivation scale is an area of future work for the teacher motivation index.

Figure 2: Teacher Motivation Index by Geography



In addition to the entire teacher motivation index, IDinsight has investigated what might be the most fruitful teacher motivation categories for STIR to target through its program. To do this, IDinsight computed the mean importance and mean situation for each sub-category of the teacher motivation index. The mean importance was defined as the sum of the values for the two importance questions for each category.²⁷ The mean situation metric was defined as the positive situation question *minus* the negative situation question.²⁸ Note that the two measures are defined differently because negative experiences in a particular category detract from that teacher's overall experience, while placing more *value* on negative experiences contributes to the teachers' nonetheless valuing that category more.

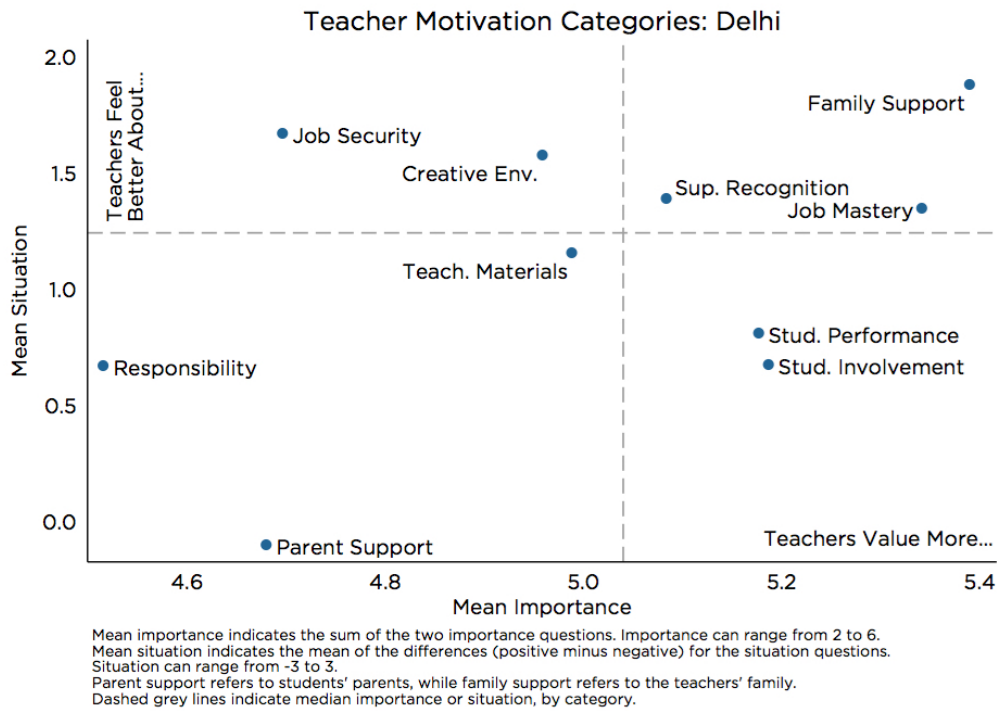
In both Delhi and U.P., teachers indicate that student academic performance and student involvement in the classroom are relatively important, but teachers rate these aspects lower, on average, than other aspects of teaching. These categories might be particularly ripe categories for STIR to focus on during its network meetings. Additionally, teachers in both Delhi and U.P. do not feel that students' parents are particularly supportive. While they place less emphasis on this

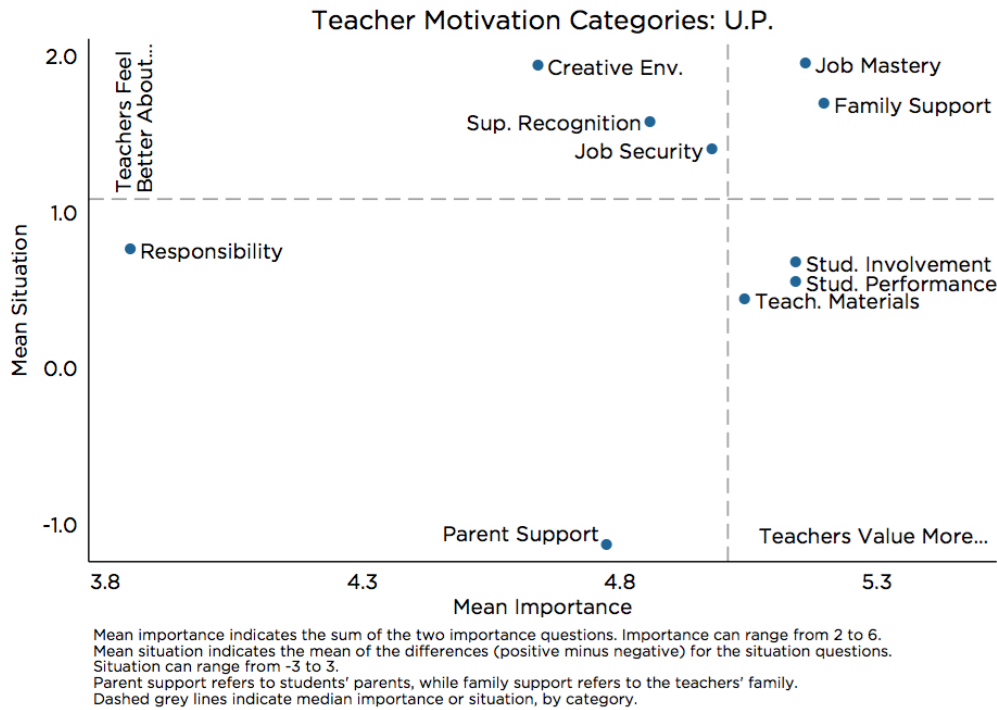
²⁷ To compute the teacher motivation index, the answer to the importance questions were given the following values: A=1, B=2, and C=3.

²⁸ To compute the teacher motivation index, the answer to the teacher situation questions were given the following values: Strongly Disagree=1; Disagree=2; Agree=3; Strongly Agree=4.

than they do on other factors, improving how teachers feel about their interactions with parents could improve teacher motivation.

Figure 3: Teacher Motivation Categories by Importance and Teacher Experience





Classroom Observation

Teacher Activities

The time teachers spend in the classroom was captured as one of three activities: teaching, classroom management, and off task.²⁹ IDinsight enumerators asked for permission to observe the classroom, and given the unexpected nature of the enumerator's visit, the data may show significant Hawthorne effects³⁰ due to teachers acting differently due to the presence of an external person within their classroom.³¹ Hence the reader should interpret the absolute numbers here with some caution. Moreover, these numbers are all conditional on the teacher being present in the school, making them incomparable to some of the studies discussed below. Finally, the reader should not interpret these numbers to indicate that all teachers spend a certain fraction of

²⁹ Teaching is defined to include time spent by a teacher in lecture, answering academic matter related questions, correction of students' academic work and engaging with all or a group of students with regards to what is being taught in the classrooms. Classroom management is defined to include non-academic clarifications of doubts and questions, general classroom discipline and behavior, and other defined administrative responsibilities such as attendance, checking uniforms etc. Finally if a teacher is spending time on any activity other than teaching and classroom management, then the teacher is considered to be off task.

³⁰ Hawthorne effects occur when study participants change their behavior in response to being observed. Given the nature of the survey, Hawthorne effects are likely present in our classroom observation data.

³¹ We try to address this by putting teachers and principals at ease in terms of the objectives of our study. In the future, we will attempt repeated observations for a smaller set of teachers to see if multiple visits will lead to more business-as-usual teaching practices by the teacher.

time on one activity or the other. Instead, the numbers presented here should be interpreted as population-level averages.

This report will refer to the best practices of time allocation as 85% time being spent on teaching and the remaining 15% on classroom management, as defined by the Stallings classroom snapshot tool.

Table 1: Teacher Classroom Activity

<i>Activity</i>	<i>Delhi APS</i>	<i>Uttar Pradesh Government</i>
Teaching	70% (66%, 74%)	76% (73%, 79%)
Classroom Management	28% (25%, 32%)	7% (6%, 9%)
Off Task	2% (1%, 3%)	17% (14%, 19%)
Number of observations	1,216	3,396

The first number listed represents the fraction of teaching time devoted to each activity. Numbers in parenthesis represent 95% confidence intervals. All standard errors are clustered at the school level. Each teacher was recorded three times during the course of the classroom observation survey.

Given STIR's focus on teachers and teaching practices this line of analysis holds special relevance to their theory of change. Trends in how teachers spend their time is observed over a number of factors, to help understand if any external factors contribute to how a teacher chooses to allocate his or her time within the classroom.

Before investigating the baseline data further, it is useful to quickly survey the literature on teacher classroom activities. Much of the literature focuses on the particular issue of teacher absenteeism, rather than what teachers are doing in the classroom. Abadzi (2009) cites an India-wide absenteeism rate of 25%;³² Chaudhury *et al.* (2005)³³ find an Uttar Pradesh-specific

³² Abadzi, H. (2009). Instructional Time Loss in Developing Countries: Concepts, Measurement, and Implications. *The World Bank Research Observer*, 24(2), 267-290.

³³ Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K., & Rogers, F. (2006). Missing in Action: Teacher and Health Worker Absence in Developing Countries. *Journal of Economic Perspectives*, 91-116.

absenteeism rate of 27%; and Muralidharan and Sundararaman (2011)³⁴ find a teacher absenteeism rate of about 25% in Madhya Pradesh. While this is similar to anecdotal evidence from IDinsight's baseline survey, IDinsight did not conduct detailed attendance data. One study that has looked at teacher classroom activities is Sankar and Linden (2014),³⁵ where they use an expanded version of the Stallings Classroom Snapshot to investigate classrooms in Uttar Pradesh, Madhya Pradesh, and Andhra Pradesh. While the numbers in this study are not directly comparable to this baseline survey due to changes we made in the Stallings snapshot tool, some of the broad findings are in line with IDinsight's experience: classroom management and time off task account for roughly 20% of U.P. government teachers' time, where the number is about 24% in IDinsight's sample. Otherwise the data reported is not directly comparable with IDinsight's study.

The baseline data show that teachers in both Delhi and U.P. spend less time teaching than the ideal time allocation. In U.P, teachers spend a substantial amount of time off task, while Delhi teachers spend nearly 30% of their time on classroom management, indicating potential inefficiencies in their administrative duties. Some part of this could be attributed to the requirements of teachers in public schools to have more out-of-class management practices such as regular student attendance and data reporting requirements.³⁶ For both Delhi APS and U.P. government schools, however, there is significant room for improvement in the amount of class time devoted to instruction.

Figure 4 shows teacher activities by grade and geography. In UP schools, there is not significant variation across grades in terms of teachers' various activities.³⁷ In Delhi, however, classroom management takes a significant amount of time in first grade, but appears to decrease by the time students are in upper primary grades.³⁸

³⁴ Muralidharan, K., & Sundararaman, V. (2011). Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, 39-77.

³⁵ Sankar, D. & Linden, T. (2014). How much and what kind of teaching is there in elementary education in India? Evidence from three States. *South Asia Region, Human Development Sector, Discussion Paper Series*, No. 67.

³⁶ Aiyar, Y., Dongre, A., Kapur, A., Mukherjee, A. N., & Raghunandan, T. R. (2014). *Rules vs. Responsiveness: Towards Building an Outcome-Focused Approach to Governing Elementary Education Finances in India*. Retrieved from: http://www.accountabilityindia.in/sites/default/files/rules_vs_responsiveness_7.5.15_revised2_0.pdf.

³⁷ A chi-squared test was used to determine whether teacher activity varies significantly by grade. It returned a *p*-value of approximately 0.60, indicating that in U.P., teachers do not vary their activities significantly by grade.

³⁸ A chi-squared test was used to determine whether teacher activity varies significantly by grade. It returned a *p*-value of less than 0.01, indicating that in Delhi, teachers vary their activities significantly by grade.

Figure 4: Teaching Activity by Grade and Geography

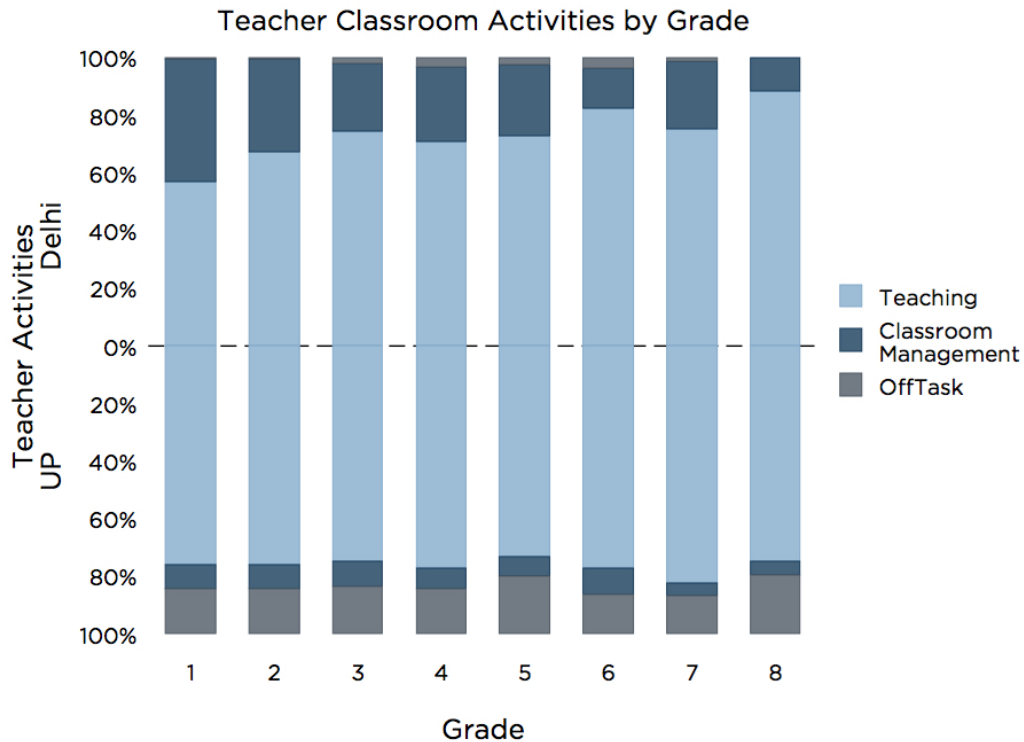
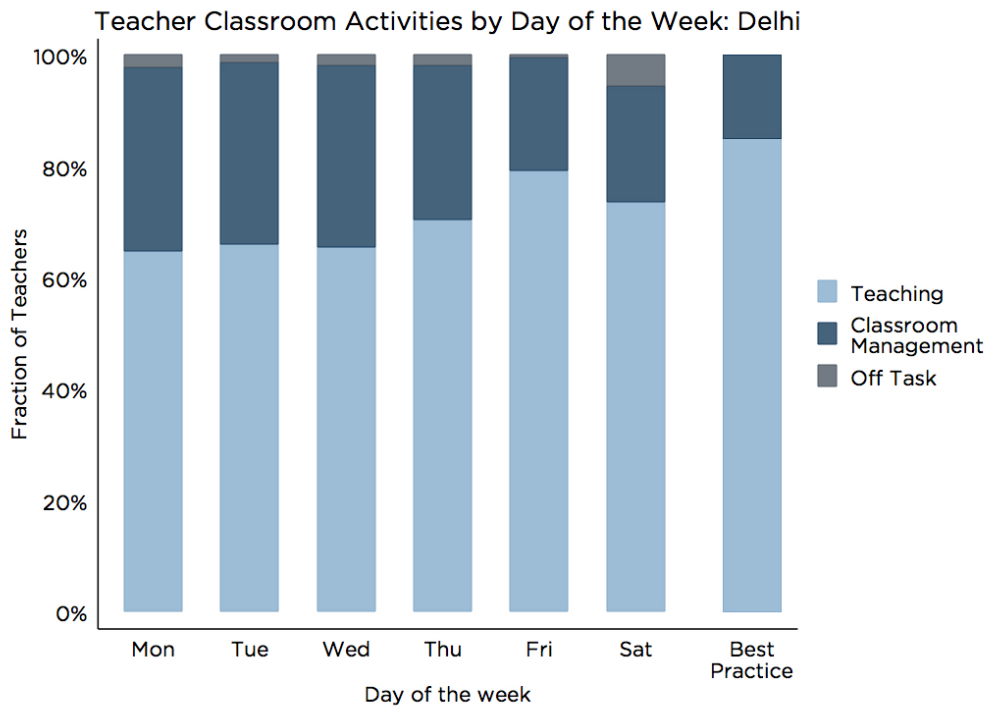
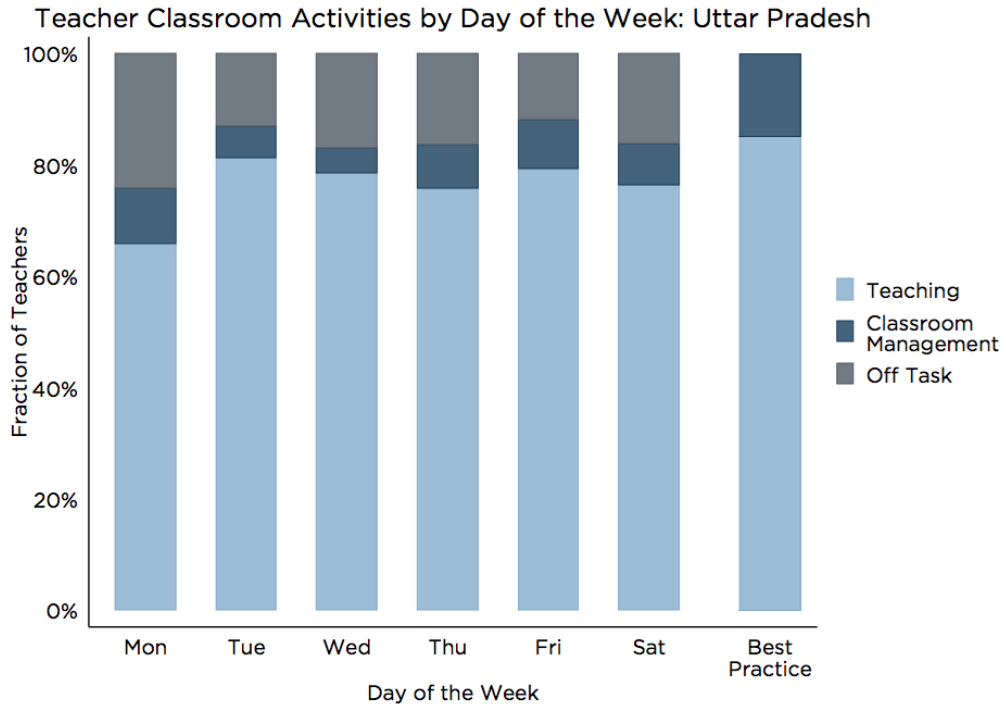


Figure 5 shows the breakdown of teaching activities by day of the week. In the Delhi APS, more time seems to be spent off-task on Saturdays, though this result is not statistically significant. On the other hand, in Uttar Pradesh, there appears to be a significant increase in time teachers spend off task on Mondays. Note that this may also be an artefact from data collection, as Mondays likely have a higher variance than other days of the week in this dataset.³⁹

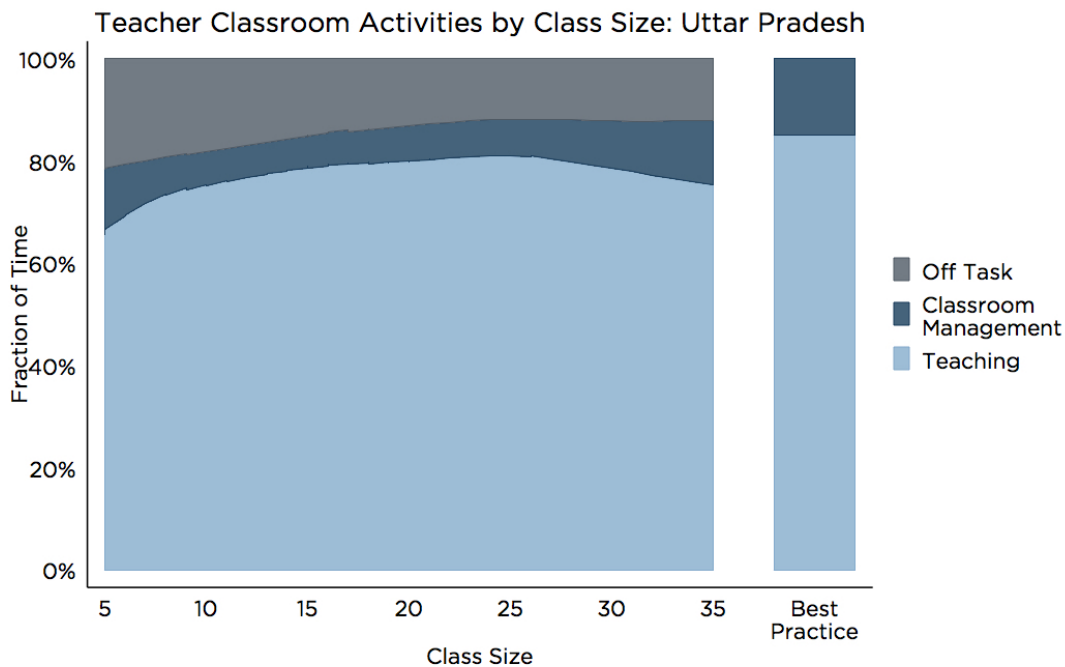
³⁹ During the course of this study, several Mondays in U.P. were local religious holidays, meaning that there are fewer Mondays than any other day of the week present in the dataset.

Figure 5: Teacher activities by day of the week: Delhi and U.P.



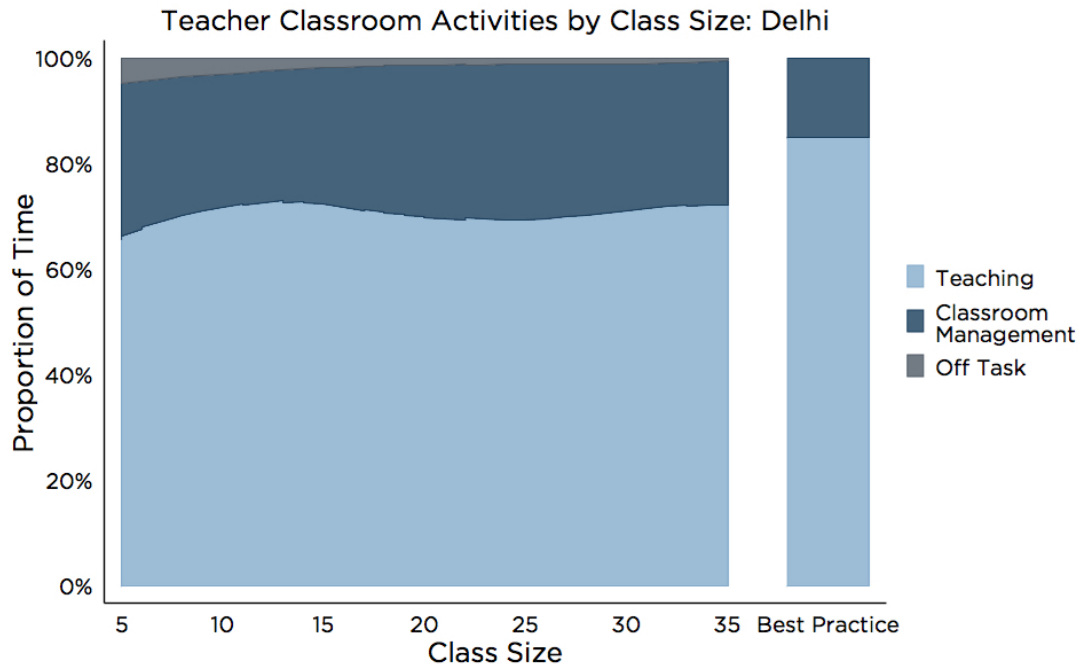
Teacher activity within the classroom does not vary much with the number of children in a particular class in Delhi.⁴⁰ In U.P. however there are certain trends in teaching activity as class size varies. The proportion of time spent off task reduces as the class size increases. Interestingly, the proportion of time spent on classroom management is higher for both small and large class sizes and more uniform in between. It is important to note that the ends of these graphs have fewer observations, and hence a larger variance, making it difficult to make concrete conclusions about very large or very small classrooms.

Figure 6: Teacher activities by classroom size: Delhi and U.P.



Data represent only times when the class is in session.
 Percentages were calculated using a local polynomial smoothing algorithm.
 Best practices for the Stallings snapshot tool are defined as 85% teaching time and 15% classroom management time.
 Data represent 3,396 classroom observations.

⁴⁰ The number of students in the class are defined to mean the actual number present during the classroom observation and not the number enrolled i.e. absent students are not included. For this report’s purposes, this seems like the more appropriate definition, as teacher activities may vary with the number of students in class.

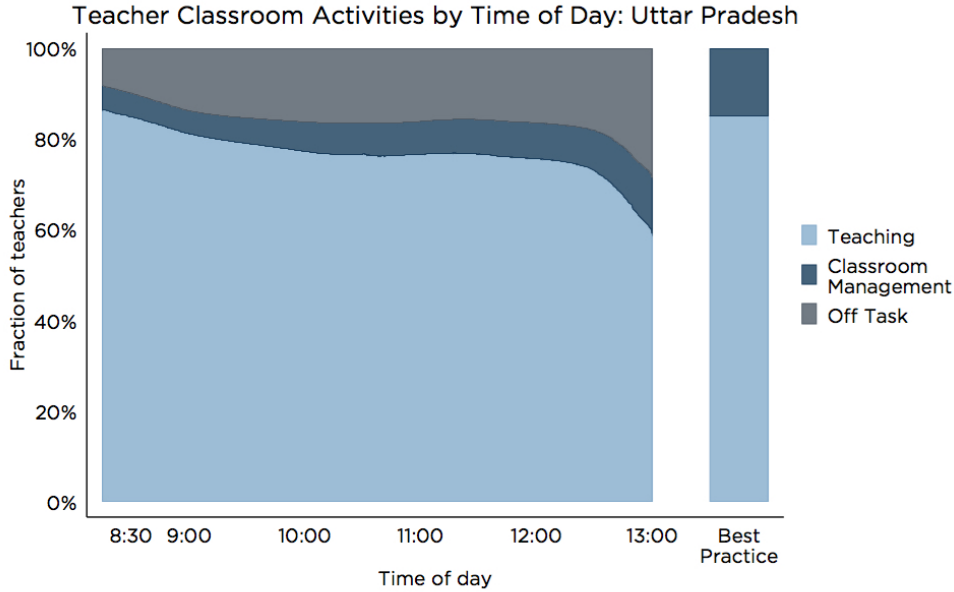


Data represent only times when the class is in session.
 Percentages were calculated using a local polynomial smoothing algorithm.
 Best practices for the Stallings snapshot tool are defined as 85% teaching time and 15% classroom management time.
 Data represent 1,216 classroom observations.

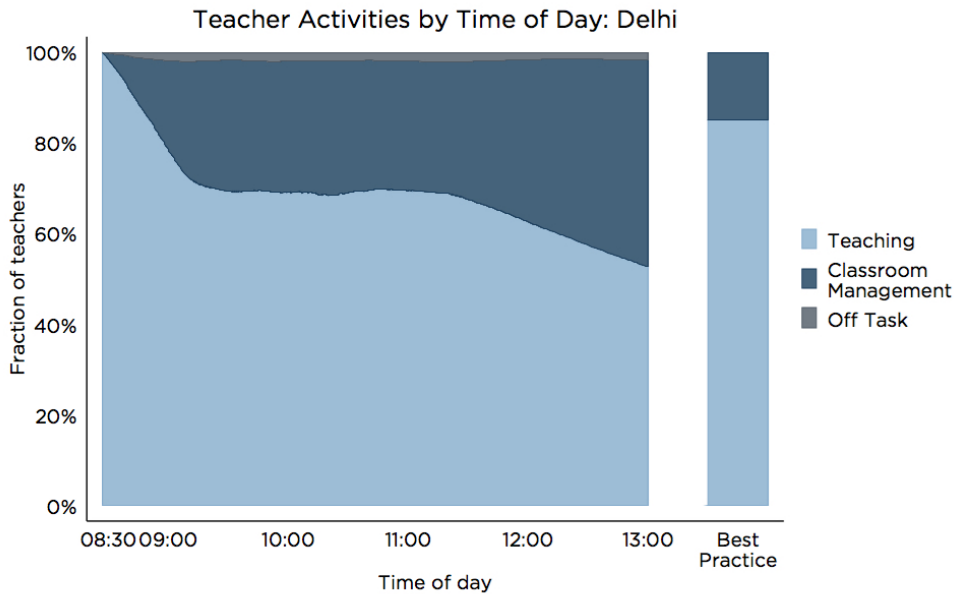
Looking at how teachers structure their classroom activity over the course of the day⁴¹ has interesting implications for STIR’s program. In our classroom observations at the beginning of the day in both Delhi and U.P., teachers spend a larger fraction of their time teaching. However, the proportion of teaching time during our classroom observations decreases in the later half of the day: teachers in Delhi spend more time on classroom management, and teachers in U.P. spend more time off task. See Figure 7 for a graphical illustration of this trend.

⁴¹ Note that IDinsight tried to conduct most classroom observations between 8:30 am and 1:00 pm. This is because most schools had assembly or prayers before 8:30 am, and most teachers were involved with schools ending and students dispersing after 1:00 pm.

Figure 7: Teacher activities by time of the day: Delhi and U.P.⁴²



Data represent only times when class is in session. Percentages were calculated using a local polynomial smoothing algorithm. Best practices for the Stallings snapshot tool are defined as 85% teaching time and 15% classroom management time. Data represent 3,396 classroom observations.



Data represent only times when class is in session. Percentages were calculated using a local polynomial smoothing algorithm. Best practices for the Stallings snapshot tool are defined as 85% teaching time and 15% classroom management time. Data represent 1,216 classroom observations.

⁴² Note that the reader should not interpret these figures as a indicating a typical day for a typical teacher. Rather, these figures represent population-level averages of teaching activities throughout the day.

Child-Friendliness of Classrooms

In order to understand how “child friendly” a classroom is, IDinsight incorporated ASER’s six-point child-friendliness matrix⁴³ into its classroom observation tool. The basic hypothesis of child friendliness is that students will respond to a more conducive and encouraging classroom environment, and will grasp more within the classroom. ASER defines six indicators of “child friendliness”:

1. Did the teacher smile, laugh or joke with at least some students?
2. Did the students ask the teacher at least one question?
3. Was children’s work displayed in the classroom?
4. Did the teacher use local information to make academic content relevant?
5. Did the teacher use any learning aides (posters, chalkboard, supplies) other than the textbook?
6. Did the teacher ask children to work in small groups or pairs?

These indicators provide a picture of the teachers’ behavior with students in the classroom, how comfortable students feel in the class, whether they feel encouraged to ask questions and if students were encouraged to work in smaller groups or teams.

The results are presented in **Table 2** below. The interpretation is consistent with that followed by ASER, which is not the proportion of time spent on each of the following indicators but rather the percentage of times the indicators was observed at least once in a classroom.⁴⁴

- **Teaching aides:** Teaching aides other than textbooks seem to be used in both Delhi and U.P. This proportion declines consistently from grade 1 to grade 8, but the fall over grades cannot be interpreted as statistically significantly different from zero.⁴⁵
- **Student participation:** Student participation is higher in Delhi than in U.P., as seen by the percentage of classrooms where students ask at least one question. In more than 80% of Delhi classrooms, at least one student asked a question. In Uttar Pradesh the incidence of the same was in just over a third of classrooms observed.
- **Students’ work displayed:** Display of the students’ work in the classrooms is low in U.P. APS schools in general have more vibrant classrooms despite the physical structures being smaller on both an absolute and per-student scale.

⁴³ Bhattacharjea, S., Wadhwa, W., & Banerji, R. (2011). *Inside Primary Schools: A study of teaching and learning in rural India*. Retrieved November 2, 2015, from the ASER centre’s website:

http://img.asercentre.org/docs/Publications/Inside_Primary_School/Report/tl_study_print_ready_version_oct_7_2011.pdf

⁴⁴ These indicators were taken as part of the classroom observation, described in the “Teacher Activities” section.

⁴⁵ A chi-squared test for difference of categorical distributions gives a *p*-value of 0.112, indicating that there is not a statistically significant difference in use of teaching aids by grade.

- **Group work:** The emphasis on group work is low in both geographies. Students worked in small groups in about 9% of U.P. classrooms and 4% of Delhi classrooms.

Table 2: ASER Child Friendliness Matrix Summary Statistics

<i>Indicator</i>	<i>Delhi APS</i>	<i>Uttar Pradesh Government</i>
Teacher smiled at least once during the course of the class	47% (39%, 51%)	14% (11%, 17%)
At least one student asked a question	84% (78%, 86%)	37% (32%, 42%)
Teacher used local information while teaching	37% (31%, 43%)	24% (21%, 28%)
Teacher used at least one learning aid	61% (57%, 68%)	58% (54%, 63%)
Teacher asked students to work in groups during the class	4% (1%, 6%)	9% (6%, 12%)
Students' work was displayed	22% (18%, 27%)	8% (6%, 11%)
Number of observations	327	842

The first number listed represents the fraction of classrooms where each indicator was observed. Numbers in parenthesis represent 95% confidence intervals. All standard errors are clustered at the school level. Only one observation per teacher was recorded for the entire classroom observation.

Student Activities

The classroom observation tool also recorded students' activities in the classroom. These activities were classified into drills or practice, group discussions/question and answer, lecture or

practice, silent work and off task.^{46,47} While there is no defined best practice as with the teacher activities, fewer students off task reflects better class room discipline and higher potential for learning within the classrooms.

Table 3: Student Activities by Region

<i>Region</i>	Delhi	U.P.
<i>Drills or Practice</i>	26% (24%, 31%)	28% (26%, 31%)
<i>Group Work/ Question and Answers</i>	18% (14%, 19%)	6% (4%, 7%)
<i>Lecture or Demonstration</i>	19% (15%, 21%)	31% (26%, 32%)
<i>Silent Work</i>	29% (27%, 34%)	17% (16%, 20%)
<i>Off task</i>	8% (7%, 9%)	18% (17%, 21%)

The first number listed represents the fraction of student time devoted to each activity. Numbers in parenthesis represent 95% confidence intervals. All standard errors are clustered at the school level.

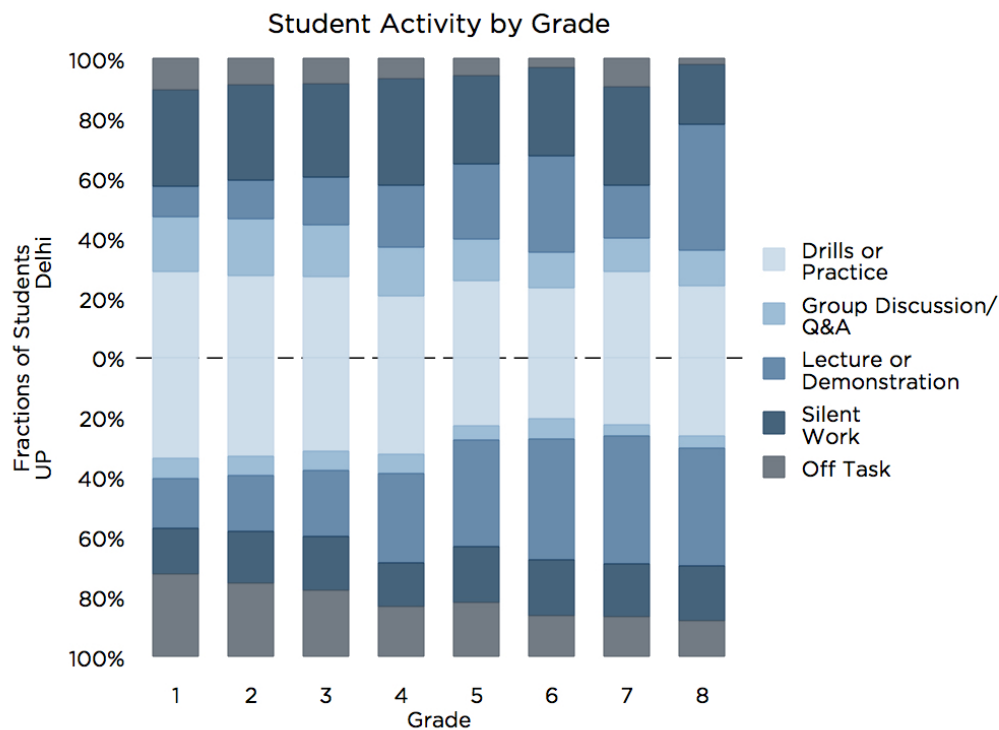
⁴⁶ Only off task is defined to include non-academic participation by students. All other activities pertain to the academic curriculum being taught in the classrooms directed by the teacher in the classroom.

⁴⁷ The difference in interpretation between student and teacher activities is that teacher activities are captured for a particular teacher and hence reflects the fraction of time a teacher spends on each of the activities. Student activities however are defined as the proportion of students engaged in each of the activities.

The distribution of students across various activities reflects the following:

- **Drills and practice:** The proportion of students engaged in drills or practice is similar in both Delhi and U.P. across all grades.
- **Group discussions:** Group discussions and questions and answers are more visible in schools in Delhi than in U.P. This is common across all grades. This is in line to trends in “child friendliness” observed in previous sections.
- **Off task:** A larger proportion of students are off task in U.P. than in Delhi. The difference is statistically significant at the 1% level at the aggregate level.⁴⁸ Interestingly amongst the grade wise difference only for grades one and two is the difference in proportions between Delhi and U.P. statistically different.⁴⁹

Figure 8: Student activities by grade and geography



Student attentiveness⁵⁰ was observed throughout the day for both the APS schools in Delhi and Government schools in U.P. There is a slight increase in the proportion of students off task in Delhi as the school day nears its end, but this change is not statistically significant, and the overwhelming majority of students are attentive throughout the day.

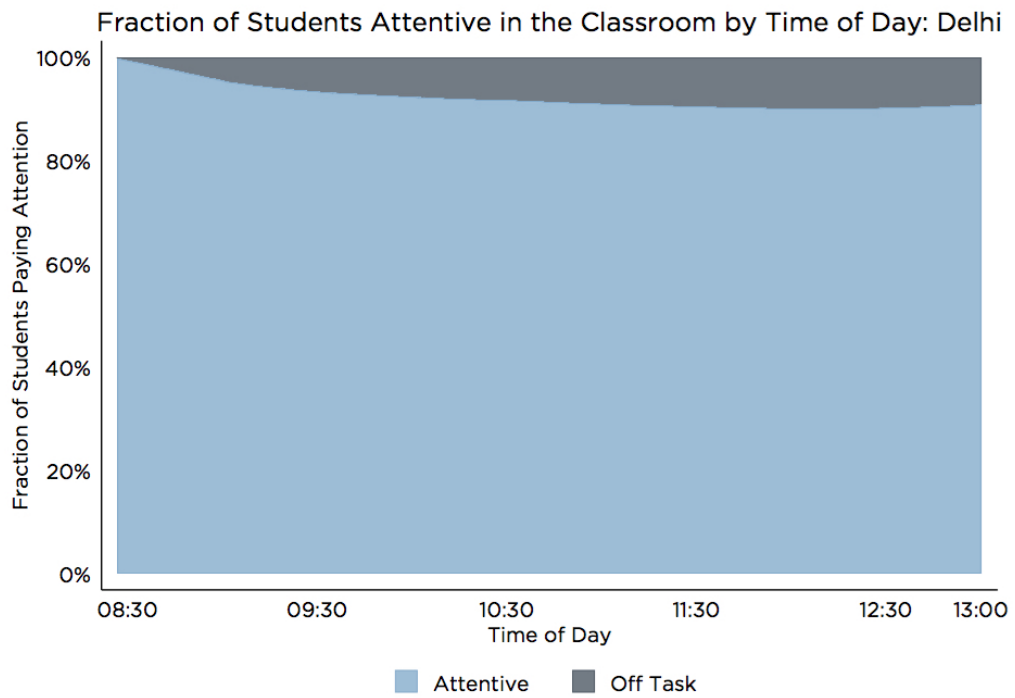
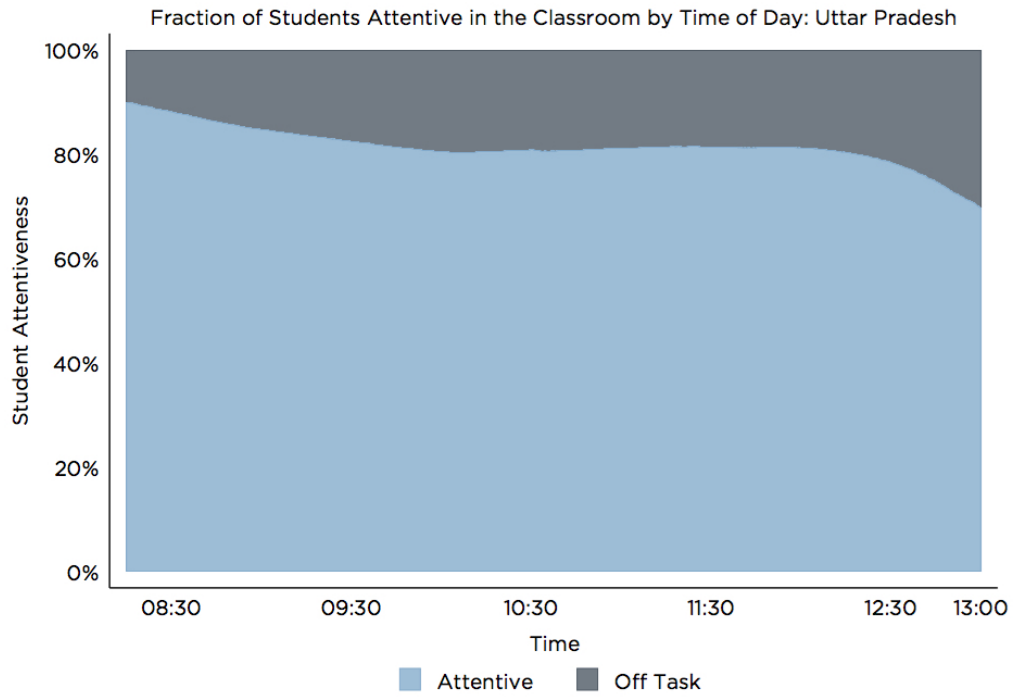
⁴⁸ P- value 0.00

⁴⁹ P value grade 1= 0.019 ; grade 2 = 0.009

⁵⁰ Student attentiveness is to be interpreted as proportion of students attentive at any given point of time. Students are defined as being attentive if they are not off task.

In U.P. a larger proportion of students are off task throughout the day, on average. There are also two significant dips in the proportion of students who are attentive, the first within the first hour of the school day and the second towards the last hour of the school day.

Figure 9: Fraction of students attentive in the classroom by time of day

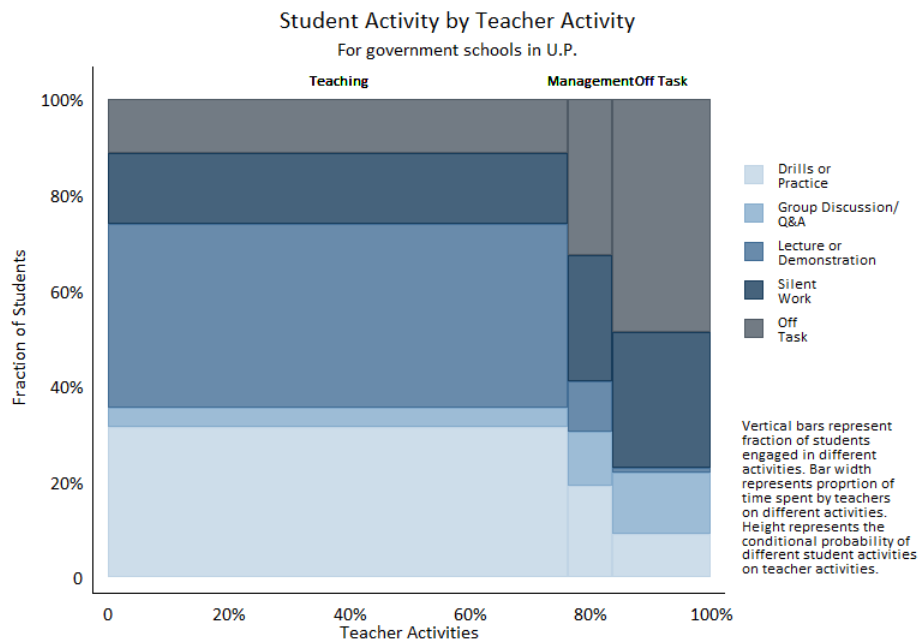
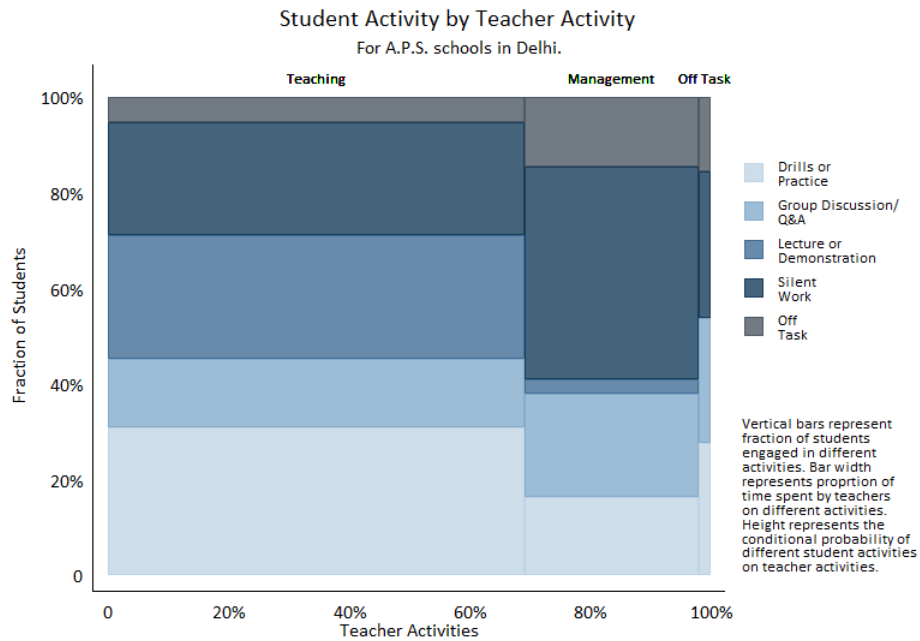


Linking teacher and student activities

Having seen student and teacher activities in isolation, combining the two allows for a conditional analysis. Note, however, that the figures here represent population-level averages, rather than an indication of the time breakdown for a typical teacher. For example, some teachers may spend all their time lecturing while others spend all their time doing drills, rather than all teachers spending some time on both lecture and drills.

When a teacher moves from teaching to classroom management, there is a visible increase in the proportion of students off task and engaged in silent work. There is an associated fall in the proportion of students engaged in drills or practice, group discussions or question and answers and lecture or demonstration. The combined proportion of students involved in these three activities is around 40% in both Delhi and U.P. However in U.P the majority of the remaining 60% children are off task whereas in Delhi they are involved in silent work. This could reflect that teachers in Delhi have a greater command on their classrooms even when they are not teaching, and thus have a more disciplined classroom environment.

Figure 10: Fraction of students engaged in different activities conditional on teacher activities⁵¹



⁵¹ Note that the estimates for students off task may be somewhat conservative, as there is some ambiguity in determining whether children are off task or engaging in a group discussion. Particularly when the teacher his or herself is off task, some of the group discussion is probably not academically relevant.

When teachers are off task an even greater proportion of students are off task as compared to when teachers are teaching or managing the classroom. More than 50% of U.P. students are off task whereas this number is significantly lower in Delhi (which may reflect that teachers assign tasks before leaving the classroom or attending to other matters). Moreover a teacher is off task less than 2% of times in Delhi classrooms. The conditional probability of Delhi students being off task at the same time represents an extremely small proportion.

Student Learning

The ultimate goal of STIR's program is to improve student learning levels. To capture learning levels, an ASER-based Hindi and math oral test was conducted for 10 students chosen at random from the classroom of each teacher within the sample.⁵² Hindi questions ranged from reading letters to reading half-page stories (of standard 3 difficulty) and math ranged from identifying single digit numbers to fractions. A student's learning level is defined as the highest level he or she achieves. In Delhi, students of grades upper kindergarten to 8th were sampled for testing and in U.P grades 1 to 8 were sampled. There are some general trends present in the data:

- **Students are learning:** There is an improvement in student levels between subsequent grades.⁵³ This is reflective both in terms median levels of students and proportion of students with higher math and Hindi levels. Overall, the improvement in levels is higher among Delhi students. For instance in U.P., students in grade 2 and 3 have the same median Hindi levels as do students of grade 4 and 5. This could reflect a higher mixed ability classroom in U.P.
- **Variance remains high, even at higher grades:** For the highest grades (grade 7 and 8), there is significant variance in student level. Approximately 15% of all students tested have answered all questions in the Hindi section and around 14% students have done the same in the math section. This gives some indication that ceiling effects of the tool are relatively limited, which is encouraging.

Math

The results for the mathematics section are presented in Table 4, and are shown graphically by grade in Figure 12. Some common themes emerge from the data:

⁵² Randomized selection of students gives a representative sample of students within sample classrooms. While students not present on the day of the survey were not sampled, this sampling methodology still gives a representative sample, weighted by student attendance.

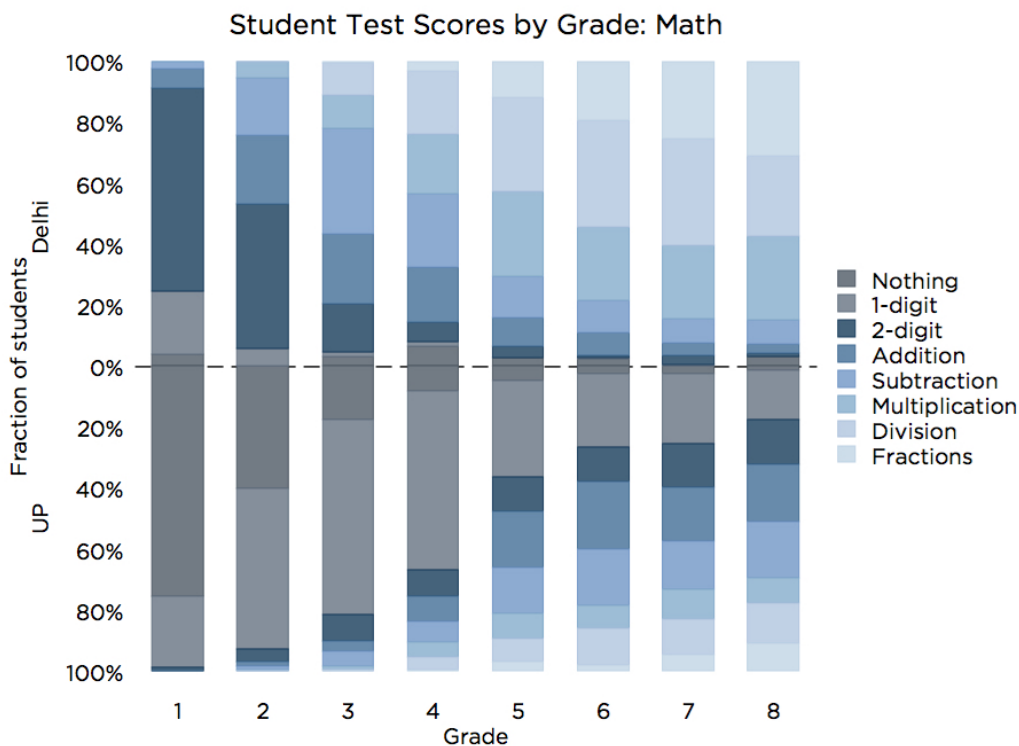
⁵³ Given the ordinal nature of our tool, this makes it easier for us to interpret results.

Table 4: Mathematics Learning Levels by Geography

<i>Region</i>	<i>Nothing</i>	<i>1-digit</i>	<i>2-digit</i>	<i>Add.</i>	<i>Sub.</i>	<i>Mult.</i>	<i>Div.</i>	<i>Frac.</i>
<i>Delhi</i>	4% (3%,5%)	9% (6%,11%)	23% (20%,26%)	11% (10%,14%)	15% (13%,17%)	14% (12%,16%)	15% (13%,18%)	8% (6%,9%)
<i>UP</i>	17% (15%,20%)	38% (36%,41%)	9% (8%,10%)	11% (10%,13%)	10% (9%,11%)	5% (4%,6%)	6% (5%,7%)	2% (1%,3%)

The first number listed represents the fraction of teaching time devoted to each activity. Numbers in parenthesis represent 95% confidence intervals. All standard errors are clustered at the school level.

Figure 11: Math Levels by Grade and Geography



- Learning levels are more uniform in Delhi:** Learning levels are more uniformly distributed across ASER test levels in Delhi as compared to U.P. In Delhi more than half the students have a learning level of subtraction and above.

- **U.P. students are clustered at the lowest two levels:** In U.P. more than half the students are concentrated amongst the lowest two ASER levels and less than 15% students have a level of multiplication and above.
- **Delhi students consistently score higher:** As seen from Figure 12, across all grades student test scores are lower in U.P. than in Delhi. Even among higher grades in Uttar Pradesh close to 20% of students scored at the single digit level or lower.
- **Students in the highest grades show significant differences across geographies:** The difference in student levels between Delhi and U.P is more visible among students of higher grades. The median math level for eight standard students is division in Delhi and addition in U.P.

Hindi

The results for the Hindi section are presented in Table 5, and are shown graphically by grade in Figure 12. Some common themes emerge from the data:

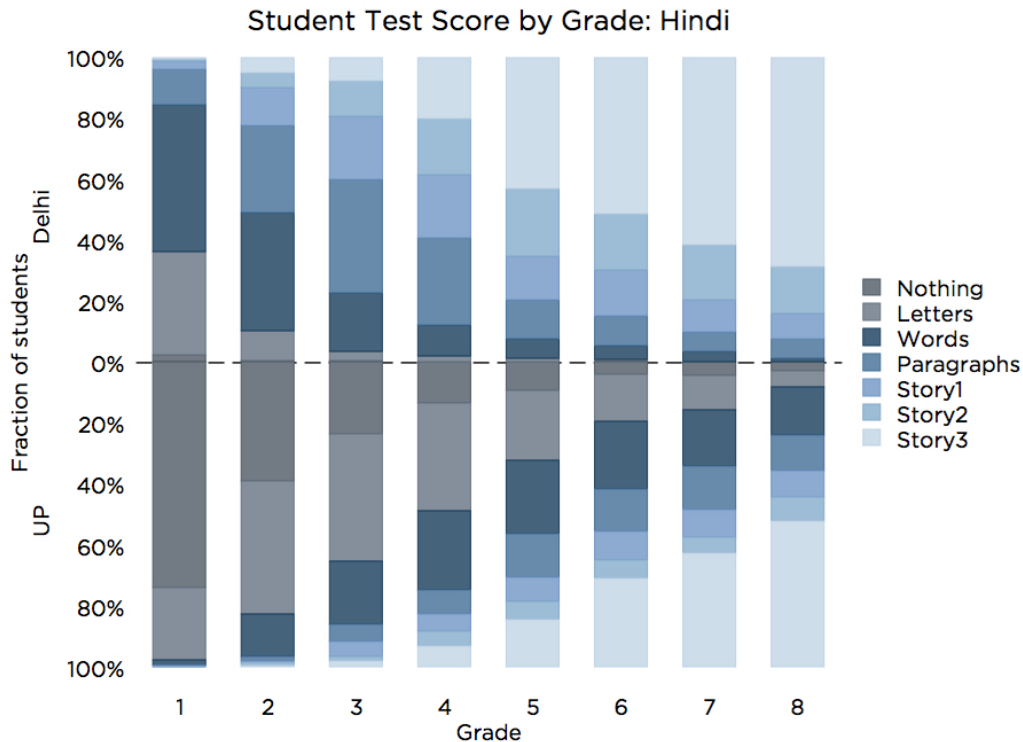
- **U.P. students struggle with reading:** In Delhi the proportion of students who were unable to answer even single question in the Hindi section are marginally lower as compared to the Math section. In U.P. around one-fifth of all students tested were unable to answer even a single Hindi question correctly.
- **Delhi students performing better:** As in Math the proportion of students with lower learning levels is higher in U.P. than Delhi across all grades. At the 8th grade level, in Delhi all students have words level or above whereas in U.P. just below 10% students have the lower two Hindi levels.
- **Hind testing tool appears easier than the mathematics tool:** In both Delhi and U.P., a higher proportion of students were able to complete the entire Hindi section as compared to the mathematics section of the testing tool.

Table 5: Hindi Learning Levels by Geography

<i>Region</i>	<i>Nothing</i>	<i>Letters</i>	<i>Words</i>	<i>Paragraph</i>	<i>Story 1</i>	<i>Story 2</i>	<i>Story 3</i>
<i>Delhi</i>	1% (0%,2%)	13% (11%,16%)	19% (17%,22%)	18% (16%,20%)	12% (11%,14%)	12% (10%,13%)	24% (20%,27%)
<i>UP</i>	20% (18%,22%)	26% (24%,28%)	19% (17%,20%)	9% (8%,10%)	6% (5%,7%)	4% (3%,5%)	17% (14%,20%)

The first number listed represents the fraction of teaching time devoted to each activity. Numbers in parenthesis represent 95% confidence intervals. All standard errors are clustered at the school level.

Figure 12: Hindi reading levels by grade and geography



Recommendations

The baseline survey has given a number of insights into the workings of APS schools in Delhi and government schools in Uttar Pradesh. While IDinsight and STIR will look into these more closely in the future, there are some key themes have emerged from the data that STIR may find useful:

- **Teachers in U.P. should spend more time with remedial students:** A significant fraction of students in U.P. government schools are well below basic reading and math levels. STIR should consider focusing teachers towards this problem and facilitate solution-generation amongst them. These students would likely benefit from teaching targeted closer to their level.⁵⁴
- **Both U.P. and Delhi teachers are teaching less than standard best practices:** As the classroom observation data have shown, both Delhi and U.P. teachers are spending less time than the recommended best practices. In Delhi, STIR might

⁵⁴ Banerjee, A., Cole, S., Duflo, E., & Linden, L. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *The Quarterly Journal of Economics*, 1235-1264.

look for ways to creatively reduce administrative burdens on teachers, allowing them to reduce the amount of time they spend on classroom management. In U.P., it will be interesting to see if STIR's intervention can motivate teachers to spend less time off task or outside the classroom.

- **There is significant room for improvement in the emotional tone of the classroom:** Both Delhi and U.P. teachers scored relatively low on the ASER child-friendliness matrix, although U.P. teachers appeared to score lower on many indicators. In both Delhi and U.P., teachers could show more children's work in the classroom and ask students to do group work more frequently. In U.P., a relatively small fraction of teachers smiled at the students. STIR should consider how best emphasize this using their network meetings and other processes.

Appendix A: Teacher Motivation Survey

Section 1

Instruction: Please tick one box for each question.

How much do you agree with these statements?	Strongly agree	Agree	Disagree	Strongly disagree
5.6 I do not worry about losing my job.				
4.6 My family members don't appreciate my job.				
5.2 I am given many extra responsibilities in my school.				
3.3 Often I need help from my colleagues, but I don't receive any.				
3.2 I have ample opportunities to show my creativity in the classroom.				
5.4 Parents are proactive about making sure that their children complete their homework.				
2.6 I work very hard but don't receive any praise for my work.				
1.2 When I work hard, students perform up to my expectations.				
2.5 I am constantly learning and improving my skills.				
2.2 Parents don't pay attention to their children's education.				
5.1 I need more teaching learning material (charts, models, etc.) but it is not available in the school.				
1.1 My supervisor praises me for my efforts in the school.				
3.1 I am not given any extra responsibilities in the school.				
4.1 My colleagues often support me.				
4.5 I work hard but students don't perform up to my expectations.				
5.5 My family members praise my efforts.				
1.5 I don't have freedom to adopt new methods in my school.				
1.3 I am satisfied with the quality of teaching learning material available in the school.				
1.6 Teaching doesn't provide enough opportunities to improve my skills.				
1.4 I feel insecure about my job.				

Section 2: Situational Questions

These questions are all situations related to teachers similar to you. Please tick on one option for each of the questions below.

- 1.8 When a temporary teacher becomes permanent in her school, how will she feel?
- Promotion is important but will not make her happy.
 - She will be happy but this will not impact her performance.
 - She will be happy and will work hard.
- 1.9 If a teacher gets fewer learning opportunities in her school, how will she feel?
- It will not matter much to her.
 - She will be disappointed but her teaching will not be affected.
 - She will be disappointed and her teaching will be affected.
- 1.10 Suppose that Suresh doesn't get support from his colleagues when he needs it. How will he feel?
- It will not matter much to him.
 - He will be disappointed and his teaching will not be affected.
 - He will be disappointed and his teaching will be affected.
- 1.12 Shyama ensures that she lets parents know of their children's shortcomings but parents don't respond well. How will she feel?
- It will not matter much to her.
 - She will be disappointed but her teaching will be not affected.
 - She will be disappointed and her teaching will be affected.
- 2.7 Suppose a supervisor praises her teachers a lot. How will they feel?
- This is important but will not make them happy.
 - They will be happy but this might not impact her performance.
 - They will be very happy and will work hard.
- 2.8 If a teacher works very hard and the students don't perform according to her expectations, how will she feel?
- It will not matter much to her.
 - She will be disappointed but her teaching will be not affected.
 - She will be disappointed and her teaching will be affected.
- 2.9 Suppose a science teacher has charts and models to aid her teaching. How will she feel?
- This is important but will not make her happy.
 - She will be happy but this will not impact her performance.
 - She will be very happy and will work hard.
- 2.11 If a teacher is not appreciated by her colleagues for her innovative methods, how will she feel?
- It will not matter much to her.
 - She will be disappointed but her teaching will be not affected.

- c. She will be disappointed and her teaching will be affected.

3.7 If a teacher takes an urgent leave and her colleagues bear her responsibilities, how will she feel?

- a. It will not matter much as it is their duty.
- b. She will be happy but this will not impact her performance.
- c. She will be very happy and will work hard.

3.8 A teacher always tells parents to pay attention to their children's shortcomings and parents follow her directions. How will the teacher feel?

- a. It will not affect her much as parents are just fulfilling their duty.
- b. She will be happy but this will not impact her performance.
- c. She will be very happy and will work hard.

3.10 Mamta's family members support her a lot with respect to her profession. How will she feel?

- a. This is important but will not make her happy.
- b. She will be happy but this will not impact her performance.
- c. She will be very happy and will work hard.

3.12 Suppose a teacher, Seeta, isn't given extra responsibilities in the school while her colleague, Pratibha, is given a lot of responsibilities in the school. How will Seeta feel?

- a. Seeta wouldn't mind this at all.
- b. She will be disappointed but her teaching will be not affected.
- c. She will be disappointed and her teaching will be affected.

4.8 If a teacher does innovative activities with students and her colleagues praise her, how will she feel?

- a. It won't affect her.
- b. She will be happy but this will not impact her performance.
- c. She will be very happy and will work hard.

4.10 Ram is working as a temporary teacher in his school. What is true for him?

- a. Being a temporary or permanent teacher doesn't matter at all.
- b. He will be disappointed but his teaching will be not affected.
- c. He will be disappointed and his teaching will be affected.

4.12 Radha is a teacher and she is not satisfied with her professional development in the current school but is otherwise happy. What will she do?

- a. She will not switch to another school.
- b. She will consider switching to another school.
- c. She will switch to another school.

5.7 If Rekha is given a lot of responsibilities in the school then how will she feel?

- a. It will not matter at all.
- b. She will be happy but this will not impact her performance.
- c. She will be very happy and will work hard.

5.9 A teacher's family is not supportive of her job and she often has to do household chores even after a tiring workday. What is true among the following?

- a. Family support doesn't matter at all.
- b. She might do better with family support.
- c. She would do better with family support.

5.10 Saurabh is a principal who never praises his teachers. How will the teachers at his school feel?

- a. It won't matter much.
- b. They will be disappointed but their teaching will be not affected.
- c. They will be disappointed and their teaching will be affected.

5.11 Suppose a teacher, Govind, gave a surprise test to his students and students performed well. How will he feel?

- a. It will not matter.
- b. He will be happy but this might not impact her performance.
- c. He will be very happy and will work hard.

5.12 Even after multiple requests, a teacher doesn't get desired teaching material for her class. How will she feel?

- a. It won't matter much.
- b. She will be disappointed but her teaching will be not affected.
- c. She will be disappointed and her teaching will be affected

Appendix B: Classroom Observation Tool

Classroom Assessment Tool

Observation date: [][]-[][]-[][]

Classroom entry time: [][]:[][]

Begin observation 1 (5 minutes after classroom entry time): [][]:[][]

Record observation 1 (10 minutes after classroom entry time): [][]:[][]

Record observation 2 (15 minutes after classroom entry time): [][]:[][]

Record observation 3 (20 minutes after classroom entry time): [][]:[][]

Record observation 4 (25 minutes after classroom entry time): [][]:[][]

STOP. Please wait on this page until it is time to begin observation 1.

OBSERVATION 1 RECORD TIME..... [][] : [][]

A. Classroom Snapshot (1): *Answer questions about this moment (now).*

1 What is the teacher doing?

- A Teaching students (discussing academic material)
- B Classroom management (discipline, attendance, or other non-academic interaction)
- C Out of classroom or off task

2 How many students are undertaking the following activities?

Drills or practice	
Group discussion/questions and answers	
Listening to lecture or demonstration	
Silent work	
Off task	

B. Classroom overview (1): *Please answer this question based on the past five minutes only.*

3 Did the teacher smile, laugh or joke with at least some students?

A Yes

B No

C Don't know

4 Did the students ask the teacher at least one question?

A Yes

B No

C Don't know

5 Was children's work displayed in the classroom?

A Yes

B No

C Don't know

6 Did the teacher use local information to make academic content relevant?

This includes use of objects, events, places or people with which students are familiar.

A Yes

B No

C Don't know

7 Did the teacher use any learning aides (posters, chalkboard, supplies) other than the textbook?

- A Yes
- B No
- C Don't know

8 Did the teacher ask children to work in small groups or pairs?

--

- A Yes
- B No
- C Don't know

9 Which materials were in use during the observation?

Textbooks	
Notebooks	
Blackboards	
Learning aids	

10 What topics were covered during this class?

Math

Single-digit numbers	
Double-digit numbers	
Addition or subtraction	
Multiplication or division	
Other math	

Hindi

Letters	
Words	
Sentences	

Stories	<input type="text"/>
Vocabulary	<input type="text"/>
Other	<input type="text"/>
<i>Other</i>	
Other subject	<input type="text"/>
	<input type="text"/>

STOP.

Please wait on this page until it is time to begin observation 2.

OBSERVATION 2 RECORD TIME..... [][] : [][]

C. Classroom Snapshot (2): *Answer questions about this moment (now).*

- 1 What is the teacher doing?
- A Teaching students (discussing academic material)
 - B Classroom management (discipline, attendance, or other non-academic interaction)
 - C Out of classroom or off task

2 How many students are undertaking the following activities?

Drills or practice	
Group discussion/questions and answers	
Listening to lecture or demonstration	
Silent work	
Off task	

D. Classroom overview (2): *Please answer this question based on the past four minutes only.*

3 Did the teacher smile, laugh or joke with at least some students?

A Yes

B No

C Don't know

4 Did the students ask the teacher at least one question?

A Yes

B No

C Don't know

5 Was children's work displayed in the classroom?

A Yes

B No

C Don't know

6 Did the teacher use local information to make academic content relevant?

This includes use of objects, events, places or people with which students are familiar.

A Yes

B No

C Don't know

7 Did the teacher use any learning aides (posters, chalkboard, supplies) other than the textbook?

A Yes

B No

C Don't know

8 Did the teacher ask children to work in small groups or pairs?

A Yes

B No

C Don't know

9 Which materials were in use during the observation?

Textbooks

Notebooks

Blackboards

Learning aids

10 What topics were covered during this class?

Math

Single-digit numbers

Double-digit numbers

Addition or subtraction

Multiplication or division

Other math

Hindi

Letters

Words	
Sentences	
Stories	
Vocabulary	
Other	
<i>Other</i>	
Other subject	

STOP.

Please wait on this page until it is time to begin observation 3.

OBSERVATION 3 RECORD TIME..... [][] : [][]

E. Classroom Snapshot (3): *Answer questions about this moment (now).*

1 What is the teacher doing?

- A Teaching students (discussing academic material)
- B Classroom management (discipline, attendance, or other non-academic interaction)
- C Out of classroom or off task

2 How many students are undertaking the following activities?

Drills or practice	
Group discussion/questions and answers	
Listening to lecture or demonstration	
Silent work	
Off task	

CONTINUE TO NEXT PAGE IMMEDIATELY

Observation 3 continues...

OBSERVATION 4 RECORD TIME.....

	[]	[]	:	[]	[]
--	-----	-----	---	-----	-----

F. Classroom Snapshot (4): *Answer questions about this moment (now).*

- 1 What is the teacher doing?
- A Teaching students (discussing academic material)
 - B Classroom management (discipline, attendance, or other non-academic interaction)
 - C Out of classroom or off task

2 How many students are undertaking the following activities?

Drills or practice	
Group discussion/questions and answers	
Listening to lecture or demonstration	
Silent work	
Off task	

G. Classroom overview (3): *Please answer this question based on the past four minutes only.*

3 Did the teacher smile, laugh or joke with at least some students?

A Yes

B No

C Don't know

4 Did the students ask the teacher at least one question?

A Yes

B No

C Don't know

5 Was children's work displayed in the classroom?

A Yes

B No

C Don't know

6 Did the teacher use local information to make academic content relevant?

This includes use of objects, events, places or people with which students are familiar.

A Yes

B No

C Don't know

7 Did the teacher use any learning aides (posters, chalkboard, supplies) other than the textbook?

- A Yes
- B No
- C Don't know

8 Did the teacher ask children to work in small groups or pairs?

- A Yes
- B No
- C Don't know

9 Which materials were in use during the observation?

Textbooks	
Notebooks	
Blackboards	
Learning aids	

10 What topics were covered during this class?

Math

Single-digit numbers	
Double-digit numbers	
Addition or subtraction	
Multiplication or division	
Other math	

Hindi

Letters	
Words	
Sentences	
Stories	
Vocabulary	
Other	
<i>Other</i>	
Other subject	

END OF SURVEY IN THIS CLASSROOM.

You have completed the survey in this classroom.

Please thank the teacher and leave the classroom. Thank you!

Appendix C: Data Quality Standards

A number of steps were taken to ensure utmost quality in conducting the survey. The data quality protocols set in helped meeting three broad objectives:

1. **Ensuring uniformity in surveying:** Given the inherent subjectivity associated with student testing and classroom observations⁵⁵ it was essential for IDinsight that the entire field team followed the same rules and maintained the same standards while surveying. This was essential to ensure that each enumerator (surveyor) interprets and codes similar situations the same way⁵⁶. In the absence of this, there would be a threat in comparing data across enumerators. This was achieved by:
 - a. **Using audio and video clips for training:** Audio and video clips from the pilot survey was used during enumerator training to recreate situations. Creating a virtual environment helped define standards for how different situations were to be interpreted.
 - b. **Regular spot checks and discussions:** In both Delhi and U.P. IDinsight field managers conducted daily debriefing sessions with the entire field team. During these sessions the field managers would bring up scenarios they had noticed during their visit to schools in the day and would ensure all enumerators were in agreement of how the scenario was to be interpreted.
 - c. **Regular data checks:** IDinsight Associates conducted data audit checks on a daily basis on the data collected during the day. This helped identify outlier cases where it seemed that the data defied the general trends at both the survey and the enumerator level.

2. **Preventing misrepresentation of data:** Given the scale of field operations it was not possible for either IDinsight Field Managers or Associates to oversee each enumerator throughout the day. To help identify and prevent enumerators from collecting incorrect data, a number of checks were incorporated within the SurveyCTO⁵⁷ survey form⁵⁸. Most of the checks were built in to appear randomly. This prevented enumerators from finding loopholes. Some of the checks built in were:
 - a. **Random selection screens:** During the course of the student testing and classroom observation often a screen would appear which prompted

⁵⁵ Activities of teachers in the classroom and whether or not a student has answered a question correctly often is not easy to classify. Hence it was important to define rules and set standards to ensure consistency.

⁵⁶ While it is not plausible to alienate all enumerator specific effects and ensure all situations are preempted during training, the data checks conducted continuously by IDinsight reflected a good quality of surveying.

⁵⁷ SurveyCTO is an offline mobile data collection tool: <http://www.surveyccto.com/index.html>

⁵⁸ The data collection was done electronically using smart phones.

- enumerators to select a specified option. This helped ensure that enumerators are focused and aware while conducting the survey. During automated data checks performed each day, selecting the incorrect answer would trigger further data checks of that survey.
- b. **Photos:** At the end of student tests, the forms would prompt the enumerators to take a photo of the student and their answer sheets. This was a good check to ensure that enumerators did not falsify the student learning data and that the survey was conducted.
 - c. **Audio files:** The most important check incorporated was to randomly record audio clips of enumerators in the schools. These clips were recorded during both the classroom observation and the student testing. These audio clips were regularly audited by IDinsight field managers, who not only used them to check the attentiveness and interaction style of the enumerators but also re-coded the survey themselves. For example, if a field manager was reviewing a student testing audio clip then he or she would fill a separate survey form based on how they would rank the students' learning level based on the students' answers. This also helped identify those enumerators who had not followed rules
 - d. **GPS coordinates:** Every time an enumerator would open a new form on SurveyCTO, his or her GPS coordinates would be recorded automatically. This allowed IDinsight staff to verify that enumerators physically visited the schools they were assigned.
 - e. **Back checks:** IDinsight field managers also conducted random back checks, in which they would visit schools and interact with principals about the visit of the enumerators.
 - f. **Time stamps:** Each section of the survey conducted had an associated time stamp. This time stamp was analyzed as part of the data audits to check how long enumerators spent on each section. This was particularly useful in checking for falsification of data in the Flanders-style verbal interaction section of the classroom observation.⁵⁹ For example, if an enumerator spends about 2 minutes on the first observation and less than half a minute in the remaining 29 observation, the observation was not conducted properly. Moreover, by analyzing the mean time taken per observation as well as the variance, IDinsight was able to address heterogeneity in the coding of observations early on.
3. **Preventing manual errors:** There are several manual entry errors that are associated with using smart phones. While these are in no way malicious it may affect data analysis. To prevent spelling errors the names of all schools and teachers were defined in the phones so that they had to be selected using radio buttons. Selection of each name had to be done twice to ensure consistency: e.g. if

⁵⁹ This section required enumerators to record the speaker and context every 5 seconds for thirty observations (and hence should last roughly two and a half minutes).

an enumerator selects name “Teacher A” the first and “Teacher B” the next time the form would prompt to check the selection.