USAID | RWANDA
FROM THE AMERICAN PEOPLE

A STUDY ON KINYARWANDA INSTRUCTIONAL TIME IN LOWER PRIMARY

February 2020

This publication was produced for the United States Agency for International Development. It was prepared by Christine H. Beggs and María José Ogando Portela of Room to Read for Chemonics International Inc. The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States government. The suggested citation for this report is Beggs, C., Ogando Portela, M.J. (2019). A Study on Kinyarwanda Instructional Time in Lower Primary. Soma Umenye/USAID Rwanda.

THIS PAGE LEFT INTENTIONALLY BLANK.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

We would like to acknowledge the leadership, dedication and hard work of the many people who made this research possible and we hope relevant and useful for the Rwandan government and its partners. We thank the Study Reference Group (Sharon Haba, Lillian Mutesi, Claudien Nzitabakuze, Gerard Murasira, Augustine Gatera, Emmanuel Sibomana, Jean Uwamahoro and James Ngoga) for their time, expertise and genuine investment in this research and its objectives. We are also indebted to the many leaders within the Ministry of Education and the Rwandan Education Board who provided the vision for this study and shared their valuable insights and guidance on the design of the study. The Soma Umenye team's design feedback and support in fielding the study was essential to the success of this work.

This study could not have been implemented without the incredible dedication and hard work of the team at Laterite Rwanda who led the data collection with diligent planning and execution. A special note of thanks to Lydie Shima, Mike Doherty, Rayner Tabetando and the talented team of data collectors who fielded the data collection with such quality and perseverance.

At USAID, we thank Luann Gronhovd and Emily Routte for their support and guidance.

And last, but certainly not least, thank you to the school leaders, teachers and students who provided their time in support of this study and are on the front lines of making the promise of literacy a reality for Rwandan children.

LIST OF ABBREVIATIONS

CWPM	Correct words per minute
EGRA	Early Grade Reading Assessment
ESSP	Education Sector Strategic Plan
GoR	Government of Rwanda
IRR	Inter-rater Reliability
LARS	Learning Assessment in Rwandan Schools
MINEDUC	Ministry of Education
NCST	National Council on Science and Technology
ORF	Oral Reading Fluency
P2	Primary Grade 2
P3	Primary Grade 3
REB	Rwanda Education Board

DEFINITIONS

A list of definitions is provided in Annex D.

EXECUTIVE SUMMARY

WHY THIS STUDY?

The overarching objective of this study is to improve early grade Kinyarwanda reading outcomes in Rwanda, sustainably and at scale. This study provides new information about the amount and nature of instructional time in P2 Kinyarwanda classes and how that time is associated with student reading skills as well as teacher and student characteristics. Developed in collaboration with the Rwandan Education Board (REB) and a reference group consisting of Rwandan education experts, this research supports the Government of Rwanda's policy goals of "enhanced quality of learning outcomes that are relevant to Rwanda's social and economic development", and that "all learners achieve basic levels of literacy and numeracy in early years." Educational research over the past 50 years has increasingly recognized that the function of "time" is critical to learning and as a result, governments and donors have increased investments in time studies. The findings of these studies have been staggering, revealing that in a number of east African countries, 50% or less of instructional time allocated by governments is actually delivered[1], undermining the very foundation of quality education delivery. Furthermore, the study of time in the educational context is intricately linked to understanding the quality and nature of what is happening in the classroom, including teachers' instructional focus and methods, students' focus and engagement and the overall interaction between teachers and students in the learning space.

RESEARCH QUESTIONS AND STUDY DESIGN

For this study we have defined "time" as a multi-dimensional construct, collecting data at multiple levels including:

- Allocated time = the scheduled time in the curriculum for Kinyarwanda lessons.
- In-classroom Time = the amount of time teachers and students are in the classroom together.
- Learning time = the amount of time students are engaged in reading instruction or reading activities that build the fundamental skills children need to learn to read.

The core research questions that are addressed by this study include:

- What is the allocated, in-classroom and learning time for P2 Kinyarwanda lessons?
- How is the learning time allocated across fundamental reading skills instruction?
- What is the association between the amount and type of in-classroom and learning time and reading skills for P2 students?
- For the questions above, how do the findings vary across strata of schools (e.g., urban, rural, low-performing schools, high-performing schools[2]

[1] Filmer, D. "System health at the frontlines: Using SDI Data to analyze education quality in SubSaharan Africa", presented at Making Systems Work: A Global Conference on Education Systems Sydney, March 1, 2016. Stable url: http://pubdocs.worldbank.org/en/812821457978473769/Session-8-Deon-Filmer.pdf

[2] Low-performing and high-performing schools determined by average reading assessment scores.

We conducted two observations of one P2 Kinyarwanda classroom/class section of students in 100 schools across all five provinces.

We surveyed 100 teachers who led the instruction in the observed classrooms.

We assessed the reading skills of and collected profile data for the 1,600 students attending the Kinyarwanda lessons we observed.

We collected administrative data from the 100 schools.

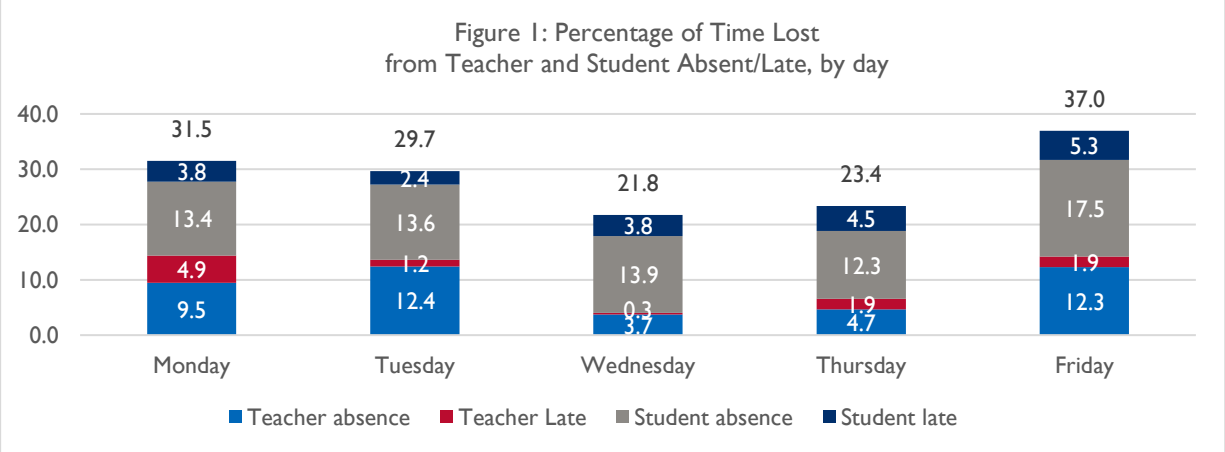
KEY FINDINGS: HOW MUCH OF THE ALLOCATED TIME IS SPENT ON LEARNING?

1. **Only 52.2% of the allocated Kinyarwanda reading instruction time is being dedicated to learning time.**
2. **28.5% of planned instructional time is lost due to teacher and student absence and tardiness**.
3. **19.3% of in-classroom time was utilized for activities other than learning.**

KEY FINDINGS: HOW DOES INSTRUCTIONAL TIME VARY BY DAY, HOUR AND LOCATION?

There are important **variations in teacher/student absence and tardiness by day of the week** as illustrated in Figure 1 below.

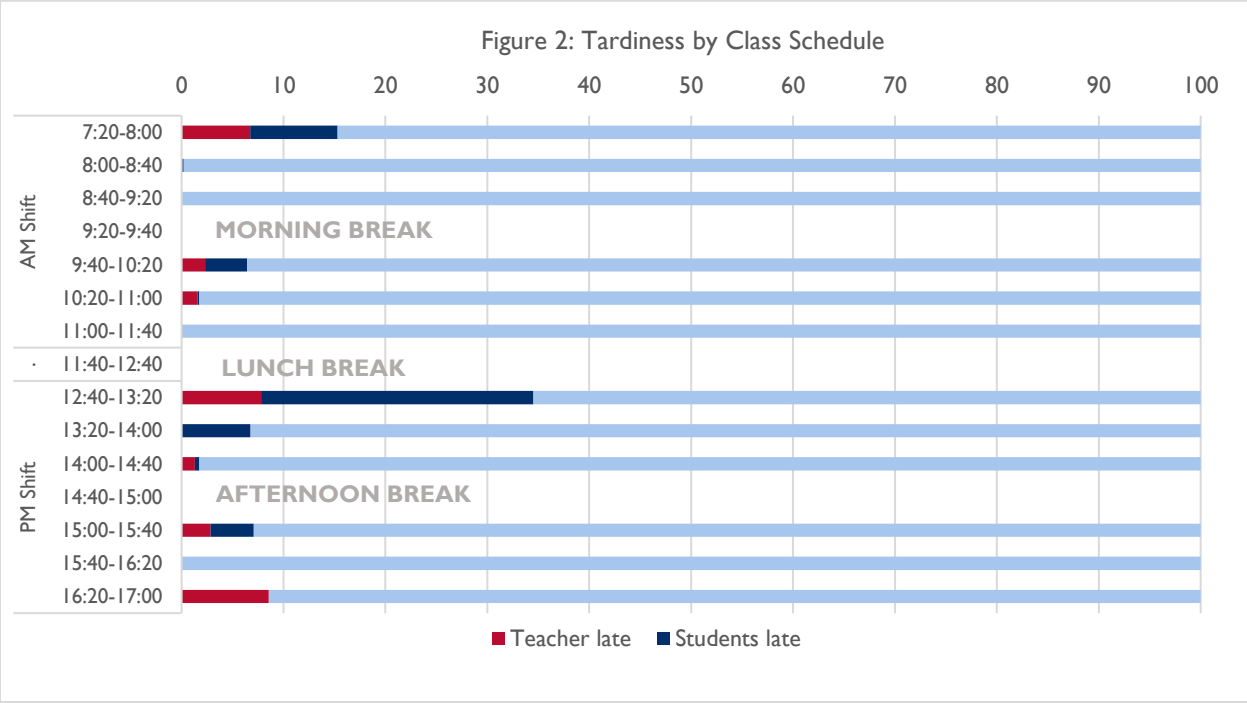
FIGURE 1: PERCENTAGE OF TIME LOST FROM TEACHER AND STUDENT ABSENT/LATE, BY DAY



Time lost due to teacher absence in urban schools was much lower than in rural schools. Student absence was about the same between urban and rural.

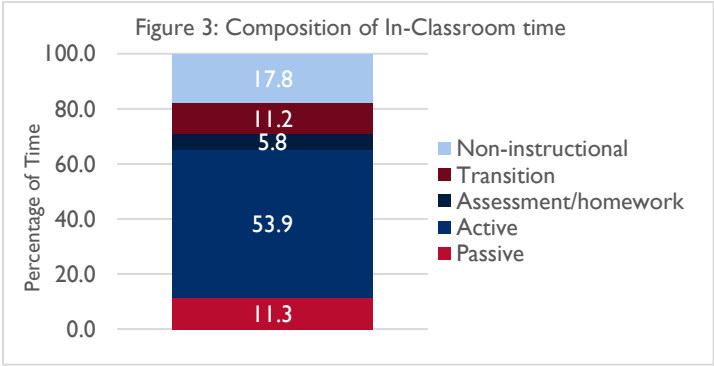
Tardiness varies a great deal by hour/class schedule. The first hour of the AM and PM shifts have the highest percentage of lost time due to tardiness. The effect of this is compounded by a high percentage of P2 Kinyarwanda classes being scheduled in these high tardiness slots.

FIGURE 2: TARDINESS BY CLASS SCHEDULE



KEY FINDINGS: HOW IS IN-CLASSROOM TIME SPENT?

FIGURE 3: COMPOSITION OF IN-CLASSROOM TIME



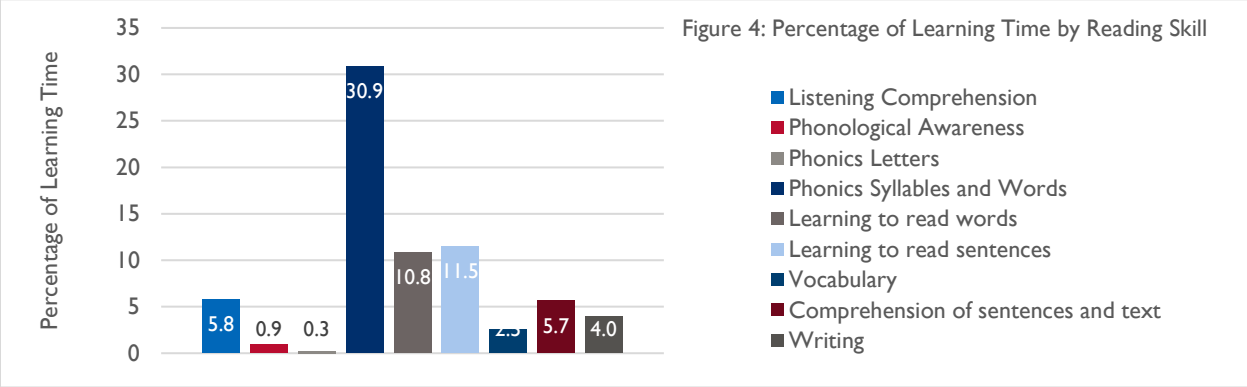
When teachers and students are in the classroom together (71.5% of the allocated time) and focused on learning activities (52.2% of the allocated time), **the majority (65.2%) of in-classroom time[3] is spent on reading instruction** (using both active and passive instructional strategies).[4]

Phonics instruction (syllables and words) dominated learning time with learning to read words and sentences the second and third most frequent reading skills focus in the classroom.

[3] In-classroom time represents 71.5% of the allocated time. Here, the 71.5% of time teachers and students are in the classroom together is taken as 100% of in-classroom time, of which 62.5% is spent on reading instruction.

[4] For example, student activity "Students are listening to the teacher blend syllables/break words into syllables" is categorized as a passive activity while "Students independently read blending syllables together or breaking a word into syllables" is categorized as an active activity.

FIGURE 4: PERCENTAGE OF LEARNING TIME BY READING SKILL

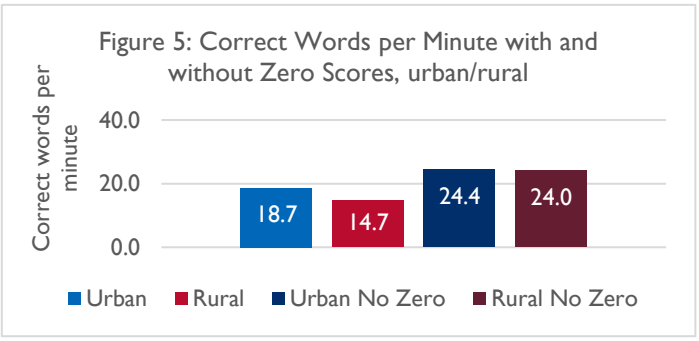


KEY FINDINGS: HOW WELL ARE THE STUDENTS IN OUR STUDY READING?

TABLE 1: STUDENT READING SCORES	% Zero Scores	Average % Correct with Zero Scores	Average % Correct without Zero Scores
P2 Syllable Sounding	30%	49%	59%
P2 Word Reading	33%	59%	87%
P1-leveled Passage (60 seconds)	35%	16.2 cwpm	25.1 cwpm
P1-leveled Comp. Q's Correct	40%	2/5	3.1/5

Students are not performing well, and zero scores are prevalent across all but the letter identification sub-task. Having a head of household who is able to read is associated with higher reading scores, but being older, a boy or repeating a grade is associated with lower reading scores.

FIGURE 5: CORRECT WORDS PER MINUTE WITH AND WITHOUT ZERO SCORES, URBAN/RURAL



When zero scores are excluded from the data, the average score between urban and rural schools is nearly equivalent, indicating that learning is taking place in rural schools, just not for all students.

KEY FINDINGS: TIME AND READING SCORES

We found that **it is not the total amount of in-classroom time that predicts higher student reading scores, but rather *how in-classroom time is spent that matters***. There is a **statistically significant[5] relationship between active instructional time and oral reading fluency scores and reading comprehension scores.**[6] When we **model replacing the 28.5% of lost instructional time due to teacher/student absence and tardiness with active instruction**

[5] "Statistically significant" indicates that the result found is not attributed to chance and that the results can be regarded with confidence.

[6] The association is significant at the 1% level without clustering the standard errors on both outcomes, and at the 10% level and 5% for oral reading fluency and comprehension scores, respectively, with clustered standard errors.

time, the improvements in reading outcomes would be meaningful, with an increase in nearly 3 cwpm in reading fluency and an increase of .35 in comprehension scores and an effect size of .19 SD.

RECOMMENDATIONS: IMPLICATIONS FOR POLICY & PRACTICE IN RWANDA

The scope of these research findings reflects that improving early grade reading outcomes at scale and with sustainability is a complex task and requires a range of actors working together toward the same goal. The full report organizes recommendations by objective and stakeholder group. Below we have presented some of the most pressing recommendations.

INCREASE OVERALL TIME TEACHERS AND STUDENTS ARE IN THE CLASSROOM TOGETHER

- National, local, school and community actors investigate reasons for teacher/student absence/tardiness and establish remediation of absence/tardiness as a priority.
- While the broader issues are being addressed, based on "by day" and "by hour" absence and tardiness data, revise scheduling of Kinyarwanda and other foundational classes so a higher percentage of classes are conducted on days and at times with lower absence and tardiness rates.
- Integrate realistic, progressive targets for reducing absenteeism and tardiness into performance contracts. [7]
- Develop information systems that provide timely and reliable data about teacher and student absence/tardiness, complemented by joint reviews at the national and local levels.
- Incentivize and motivate teachers to reduce absences and tardiness. This will require gaining a better understanding, for each school, the drivers of absenteeism and tardiness for teachers.
- Identify strategies for communities and parents to reduce student absenteeism and tardiness.

INCREASE "ACTIVE" READING INSTRUCTIONAL TIME

- Teacher preparation institutions and in-service teacher professional development programs should train and equip (e.g., with materials) teachers to deliver more "active" instruction.
- Teacher coaching strategies should include a focus on the use of active instructional strategies and engaging distracted or unengaged students.

ADDRESS THE DISPARITY IN READING OUTCOMES ACROSS SCHOOLS AND STUDENTS

- Remediation strategies targeting children with zero scores should be developed and implemented as a priority to address the children that are "internally excluded" from learning. This will help reduce repetition and drop-out rates and server as the most efficient way for Rwanda to increase the overall proportion of students meeting reading benchmarks.
- Develop practical and timely strategies for classroom-based assessments for timely identification of struggling readings and to close the gap between teachers' estimation of student reading skills and actual student reading skills.

---

[7]For additional information on imihigo/performance contracts, see http://documents.worldbank.org/curated/en/833041539871513644/122290272_20181348045807/additional/131020-WP-P163620-WorldBankGlobalReport-PUBLIC.pdf

SECTION 1: INTRODUCTION

The overarching objective of this study is to improve early grade Kinyarwanda reading outcomes in Rwanda, sustainably and at scale. According to the 2018 Learning Assessment in Rwandan Schools (LARS 3), only 55% of P3 students are meeting or exceeding reading benchmarks as defined by MINEDUC/REB. This study provides new information about the amount and nature of instructional time in P2 Kinyarwanda classes and how that time is associated with student reading skills as well as teacher and student characteristics. Developed in collaboration with the Rwandan Education Board (REB) and a reference group consisting of Rwandan education experts, this research supports the Government of Rwanda in their efforts to advance their Education Sector Policy[8] and assist MINEDUC and REB to meet ESSP 3 Strategic Priority 1 "Enhanced quality of learning outcomes that are relevant to Rwanda's social and economic development" and associated Outcome 1.1 "All learners achieve basic levels of literacy and numeracy in early years." More specifically, the study aimed to help explain the current low levels of early grade Kinyarwanda reading outcomes and to inform Kinyarwanda curriculum development and teacher professional development policies and investments. Findings will also inform the USAID-funded Soma Umenye project interventions to improve reading outcomes. The study received research approvals from the Rwandan National Council on Science and Technology (NCST) and REB.

Current evidence about improving literacy outcomes indicates that many elements are required, including effective instruction, quality curricula that is supported by appropriate teaching and learning materials, as well as effective school leadership and parental engagement. Even with these elements in place and functioning, early grade reading outcomes can remain low if children do not have sufficient time to learn and practice fundamental reading skills.

Educational research over the past 50 years has increasingly recognized that the function of "time" is critical to learning and that time needs to be understood and studied in all its variant forms. The study of time in the educational context is intricately linked to understanding the quality and nature of what is happening in the classroom, including teachers' instructional focus and methods, students' focus and engagement and the overall interaction between teachers and students in the learning space. With increasing frequency, governments and their donor and research partners have endeavored to understand how much of the intended learning time is actually dedicated to learning, and what the quality of that time is with respect to producing learning outcomes for children. The research to date includes fairly straightforward measures of the amount of time teacher and students are at school and/or in the classroom together. More nuanced studies attempt to gain a deeper understanding of how time is spent[9] in the classroom, and even more complex and multi-dimensional studies include measures of constructs such as the "opportunity to learn,"[10] which includes time as one element. Importantly, a select few of these studies have begun to link data collected on the amount and nature of learning time to student learning outcomes so that associations can be made between the two to inform policy and practice more directly.

---

[8]Republic of Rwanda, Education Sector Policy, 2003. http://mineduc.gov.rw/fileadmin/user_upload/EDUCATION_POLICY.pdf

[9] For example, the Stallings Observation System, https://www.worldbank.org/en/programs/sief-trust-fund/brief/the-stallings-classroom-snapshot,

[10] Schuh Moore, A., DeStefano, J., Adelman, E., FN (2030). Opportunity to Learn: A High Impact Strategy for Improving Educational Outcomes in Developing Countries. USAID. https://www.epdc.org/sites/default/files/documents/EQUIP2%20OTL%20Book.pdf

Several cross-national studies have been conducted on multiple aspects of time in the educational context, including the World Bank's Service Delivery Indicator[11] assessments. These assessments collect national-level data on both scheduled and actual learning time to understand potential drivers of poor learning outcomes in developing countries. The findings have been staggering; with large gaps between scheduled school time, time when both teachers and students are in the classroom and time when students are engaged in effective learning.[12] For example, in Kenya the scheduled daily teaching time was 5 hours 41 minutes, but after deducting time for teacher absence at school, absence in the classroom, and time in the classroom but not teaching, the average daily teaching time was just 2 hours and 25 minutes. Calculated over time, this amounts to less than four years of education over the course of an eight-year primary cycle, wasting both government resources and students' learning opportunities. Other countries in eastern Africa such as Uganda and Tanzania have even more striking gaps in scheduled teaching time versus actual teaching time, with a ratio of 7 hours 13 minutes of scheduled time to an average of 2 hours and 56 minutes of actual teaching time in Uganda. These data have illuminated the prevalence of insufficient and inadequate learning time in school in many developing counties, and policy discussions about school schedules, interventions to reduce teacher and student absenteeism, and strategies to increase learning time in the classroom have ensued.

As the aspirations for research have grown to include more nuanced understanding of what is happening in classrooms, classroom observations have become even more critical to improving instruction and learning outcomes. This is especially true given that other proxy measures, such as teacher characteristics (e.g., qualifications, experience, pay),[13] school infrastructure variables, etc., have not been shown to consistently predict the quality of instruction or learning levels. Researchers have acknowledged that "classroom-level 'educational production' remains something of a 'black box', partly because of data limitations but also because 'what matters' depends so heavily on the setting and context."[14] The challenge of determining and measuring "what matters" in the classroom is further complicated by the typical approach of developing a "normative framework for quality judgements," which often carries with it inferential judgement by the observer to "score" the nature or quality of the activities, teacher methods, etc. The inferential nature of many classroom observation frameworks and instruments brings into question the reliability and validity of observation data, especially when observations are conducted by single observers. More rigorous research projects will assess inter-rater reliability[15] and validity[16] of observations during training and instrument piloting. Some research designs include video recordings for a subsample of lessons

---

[11] World Bank. Service Delivery Indicators. http://datatopics.worldbank.org/sdi/

[12] Filmer, D. "System health at the frontlines: Using SDI Data to analyze education quality in SubSaharan Africa", presented at Making Systems Work: A Global Conference on Education Systems Sydney, March 1, 2016. Stable url: http://pubdocs.worldbank.org/en/812821457978473769/Session-8-Deon-Filmer.pdf

[13] Hanushek., E., Rivkin, S., The Distribution of Teacher Quality and Implications for Policy Annual Review of Economics 2012 4:1, 131-157 and "Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps and Policy Implications" (with Karthik Muralidharan), in E Hanushek, S. Machin and L. Woessmann, eds., Handbook of the Economics of Education, Volume 5. North Holland/Elsevier. 2016.

[14] Rolleston, C. et al "Unlocking the Black Box: To what extent are interactive classrooms effective classrooms in the Indian context?" (2018). https://www.slideshare.net/YoungLivesOxford/unlocking-the-black-box-whats-happening-in-more-effective-classrooms-in-india

[15] Inter-rater reliability measures the rate of agreement between two observers collecting data on the same subject/respondent.

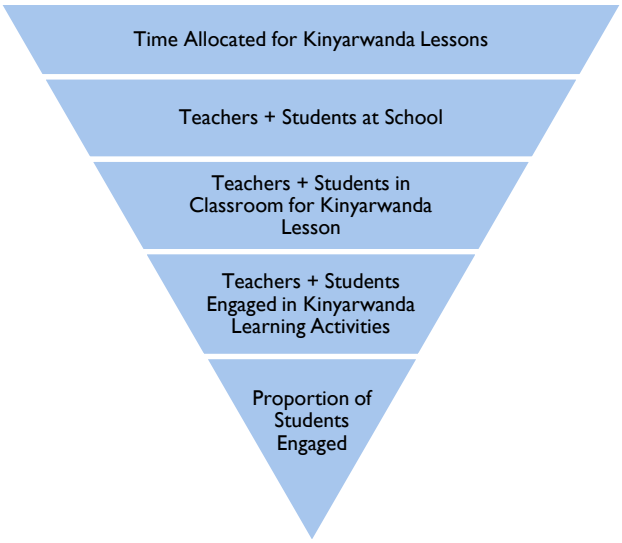
[16] In this discussion, validity refers to accurately measuring what an instrument is intended to measure.

during data collection that can be later scored by another observer[17] and spot checks in the field by supervisors, but these methods are still not widely utilized or comprehensive in nature. This has led to broad questions about how precise observation data are, reflected in the fairly common accepted standard of 80% inter-rater reliability for classroom observations. In most instances, researchers and project evaluators do not have the data to determine what the inter-rater reliability is during full data collection, and as such, classroom observation data have been difficult to trust.

We have endeavored to address these critical aspects of collecting reliable, detailed, and unbiased data at the classroom level by developing an observation instrument with minimal subjective/inferential properties and structuring the data collection with paired observers to verify our data are reliable. The instrument is highly focused on understanding reading instruction in detail and we have connected time and observation data to student reading scores, student characteristics, and a comprehensive profile of teachers.

In keeping with other time study approaches, we have defined "time" as a multi-dimensional construct, collecting data at multiple levels including: allocated time, in-classroom time, and learning time. Given this study is focused specifically on Kinyarwanda reading lessons, we have included very detailed information about the activities taking place in classrooms, disaggregated by reading skill and type of pedagogical approach. There is further discussion about the definitions and calculations used for this study in Annex D.

- Allocated time = the scheduled time in the curriculum for Kinyarwanda reading lessons.
- In-classroom Time = the amount of time teachers and students are in the classroom together.
- Learning time = the amount of time students are engaged in reading instruction or reading activities that build the fundamental skills children need to learn to read.



[17] These later video checks can be helpful to improve instrumentation and training but are typically not implemented in time to improve ongoing data collection. In some instances, video alone is used to code classroom observations.

The research questions that are addressed by this study include:

- What is the allocated time for P2 Kinyarwanda lessons per week and how is that time structured?
- What is the in-classroom time for P2 Kinyarwanda lessons per week and what is the nature and quality of that time?
- What is the learning time for P2 Kinyarwanda lessons and what is the nature and quality of that time?
- Who are the teachers leading P2 Kinyarwanda classes, how do they perceive their jobs, what kind of support do they receive and what are they looking for to do their jobs more effectively?
- What is the association between amount and type of in-classroom and learning time and reading skills for P2 students?
- For the questions above, how do the findings vary across strata of schools (e.g., urban, rural, low-performing schools, high-performing schools)?

TABLE 2: SUMMARY DESCRIPTION OF STUDY

	Early Grade Kinyarwanda Instructional Time Study
Policy Focus	Delivery of ESSP3, improving early grade Kinyarwanda instructional time/quality, and Kinyarwanda subject teacher professional development/coaching.
Subject Focus	P2 Kinyarwanda reading lessons (as indicative of P1-P3 classes)
Research Questions	Focus on understanding the amount of instructional time for early grade Kinyarwanda lessons on average and across higher and lower performing classrooms; the nature of reading instructional activities in the classroom, the quality of teacher: student engagement and the connection with student reading outcomes.
Sample	One P2 classroom in 100 schools observed twice (200 P2 classroom observations); 100 Kinyarwanda teachers surveyed and 1,600 students assessed and surveyed.
Data Collection Focus	Primarily quantitative: school data, classroom data, P2 Kinyarwanda lesson observations, teacher survey, student survey, and reading skills assessment.

SECTION II: STUDY DESIGN AND METHODOLOGY

OVERVIEW OF STUDY DESIGN

To address the research questions above, we fielded four data collection instruments over the course of four weeks during Term 3 of the 2018 school year.

We conducted two observations of one P2 Kinyarwanda classroom/class section of students in 100 schools across all five provinces.

We surveyed 100 teachers who led the instruction in the observed classrooms.

We assessed the reading skills of and collected profile data for the 1,600 students attending the Kinyarwanda classes we observed.

We collected administrative data from the 100 schools.

The analysis focused on understanding the amount of time teachers and students are in the classroom together (in-classroom time) and how they spend that time together (learning time, non-instructional time, etc.). The original design called for all of the observations to be conducted by paired observers but given changes in the school exam calendar during data collection, 60% of observations were conducted by single observers. The inter-rater reliability between paired observations was high (94% raw agreement at the teacher heading level).[18] We also analyzed the amount of time spent across different reading skills and student engagement levels. These time-related data were analyzed in the context of student reading scores, teacher characteristics and attitudes, and student characteristics. This analysis was conducted for a set of 85 randomly selected schools and for four case studies of five schools each (urban high and low performing and rural high and low performing).[19] These case studies allow us to understand more about classrooms in similar contexts that are delivering different outcomes for students.

SAMPLE

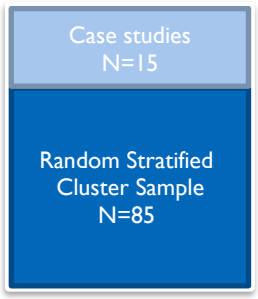
The total sample included 100 schools, with one P2 Kinyarwanda classroom/teacher per school and 16 P2 students per observed classroom. The total sample size was driven by a combination of factors, primarily the priority of visiting each sampled classroom twice, maintaining a small cohort of observers to increase the reliability of observation data, and the need to finish data collection by the end of the 2018 school year.

[18] Four levels were captured in the classroom observation which are described in the instrument description section.

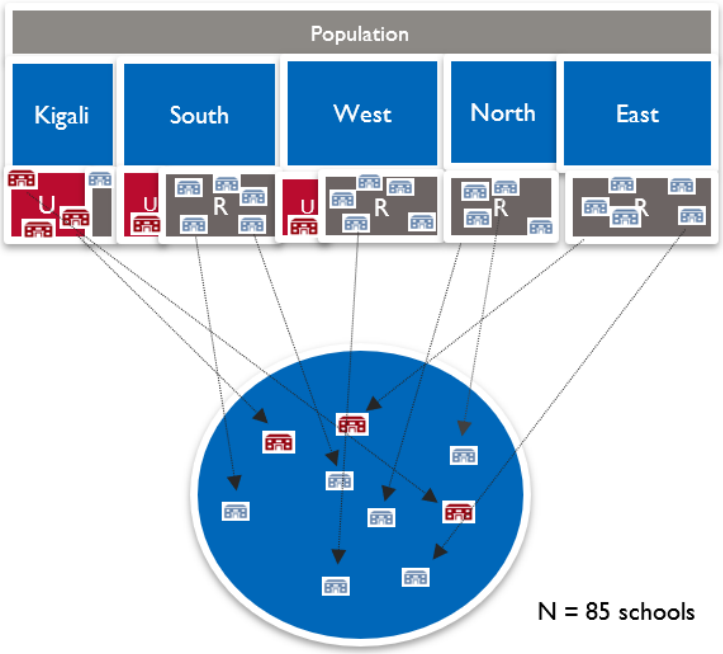
[19] There are five schools in both the 85 randomly selected schools and the 15 purposefully selected case study schools for a total of 100 schools.

The sample of students allows for a confidence level of greater than 95%, indicating there is a 95% probability that the population proportion is between our confidence intervals and our sample is very precise with respect to student data. The sample at the classroom level is proportionately allocated across provinces (with an oversampling of urban schools). A full discussion of the precision of the classroom and student samples can be found in Annex B.

The sample of 100 schools was divided into two sections with 85 schools randomly selected and 15 schools purposefully selected for case studies. The random sample was pulled from 15 randomly selected districts with the number of schools proportionate to the district school population. If sampled proportionately, only eight urban schools would have been selected and as such, urban schools were oversampled under a stratified design resulting in a total of 18 urban schools and 67 rural schools included in the random sample. The random sample of schools was clustered at sector level with two schools selected per sector to facilitate data collection at two schools per day. Annex A describes in detail the sampling procedure.



Time tables were collected for the 100 schools selected. The data collection firm selected the P2 class section to be observed with best efforts to stratify observations across days of the week and hour of the day. Each classroom selected was observed twice over a several-week period and across different days of the week and hours of the day.[20] All schools in the sample are government or government-aided schools. The random sample excluded schools selected for the 2018 national EGRA. Changes in the exam and end-of-school schedule were introduced early in data collection. As a result, adjustments were made to the field plan, and six replacement schools were required to ensure the team was able to conduct two observations for each class and complete fielding all data collection.



[20] See Annex A and C for additional details about the data collection timing, changes in the school calendar, and replacement schools.

The 15 schools for case studies were originally selected using the results of the LARS reading outcome data, with the highest and lowest performing urban schools and the highest performing rural schools per province selected to serve as case study schools. The case study sample design was subsequently updated to include a total of 20 schools clustered into four case studies (urban high and low performing and rural high and low performing) of five schools each and the selection was based on reading scores from this study across the entire sample (not per province). Further discussion of the case study methodology can be found in Annex A.

INSTRUMENTS

This study's design relied upon the use of four survey instruments: (i) classroom observation for the P2 class section of interest, (ii) teacher survey for the Kinyarwanda teacher of the sampled P2 class section; (iii) student profile and reading assessments for P2 students from the class section of interest; and (iv) school profile for the head teacher or any relevant school authority. All instruments (with the exception of the reading assessment[21]) were developed in English, translated to Kinyarwanda and then independently back-translated into English. Areas of discrepancy between translations were fully reviewed and reconciled. The classroom observation, teacher survey, and school profile were coded in SurveyCTO and the student reading assessment and student profile were coded in Tangerine.

The key constructs of each instrument are discussed below. Additional information about instrument development, piloting, reliability testing, and administration can be found in later sections and in Annex H-L.

THE CLASSROOM OBSERVATION INSTRUMENT collected information about the amount of time the teacher and students are in the classroom together and the activities they are engaged in to answer:

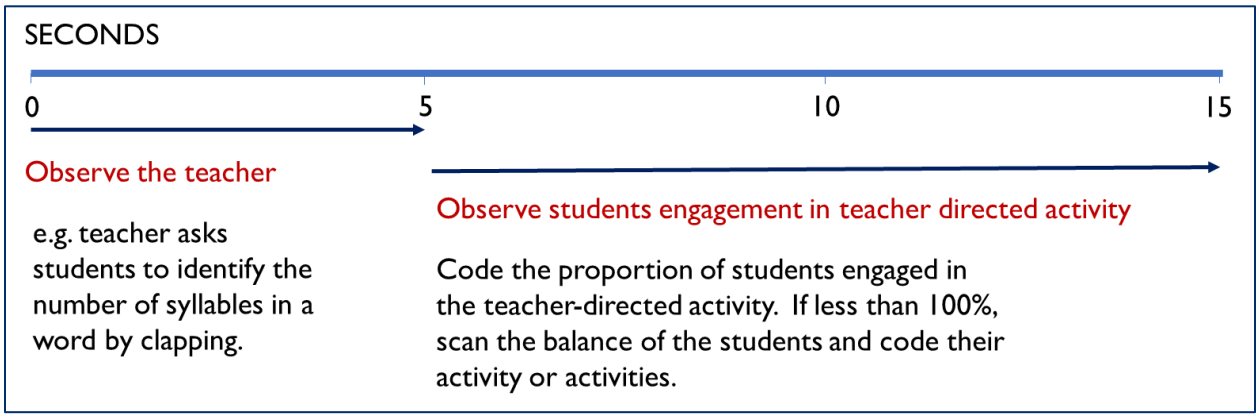
- How much of the allocated time are teachers and students present in the classroom?
- How the teacher is spending instructional time across skills: listening comprehension, phonological awareness, letters phonics, syllables and words phonics, word reading, sentence reading fluency, vocabulary, reading comprehension, and writing?
- What proportion of students are actively engaged and in what activity?
- How much time is spent on non-learning activities?
- What is the classroom management, student engagement and classroom climate like?

The structure and protocol of the classroom observation instrument is modeled after the Stallings observation instrument and deploys a snapshot approach to observations, recording an observation every 180 seconds over the course of a 40-minute Kinyarwanda lesson (average of 13 intervals per Kinyarwanda lesson) or until the class ended. The observation data were collected using a tablet.

The actual observation of teacher-directed activity and student activities and engagement levels occurs in the first 15 seconds of each 180 second interval. Figure 6 displays an example of a 15-second scan of the classroom. The balance of the interval time is used for coding, noting teacher-directed activity by hand in observers' own words (for quality checks and reconciliation of coding), and recording late student arrivals. Each classroom included in the sample was observed twice in unannounced visits over the course of several weeks.

[21] Reading assessment items were developed and validated in Kinyarwanda, so no translation was required.

FIGURE 6: 15-SECOND OBSERVATION SCAN



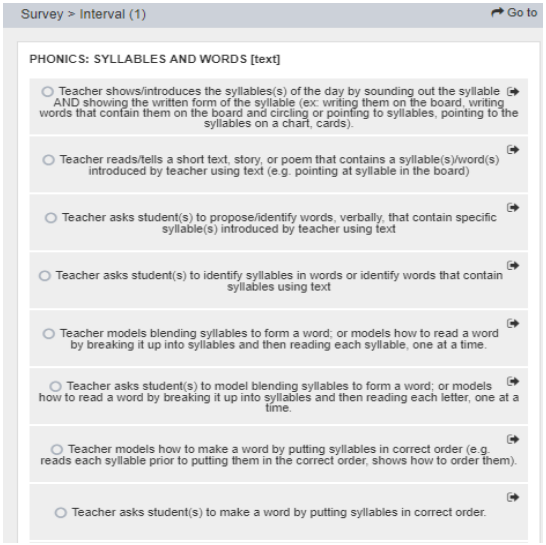
The instrument is organized into four general levels: Teacher Heading (category of teacher-directed activity), Teacher-Directed Activity (specific activity the teacher is leading), Student Activity (specific activity the students are engaged in), Student Engagement Level (proportion of student engaged in any specific activity).

From the Teacher Heading menu, the observer selects either the reading skill the teacher is focused on or the General Teacher category. The latter category includes a range of items, including transition (some level of transition time is necessary for instruction) and non-instructional time. Examples of transition time include: Teacher is setting up classroom/handing out materials/preparing exercise; Teacher is restating/asking questions about a previous lesson; and Teacher is introducing the lesson or activity and providing students with instructions. Examples of non-instructional time include: Teacher is doing administrative paperwork in classroom; Teacher is not in the classroom; and Teacher is in the classroom but distracted and not engaged.

If the teacher is directing reading-focused activities, the observer selects the reading skill being focused on and then the specific activity. Examples of teacher-directed reading instruction activities include:

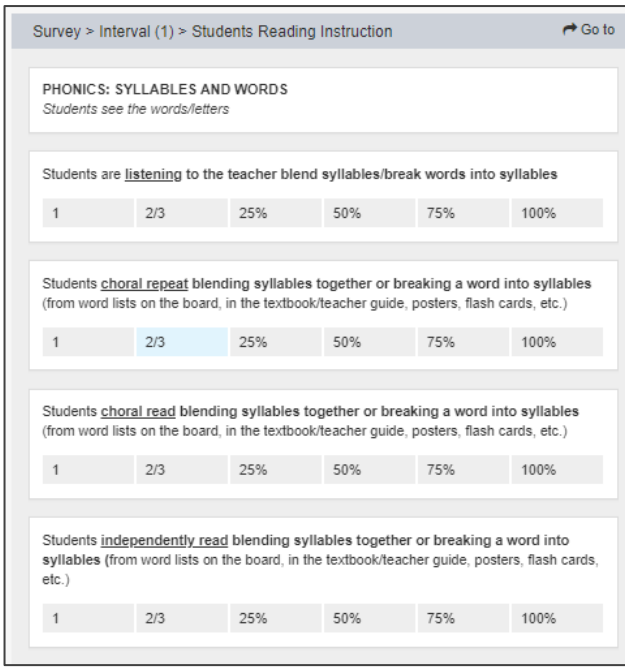
- Phonological awareness: Teacher asks students to propose words, orally, that contain specific sounds or syllables (either from students' memory or by looking at pictures of objects)
- Phonics letters: Teacher is showing/introducing the letter of the day
- Reading words fluently: Teacher models how to read a word fluently (syllables are not pronounced/distinguishable) (From word lists on the board, in the textbook/teacher guide, posters, flash cards, etc.) and asks students to choral repeat, read, or read independently
- Comprehension: Teacher asks students to summarize or explain in their own words what a sentence or text is about.

FIGURE 7: TEACHER-DIRECTED ACTIVITY – PHONICS: SYLLABLES AND WORDS



The classroom observation instrument also collects data about student activities and the level of engagement in those activities. Student activities are also organized into learning time (active/passive/assessment), transition time, and non-instructional time. See Figure 8 for examples of student learning activities under phonics: syllables and words. Examples of student transition activities include: Students are copying homework assignment from board/writing teacher's dictation of homework into notebook; and Students are preparing for the lesson as directed by the teacher (getting notebooks/texts, rearranging desks). Examples of student non-instruction activities includes: Students are engaged in social interaction with other students; and students are inattentive (head down on desk, off-task) and unengaged.

FIGURE 8: STUDENT ACTIVITY – PHONICS: SYLLABLES AND WORDS



The observer first codes the student activity that corresponds with the teacher-directed activity and the proportion of students engaged in that activity (1, 2/3, 25%, 50%, 75% or 100%). If less than 100% of students are engaged in the teacher directed activity for that interval, the observer then determines the other activities (and proportion) that students are engaged in. There are a total of 109 possible student activities included in the instrument, which constitute a comprehensive item bank that includes most of the activities/tasks given to students in Kinyarwanda classes.

In analysis, student activities selected by observers are categorized by reading skill or into non-instructional, transition and assessment/homework categories. All student reading activities selected are further categorized as "active" or "passive" activities to facilitate analysis of instructional practices and the relationship to student engagement and reading outcomes. For example, student activity "Students are listening to the teacher blend syllables/break words into syllables" is categorized as a passive activity while "Students independently read blending syllables together or breaking a word into syllables" is categorized as an active activity. All activities under each reading skill (teacher-directed or student level) are categorized as active or passive.

Following the lesson, the observer codes summary information about the classroom, overall teacher-student interactions, materials availability and use, etc. Example items include:

- Teacher calls on a variety of students during the lesson
- Teacher makes an effort to engage inattentive students
- Teacher's writing on the blackboard was legible and visible for all students

Teachers were also briefly interviewed after the observation about use of lesson plans, estimations of that class' reading skills, etc.

The following table outlines the inter-rater reliability results during data collection (60% of observations conducted with paired observers) across levels of the observation data: teacher heading, teacher main activity, student main activity and student corresponding activity. The literature indicates 80% agreement is an acceptable inter-rater reliability (IRR) for relatively simple classroom observation instruments and as such, we feel confident in the reliability and validity of our observation data given these IRR scores. See latter sections and for additional explanation and inter-rater reliability results for each pilot/field test phase.

TABLE 3: INTER-RATER RELIABILITY SCORES FOR CLASSROOM OBSERVATION INSTRUMENT DURING DATA COLLECTION

Scores	Teacher Heading	Teacher-directed Activity	Student Corresponding Activity
Agreement	0.938	0.872	0.863
Expected Agreement	0.184	0.044	0.042
Kappa	0.924	0.866	0.857

THE TEACHER SURVEY INSTRUMENT collects information about the teachers associated with the classroom sections observed in the study. The study aimed to provide a nuanced understanding of teachers to connect their classroom practices with characteristics such as experience, education, attitudes, etc. The teacher survey includes the following areas of measurement:

- Background profile (e.g., gender, age, number of years teaching, academic qualifications, non-teaching jobs, transportation)
- Perceptions of student reading skills
- Perceptions of teaching
- Lesson preparation and administrative tasks
- Motivation related to Kinyarwanda lesson planning
- Individual self-efficacy related to reading instruction and class management
- Support and feedback received
- Expectations of student learning capacity

THE STUDENT PROFILE was administered after student assent was received and in keeping with protocols for the privacy and comfort of young respondents. Students provided information on the following subject areas: basic demographics, household assets/infrastructure, availability of reading materials at home, support for schooling at home, prevalence of tardiness and absence, and the reasons associated with tardiness and absence.

THE STUDENT READING ASSESSMENT was modeled after the Early Grade Reading Assessment yet was designed explicitly to address the questions in this study, rather than provide representational data about P2 student reading skills in Rwanda. The reading assessment design strategy emphasized obtaining a large distribution of scores and ensuring a large percentage of non-zero scores to bolster the opportunities for regression analysis. Reflecting on the 2018 LARS and 2017 P1 EGRA results (high percentage of zero scores), we decided to include P1-P3 items (with variations across sub-tasks, see

table below) to minimize both floor and ceiling effects. Assessing students at levels that allow more students to score above zero offers a greater distribution of scores to compare with time, a larger sample of non-zero scores and as a result, narrower confidence intervals.

TABLE 4: READING ASSESSMENT SUB-TASKS AND LEVELS

Sub-task[22]	Level	# Items
Letter Identification	P1	20
Syllable Sounding	P1	20
	P2	30
Familiar Word Reading	P1	10
	P2	10
	P3	10
Passage Reading	P1	5 sentences, 25 words
Comprehension Questions	P1	5 questions

Twenty students (10 male/10 female) were randomly selected from each observed classroom (this included four replacements in the event students declined to participate) for assessment. Administration of reading assessments took place after the initial classroom observation to prevent revealing which P2 classroom was included in the study prior to administration. This mitigated the risk of schools paying special attention to teacher and student absence or tardiness for the observed Kinyarwanda lessons.

The reading assessment was administered in keeping with the typical EGRA protocol[23] with a few exceptions that were intended to align with the protocols being used for the national EGRA in Rwanda. Specifically, students were allowed 180 seconds to the read the passage, with time being marked at 60 seconds to facilitate two methods for calculating correct word per minute (based on 60 seconds and 180 seconds). Both are reported in the reading assessment findings, but in keeping with the 180 second passage reading timing for the Rwandan national EGRA, all regression analyses utilize students' scores at 180 seconds for calculating correct words per minute.

[22] Assessment items were based on the metrics and assessment items produced at the July 9-14, 2018 Soma Umenye-organized workshop with MINEDUC, REB, URCE, and other stakeholders. Several items were updated after piloting and testing and the passage and questions were updated using passages and questions from the national EGRA pilots. See Annexes H-L for instruments and further discussion.

[23] See RTI International (2015). Early Grade Reading Assessment (EGRA) Toolkit, Second Edition, Washington, DC: United States Agency for International Development.

THE SCHOOL PROFILE is administered for each school in the sample and utilizes school records to collect information, primarily focused on the following:

- Number of P1-P3 Kinyarwanda sections
- Number of P2 teachers
- Absence rate of P1-P3 Kinyarwanda teachers
- Absence rate of P2 students in observed class
- School facilities (e.g., library, electricity, working latrines)

These data are used to complement the classroom observation data and to offer summary statistics about the schools in our sample.

SECTION III: FIELDING THE STUDY

TRAINING AND FIELD PREPARATION

Laterite, a firm based in Kigali, was selected through a competitive process to serve as the data collection firm for the study. Prior to the start of field preparation activities, Laterite trained a team of 22 enumerators selected from its pool of experienced field staff to carry out field preparation activities, which included obtaining local authorities' approvals at the district and sector levels and confirming school-level data (e.g., location, time table).

General training was provided to all enumerators and supervisors initially, including;

- Background of the project and the research objectives and questions.
- Research methodology: The study area and sampled schools.
- Field preparation protocols: Specific guidelines for each phase of field preparation.
- Staffing: Team structures and responsibilities.
- Ethics: Fieldwork/calls etiquette and professional code of conduct.

Enumerators were grouped to focus on specific instruments and training was tailored for each enumerator group. The team of 11 observers (administering the classroom observation tool), 17 assessors (administering the reading assessment and student profile tool), 10 surveyors (administering the teacher survey and school profile tools) and 6 field supervisors were trained over a period of 15 days. Training included reviewing the logic of the questionnaires and potential risks and bias. The following approaches were undertaken during instrument-specific practice sessions:

- Reading assessments: Enumerators referred to reading assessment videos provided by Room to Read to gain a better understanding of how they are administered and identify best practices. Enumerators conducted several rounds of practice with pre-established scripts to assess the consistency in their scoring. Practice sessions focused on the following exercises: (i) the field supervisor plays the role of a student while enumerators score separately, (ii) two enumerators administer the reading assessment to one enumerator playing the role of a student and score separately, (iii) enumerators take turns playing the role of a student and assess each other in pairs.
- Classroom observation: Enumerators performed the following exercises to familiarize themselves with various instructional and non-instructional activities performed by the teacher and students: (i) identification (in their own words) of teacher and students' activities within each heading and; (ii) matching a provided set of teacher and students' activities to relevant headings. The team also conducted several rounds of practice by observing classroom videos, coding the teacher-directed activity along with students' activities and corresponding levels of engagement prior to transcribing on the comment sheets what was happening in class during each interval. Synchronization checks were performed throughout the observation exercise by field supervisors and an assessment of codes' similarity was conducted at the end of each classroom video. Areas of discrepancy were discussed and clarified during each practice session. A knowledge assessment focused on the correct coding of reading skills was administered to all trainees and results reviewed and discussed in group format.
- Teacher survey and school profile: Enumerators practiced both surveys through exercises involving: (i) the self-completion of surveys and (ii) mock interviews in pairs.

The training activities were punctuated with two instrument pilots and two field tests in advance of data collection. After each pilot and field test, all enumerators participated in comprehensive group de-briefs with Laterite supervisors and the researchers to reflect on and improve the instruments, protocols, and field organization.

INSTRUMENT PILOTING, FIELD TESTS AND ADAPTATION

All survey instruments and associated administration protocols were piloted/field tested in a total of 24 schools located within Muhanga District. Pilot and field testing activities were run under real conditions over four separate days (September 7th, 18th, 28th and October 2nd).

The main objectives of the pilots were to (i) collect feedback from the team and respondents on surveys and associated protocols (understandability and relevance of questions, unclear translations, coding errors, lists of options) and, (ii) familiarizing enumerators with the data collection setting. Pilot and field testing activities also aimed at testing aspects of field surveying such as (i) interview length for the teacher survey, school profile, student profile, and reading assessments and; (ii) the feasibility of visiting two schools per day. The number of surveys conducted for each round of pilot is as follows.

TABLE 5: SURVEYS PER PILOT ROUND

Pilot	Date	Schools	Classroom Observations	Reading Assessments	Teacher Surveys	School Profiles
Instrument Pilot 1	September 7th	11	36	48	28	11
Instrument Pilot 2	September 18th	12	38	60	29	11
Field Test 1	September 28th	5	24	206	0	0
Field Test 2	October 2nd	10	33	92	21	8
TOTAL		**24**	**131**	**406**	**78**	**30**

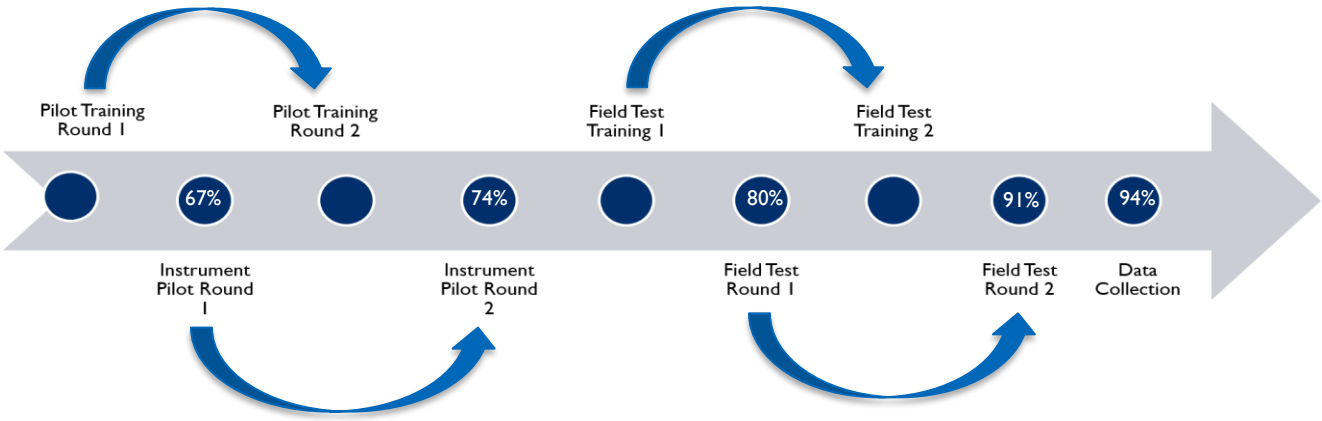
The team engaged in debriefing sessions after each round of pilot to provide feedback and list any issues that occurred during interviews, reading assessments and classroom observations. Issues were then recorded, and appropriate updates were made to instruments, protocols, and training. Updates to instruments during piloting and field testing primarily focused on the following:

- Classroom observation instrument: additional items that were more detailed and specific to minimize interpretation and "other" coding by observers; fine-tuning observation and coding timing and protocol to facilitate synchronization between paired observers.
- Teacher survey: shortening the survey based on initial admininstration timing; updating translations and enumerator explanation to faciliate comprehension of questions; adding response options based on teachers' responses.
- Student reading assessment: refinement of enumerator instructions, updating of passage and comprehension questions for passage, and questions with higher reliability.
- Student profile: shortening the survey based on initial administration timing; adding response options based on students' responses; clarifying terminology/translations.

- School profile: clarfiying data and source required.

The classroom observation instrument underwent the most significant changes as a result of the pilots and field testing, with the majority of changes occurring based on debrief sessions following the first two pilots. We tested several different approaches to tracking interval times and observing the teacher and student activities. The feedback from observers was consistent in the desire for more precise items to reflect all the activities that occur in Kinyarwanda classrooms. Given that observations were coded on a tablet and that the software allowed for a cascading menu, we were able to add items without creating undue complexity for observers. Figure 9 below depicts the rounds of instrument pilots and field tests with inter-rater (raw agreement) reliability scores for paired classroom observations.[24]

FIGURE 9: CLASSROOM OBSERVATION INSTRUMENT INTER-RATER RELIABILITY



DATA COLLECTION

OVERVIEW

At the outset of data collection, the team was notified that the school calendar had changed and that exams and the end of the school term would start several weeks earlier than originally scheduled.[25] Given the need to complete data collection (with two classroom observations for each classroom), the research team assessed the options and decided on two strategies to ensure data collection was completed before exams interrupted lessons. First, the data collection firm contacted all sample schools to confirm their exact timing for exams and revised the field plan to prioritize visiting schools with the earliest exam schedule. Second, the research team reassessed the observation approach to determine if the original plan for all 200 observations to be conducted in pairs was still viable given the new school

[24] All pilot and field test cycles included pair observers to test inter-rater reliability for classroom observations and reading assessments. Percentages shown in blue circles are the raw agreement between observers for teacher heading.

[25] REB's decision, communicated in a memo dated 8-9th of October 2018, brought forward the end of term from the 23rd of November 2018 to the 9th of November for schools in the primary level.

calendar. A decision was taken to split the pairs (single observer per classroom) for the remaining 60% of the observations. This decision was based on the high inter-rater reliability data for the 79 paired observations that had already taken place (in addition to the high inter-rater reliability results during the final field test), resulting in high confidence that having paired observations was not necessary.

In summary, and including replacement schools, the final data collection included:

- 286 classroom observations
- 79 observations were conducted by paired observers with agreement between the observers for teacher heading at 94% across 2,054 observation points (~13 intervals per class).
- 104 teacher surveys
- 104 school profiles
- 1,664 reading assessments with P1-P3 leveled items and student profile.

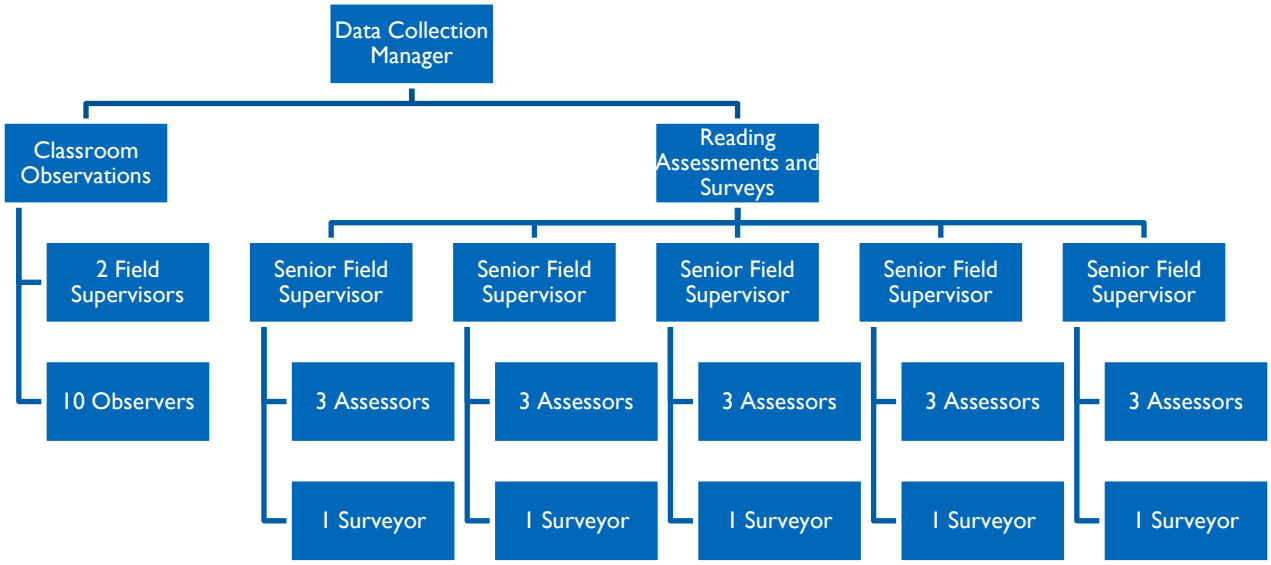
FIELD TEAM STRUCTURE

The structure of the data collection team was as follows: Laterite used three separate teams (observers, assessors, and surveyors) to ensure sufficient staffing capacity for efficiently conducting multiple data collection activities. Laterite selected enumerators for this study from its pool of experienced field staff based on (i) their performance on past projects with Laterite; (ii) their experience working with Laterite on education-related projects; and, (iii) their availability during the period of the research. Laterite staffed the reading assessments team with only female enumerators in line with their internal policy and best practice that girls below the age of 18 should always be interviewed by female enumerators.

The three teams of enumerators visited schools on three separate days to minimize disruptions. A team of observers was scheduled to conduct two school visits in pairs and carry out the classroom observation on the same class. As noted above, this plan was altered when exam schedule changes were introduced part way through data collection and we transitioned to single observers for 60% of the observations. A sub-team of assessors and surveyors (four assessors and one surveyor each) was scheduled to conduct a separate school visit and carry out 16 students' reading assessments, one teacher survey and one school profile.

Each sub-team of assessors and surveyors was led by one field supervisor/senior field supervisor while observers were led by two field supervisors. The supervision team was in the field throughout the duration of the study and their main responsibilities consisted of (i) monitoring the conduct of field operations and troubleshooting issues, (ii) assessing the team's conformity to the instruments' protocols and adherence to the study professional and ethical code of conduct, (iii) leading debriefing sessions and providing regular feedback on areas of improvement, and (iv) reporting issues.

FIELD TEAM STRUCTURE



CLASSROOM OBSERVATIONS BY DAY OF WEEK AND HOUR OF DAY

After adjustments for the changes in the exam schedule and school calendar, the final schedule for classroom observations is listed in the table below. The research design called for a distribution of observations across days of the week and hours of the day to capture variations and also in recognition that the Kinyarwanda curriculum is progressive throughout the week in terms of its focus on reading skills. For example, early in the week the curriculum calls for the introduction of a new story or content that tends to produce an emphasis on listening comprehension and phonological awareness. As detailed in the table below, the distribution of classroom observations is not balanced across day and hour, but that is in large part a function of how P2 Kinyarwanda classes are distributed across day and hour within schools' time tables. The distribution of classroom observations is roughly representative of the distribution of P2 Kinyarwanda class schedules across our sample schools, with some exceptions due to logistics (geographic clustering of observation visits) and revision of the field plan due to changes in the exam schedule and school calendar. School head teachers were notified that we would be visiting schools in the month of October, but all lesson observations were unannounced.

TABLE 6: CLASSROOM OBSERVATIONS BY DAY AND HOUR

Scheduled Time	Monday	Tuesday	Wednesday	Thursday	Friday	Total
7:20:00	13	1	4	3	5	26
7:30:00	0	2	0	0	0	2
8:00:00	3	7	7	0	5	22
8:40:00	5	7	3	6	3	24
9:00:00	0	0	0	0	1	1
9:30:00	0	0	0	1	0	1
9:40:00	9	5	4	4	3	25
10:20:00	5	2	10	6	5	28
11:00:00	0	6	0	7	5	18
11:00:00	0	0	0	0	1	1
12:40:00	7	8	2	4	3	24
13:20:00	7	5	5	6	5	28
14:00:00	6	2	4	6	1	19
15:00:00	4	9	5	6	3	27
15:40:00	4	3	7	3	8	25
15:50:00	0	2	0	0	0	2
16:20:00	4	4	4	1	0	13
Total	67	63	55	53	48	286

SECTION IV: DATA PREPARATION

DATA CLEANING ACROSS INSTRUMENTS

Running concurrently with data collection, the data cleaning was done using Stata 15. Though data for each survey instrument was cleaned separately, the following standard data cleaning operations was implemented on all data.

- Labelling all variables by executing a SurveyCTO generated do-file (Tangerine used for EGRA);
- Dropping all test/mock data by looking at the enumerator ID and submission date;
- Dropping observations/data from schools that were observed once and subsequently replaced during fieldwork.
- Assignment and cross-checking enumerator, school, teacher, and class ID's (within and across instruments and observations);
- Dropping cases of duplicate submissions;
- Translating other specified options and enumerators' comments into English;
- Anonymizing data by removing any directly identifiable information of participants;
- Dropping variables with no information/data;

- Calculating and analyzing summary statistics (mean, median, maximum and minimum values, standard deviation) for each numeric variable and inspection of plausible values.
- Tabulation of each categorical variables, inspection of plausible values, and assignment of new coding categories for answers with high frequency specified under 'other, specify.'
- Cross-tabulation of variables to ensure 'skips' questions were followed correctly and are consistent.

In summary, the final data sets include:

- 200 classroom observations (one P2 classroom in 100 schools observed twice)[26]
- 100 teacher surveys
- 100 school profiles
- 1,600 reading assessments with P1-P3 leveled items and student profile.

EGRA DATA CLEANING

The EGRA was administered using the Tangerine software and most of the cleaning processes were done manually. The EGRA data was downloaded daily as a CSV file and imported into Stata for cleaning. The following data cleaning operations were undertaken:

- Cross-checking that 16 assessments were administered and submitted per school;
- Ensuring that assessment protocols were respected by verifying that students were only administered segments of the assessments for which they were eligible given their response/performance in the preceding subtask;
- Calculating Cronbach Alphas for different sub-sample of students by subtask.

CLASSROOM OBSERVATION DATA CLEANING

TREATMENT OF OBSERVATION DISAGREEMENTS AND "OTHER" CODING

The classroom observation data set includes over 3,600 unique observation points, with 79 paired class observations and 121 single class observations, both with ~13 intervals per observation. Across the ~2,000 observation points from paired class observations, there were a total of 233 disagreements (6.42%) on coding between observers. The observation protocol includes observers hand writing, in their own words, the teacher-directed activity for each interval during an observation. These hand-written notes along with the original coding were entered into Excel for all observations with discrepancies between observers (supervisor notes were added when available). Researchers reviewed each discrepancy and finalized the coding based on observers' and supervisor's (as available) notes. This means all disagreements were resolved and one consolidated classroom observation, among the paired observations, was kept for analysis. Additionally, there was a total of 382 teacher and student activities across the ~3,600 observation points that were coded as "other." Using the data entered under 'other' alongside hand-written notes of observers, "other" items were recoded into the appropriate teacher or

[26] 100 P2 Kinyarwanda classes were observed twice. If teachers were absent during unannounced lesson observation visits, the lesson observation was re-scheduled.

student activity. Only one code was added to the questionnaire after data collection to account for an activity that was not included in the questionnaire.[27]

TIME VARIABLES

Several checks were performed on the data to ensure the amount of time was accurately measured: comparing the scheduled start time of the class, confirming whether the first interval was taken around 3 minutes after the start of the scheduled time, and confirming the end time of the lesson. Both the scheduled start of the class and the end of the class were collected automatically by the tablet and inputted manually by enumerators. In addition, in paired observations, this information was captured independently. This whole array of information was analyzed jointly, along with information on the schedule of classes collected through the timetables to deal with any discrepancies in timings. For example, a few class observations were conducted only after the teacher arrived (in two occasions where it was not clear whether the class was going to take place or not), instead of starting three minutes after the scheduled class start time. In this instance, intervals were adjusted so that start of the class reflected that time lost due to teacher tardiness.

Because all intervals were of length +/- 3 minutes, in analysis, each interval received a total of three exact minutes, and the last interval was then adjusted for the remaining time so that the sum of all intervals equaled the exact duration of the class observed. For this, the information collected on the first interval (taken starting minute 3 of each class, up to minute 6) was duplicated for the first 0-3 minutes where no observation is yet conducted. This allowed us to categorize the whole time of the class (from scheduled start to finish) to a learning/non-learning activity.[28]

ENROLLMENT/ATTENDANCE

Enrollment, attendance data and late arrivals are crucial variables employed in the calculation of amount of time lost due to student absence/tardiness. Enrollment and attendance figures were compared and reconciled. First, when discrepancies in enrollment data collected occurred between the first and second observations, pictures that had been taken on the tablets of the enrollment list were examined and used to reconcile these inconsistencies. Secondly, attendance numbers were capped to equal the number of students enrolled. For example, if the enrollment list had 52 students registered in the class, but 55 students were present in the classroom, attendance was then taken at 52. This happened when children from another shift or class sections were present in the class, a situation which was noted on certain occasions by fieldworkers in their class comments sections. In addition, corrections to the number of students arriving late were also done, which happened particularly when the whole class arrived late and enumerators did not manage to count with precision the number of children arriving late.

---

[27] 'Teacher is congratulating students /asking students to congratulate their classmate/ asking their grades/asking them to raise their hands based on the marks they received' (code1017) and corresponding 'Students are being congratulated and/or raising their hands based on the marks they received' (code S110).

[28] For example, if teacher is introducing the syllable of the day during the first interval, we assume the teacher was doing the same exercise during the first three minutes.

DEFINITIONS AND CATEGORIZATION OF TIME

As noted in the introduction, time has been recognized through multiple studies as a key element in the delivery of quality education. Across these studies there have been variations in how time is defined and at what levels it is measured. Time studies have typically focused on the amount of time allocated by the education system for instruction and how much of that time is lost due to teacher/student absence and tardiness. This study also begins the measurement of time with the amount of time MINEDUC and REB have allocated to P2 Kinyarwanda instruction (eight 40-minute sessions per week). This constitutes "allocated time", and is quantified proportionate to the number of students enrolled per class.[29] This means that any given class is measured in student minutes (SM) and the amount of SM can vary from class to class.[30]

We then measure the amount of time teachers and students are in the classroom for P2 Kinyarwanda lesson. For this, we discount the lost time due to teacher absence and time out of the classroom for each three-minute interval. For time that the teacher is fully absent and the lesson scheduled to be observed does not take place, we multiply 40 minutes times the student enrollment for that class. For the time that the teacher is either late or is not in the classroom (and the students are not engaged in any learning activity), we multiply the amount of teacher late/not in-classroom time by the number of students in the classroom at that time. For instance, if a teacher is 6 minutes late for class, the student enrollment is 40 children, but only 35 children are present for those 6 minutes, the lost time due to teacher not in classroom is calculated as 210 minutes (6 minutes x 35 students).

We collect data in three-minute intervals about the number of students who arrive late and include that as lost time for that student. At the end of the lesson, we count the number of students present and compare that to the student enrollment data in the classroom to identify the number of absent students. It was determined that the classroom register was more accurate and up to date than the register in the schools' front office, though we did collect student enrollment data from the front office on our observed classes for comparative purposes.

The sum total of teacher time in the classroom and student time in the classroom constitutes "in-classroom time" in this study (see Figure 10).[31] A graphical example of these time calculations is provided below for a given class of 50 students. Scenario 1 assumes that the teacher and all students are on time for class, while Scenario 5 assumes that the teacher is absent and that results in a loss of all in-classroom time. Scenario 2, 3 and 4 show the cumulative effect of student absence, teacher tardiness, and students' tardiness, respectively.

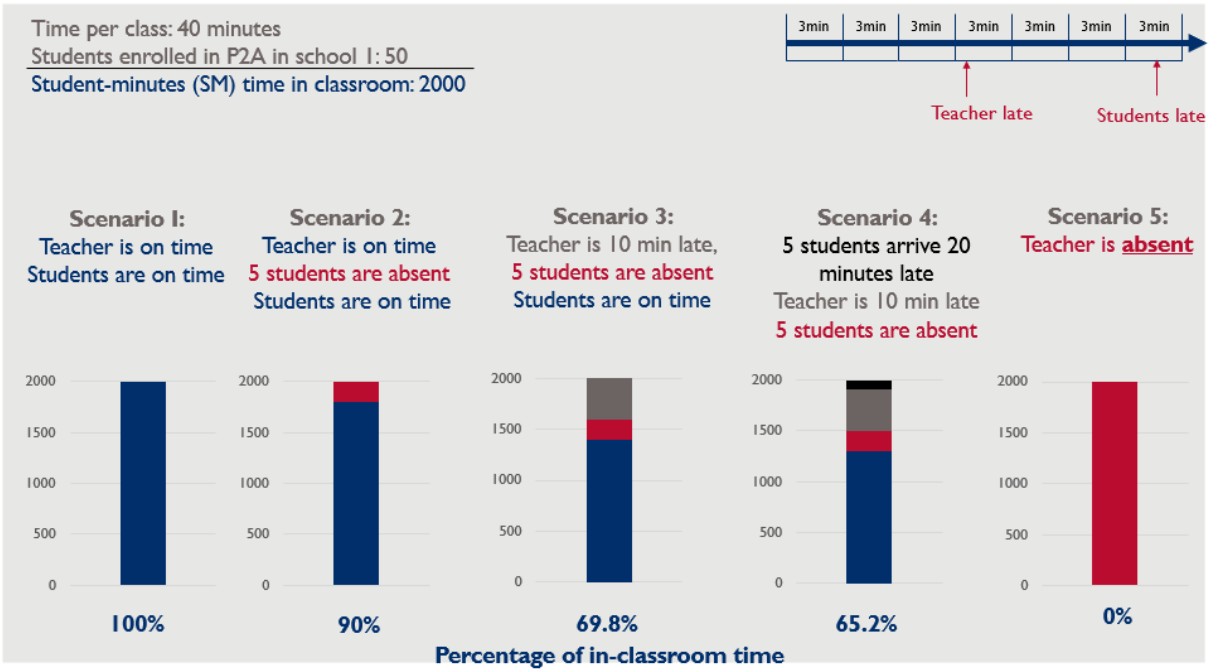
---

[29] This is equivalent to 354,400 student minutes of REB allocated time examined (40-minute session *x* 44.3 average class size/enrollment in sample *x* 100 class sections *x* 2 observations) in this study. In practice, the duration of classes varied from class to class, but given that the average duration of class was of 43.2 minutes,[29] time observed represented 108% of REB allocated time. As a result, we use the terminology of "observed" time throughout the report. See Section II for a full discussion of the classroom observation protocol.

[30] For example, in a class of 40 students, the time allocated for a lesson is 1600 SM, while that increases to 2000 SM in a class of 50 students.

[31] For the calculation of in-classroom time, we used the teacher and student absence/tardiness data derived from unannounced lesson observations even if that lesson was ultimately rescheduled for another time or day. For analysis of the composition of in-classroom time (e.g., what reading skills time was spent on), we took the time for all regularly observed lessons and those lessons we observed that were rescheduled.

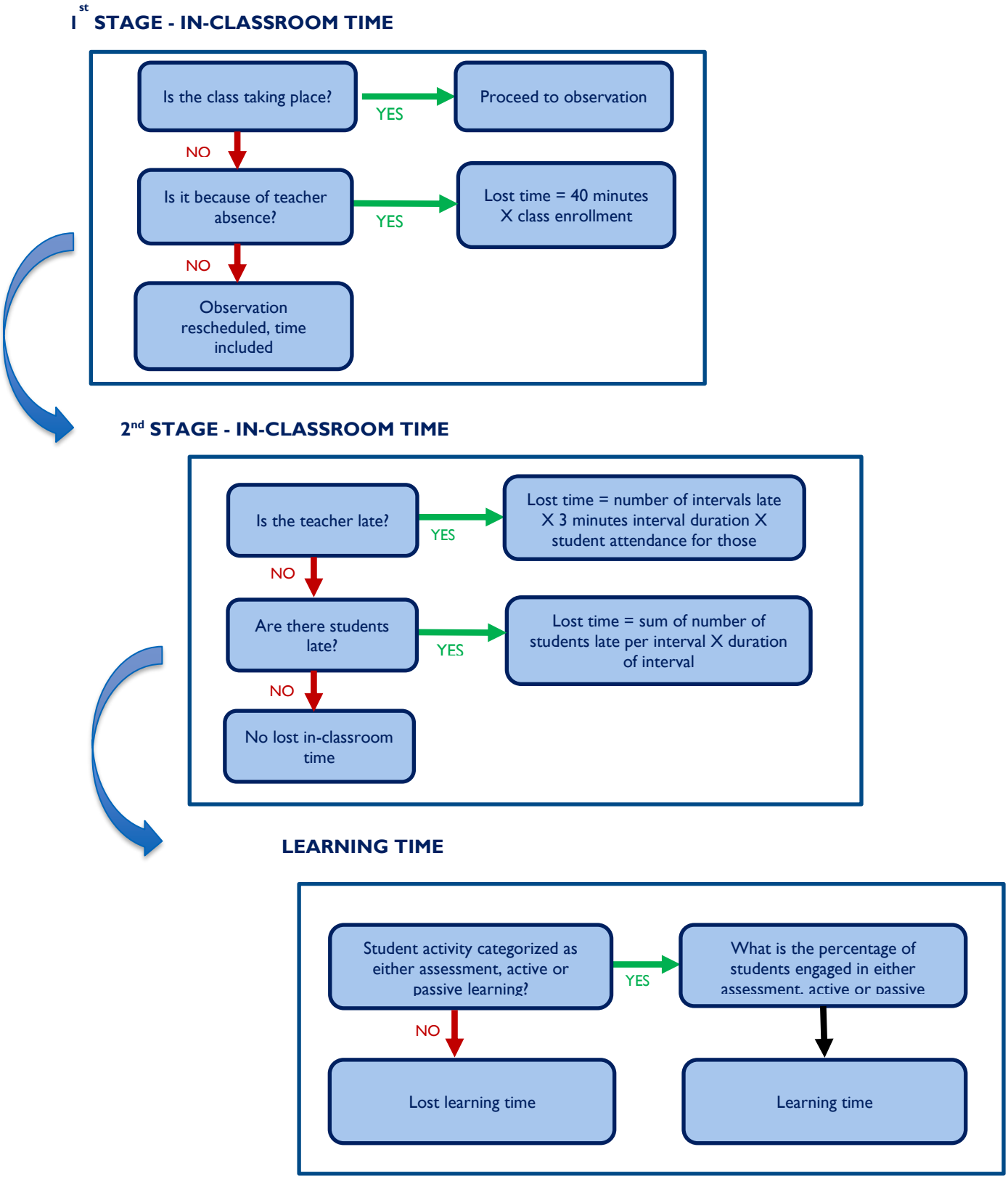
Example of Time Calculation (for a 40-minute class of 50 students):



We then move to measure how time is being spent when the teacher and students are in the classroom together. As noted in the instrument discussion section, there are a total of five categories of in-classroom time: learning time (active/passive reading activities), assessment/homework time, transition time, and non-instructional time.

Importantly, we utilize the student activities and level of engagement of students to calculate all categories of time within in-classroom time based on the student engagement levels. For example, in a classroom of 40 students where the teacher-directed activity is "introducing the syllable of the day" during a particular interval and 75% of the students are engaged and paying attention to the teacher while 25% are off-task (e.g., talking with another student, disengaged), we code the learning time as 90 minutes (the 30 students paying attention times 3 minutes –this is the interval time— and the non-instructional time as 30 minutes—the 10 students off-task time 3 minutes). The graphic below illustrates these steps:

FIGURE 10: CALCULATION OF DIFFERENT LEVELS OF TIME



SECTION V: STUDY LIMITATIONS

This study does have a range of limitations and the findings reported should be considered with these limitations in mind.

- The regression analyses that explore how different data are related and what factors are associated with student outcomes estimate the relationship between data at one point in time and do not make assertions about causality.
- The disaggregated findings (e.g., urban/rural, gender) do not have the precision that the overall sample does. Additional information about the precision of the study can be found in Annex A.
- The study included observations of one P2 Kinyarwanda class and in-depth surveys of one P2 Kinyarwanda teacher per school in the sample. As such, there may be limitations on the generalizability of classroom and teacher findings, especially for those 44% of schools with more than one P2 teacher.
- Schools were informed that they were part of a study but were not informed in advance about the P2 Kinyarwanda section that would be observed or the date of observations. Notwithstanding this, the section to be observed was made clear after the first observation so it is possible that teacher and/or student absence/tardiness was mitigated for the second observation (which was also unannounced).
- The presence of an observer in a classroom can induce the Hawthorne Effect, a phenomenon where people perform differently, and typically at their best, while being observed. While it is possible our observations induced this effect, there were significant efforts made to minimize the observer effect and to document when the teacher or students were distracted by the observer. There is no indication that there was substantive observer effects. Even if there are some observer effects, there is also value in knowing what teachers believe is optimal instruction and where the frontier of their knowledge and skills exists.
- The student profile data is sourced from respondents with an average age of nine years. Every effort was made to verify respondents' understanding of profile questions during piloting and utilize appropriate terminology and examples, but the respondents' age and likely inexperience answering such questions should be taken into consideration.
- The school-level teacher and student absence data for Term 2 and Term 3 is based on school administrative data collected by our enumerators and not direct observation.
- The Teacher Motivation scales were not as reliable (based on Cronbach Alpha scores) as we would have liked, especially at the sub-scale level. Additionally, some of the items within sub-scales have opposing values when they should be pointing in the same direction. This could be driven by lack of comprehension, unfamiliar concepts, insufficient contextualization or simply measures that do not hang together to form a cohesive construct.
- As with most surveys, there is the risk of social desirability bias in responses, particularly with the teacher survey. Efforts were made to encourage candid responses and respondents were reminded their responses would be confidential. We did receive responses counter to social desirability (e.g., 50% of teachers stated they would consider other job opportunity).
- The reading skills assessment, while modeled after EGRA, should not be compared directly with other EGRAs conducted in Rwanda because of the multi-leveled items included in syllable sounds and familiar word reading sub-tasks and the use of P1-leveled passage and comprehension questions.[32]

[32] As noted elsewhere, based on previous EGRA results, a P1 leveled passage and set of questions was utilized for this study to increase the non-zero scores and reading assessment data available for analysis. The reading assessment for this study was developed to respond to targeted research questions and not to serve as a measure of the proportion of P2 students meeting Kinyarwanda standards.

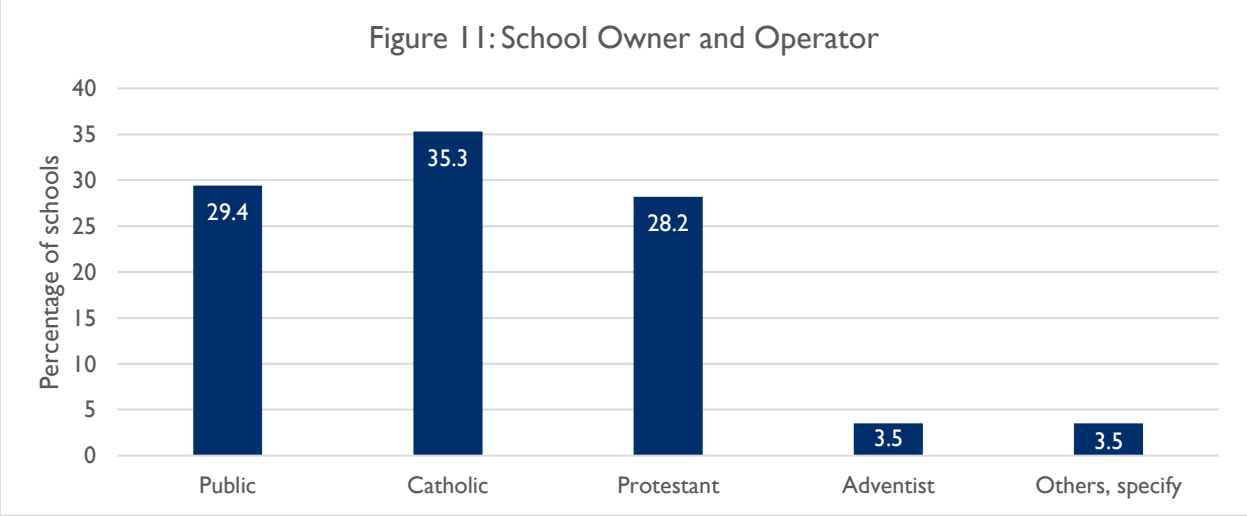
SECTION VI: FINDINGS

FINDINGS: SCHOOL PROFILE

In this section we provide an overview of the schools included in our random sample. Data for this section were collected from the administrative records in the school office and included all P2 Kinyarwanda class sections in the school, with data recorded separately for the observed P2 Kinyarwanda class.

In our random sample of 85 schools, 71% are government-aided and 29% are public. Within government-aided, there are a range of owners/operators as outlined in Figure 11 below:

FIGURE 11: SCHOOL OWNER AND OPERATOR



Through the school profile instrument, we collected data on basic school infrastructure and resources. In terms of learning resources, we see that a significant majority of rural schools (63%) do not have libraries/non-textbook reading materials available for P1-P3 student use. 50% of urban schools report having computers that are accessible for student use while only 1% of rural schools report this.[33] P1-P3 student access to the internet at school is nearly non-existent, with 96.5% of schools (both urban and rural) reporting no student access to internet.

With respect to infrastructure, 33% of rural schools still do not have electricity but nearly 100% of rural schools had working toilets that were accessible to students (100% of urban schools did). The following table provides the averages across the random sample and the urban/rural percentages for each variable.

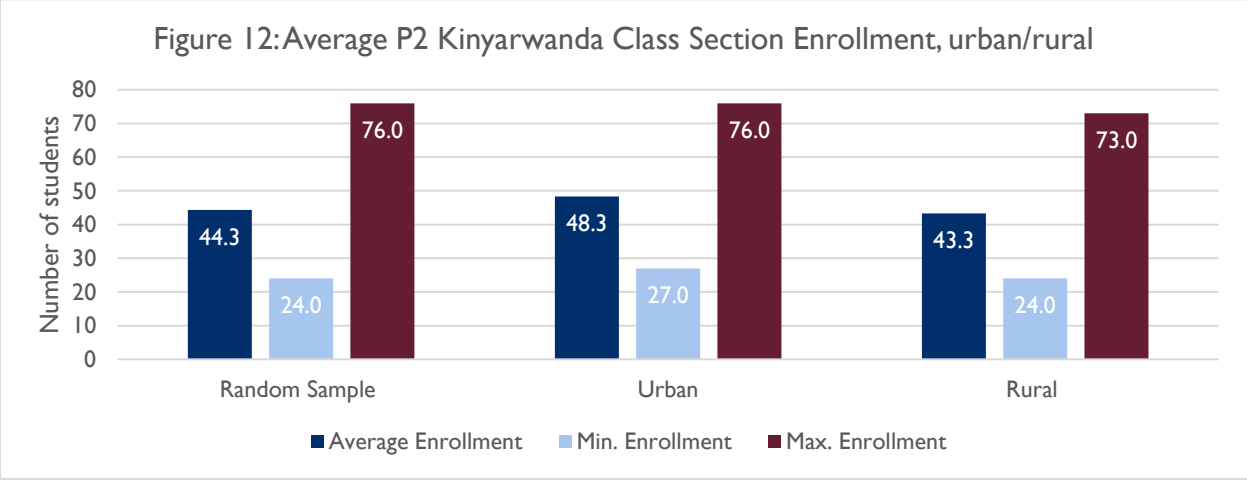
[33] The scope of data collection did not include verification of school-reported access to computers or estimation of frequency of use.

TABLE 7: SCHOOL INFRASTRUCTURE AND RESOURCES FOR RANDOM SAMPLE, URBAN/RURAL

Table 7: School Infrastructure and Resources for Random Sample, urban/rural	Yes (%)
Books (non-textbooks) available to P1-P3 students	37
Urban/Rural	*56/31*
Computers that are accessible to P1-P3 students' use	26
Urban/Rural	*50/19*
Internet accessible to P1-P3 students	4
Urban/Rural	*6/3*
Electricity	74
Urban/Rural	*100/67*
Working/useable toilets accessible to students	98
Urban/Rural	*100/97*
Drinking water available to students	74
Urban/Rural	*78/73*

There is a wide range of class section size across the random sample, with the lowest enrollment at 24 students and the highest enrollment at 76 students. The average, minimum and maximum enrollment does not vary much between urban and rural schools.[34]

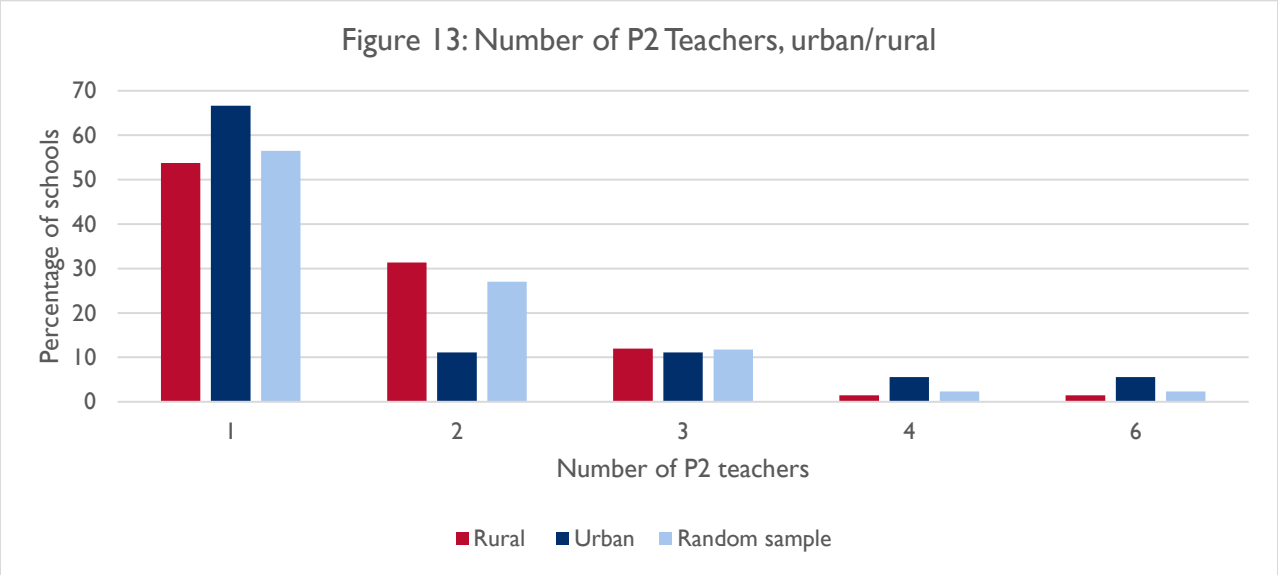
FIGURE 12: P2 KINYARWANDA CLASS SECTION ENROLLMENT, URBAN/RURAL



[34] Enrollment data based on school administrative records collected with the school profile instrument.

In terms of number of P2 teachers (all subjects), the random sample of schools is comprised of 57% single-teacher schools (54% rural and 67% urban), with the remaining schools having between 2 and 6 teachers teaching at the P2 level.

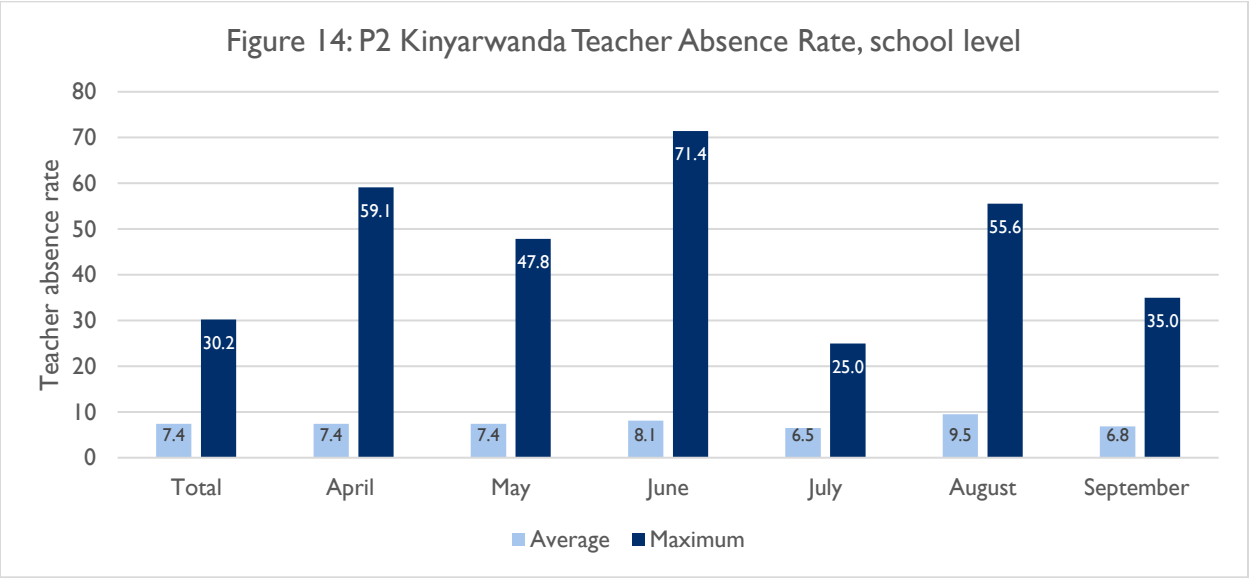
FIGURE 13: NUMBER OF P2 TEACHERS (ALL SUBJECTS), URBAN/RURAL



Using the school profile instrument, we also collected administration data on absences for all P2 (not just our observed teachers) Kinyarwanda teachers in each of the sample schools. We collected these data by month for Term 2 and for August and September of Term 3 given data collection took place during the month of October.[35] The average teacher absence[36] rate of 7.4% across all P2 Kinyarwanda teachers at the school level roughly corresponds with the 8.2% teacher absent rate in the observed lessons. We see general consistency in the average absence rate across months, with the highest rate in August at 9.5%. The maximum absence rate varies greatly across months however, with 71.4% of the days in June and 25% of the days in July.

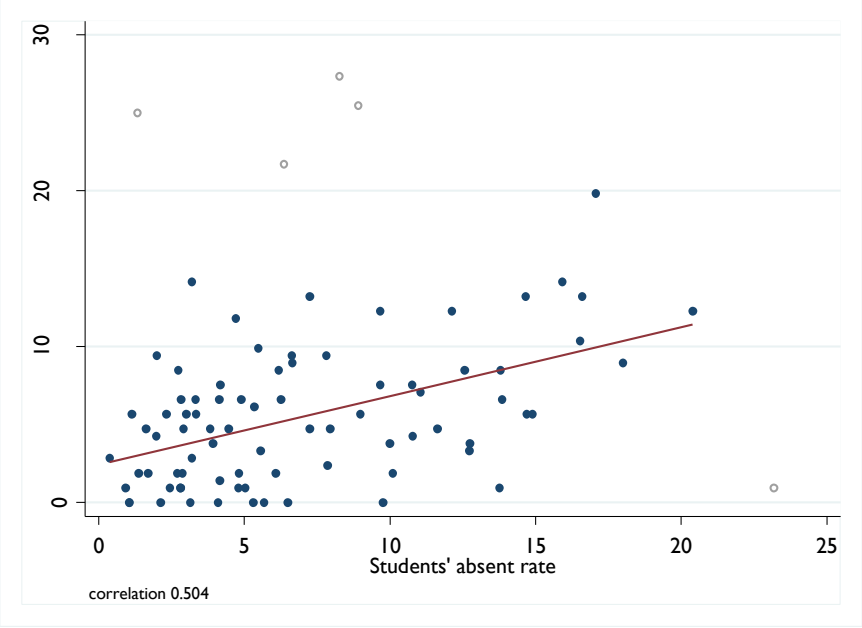
[35] The total number of school days in that period sums 106, distributed across months: April (11), May (23), June (21), July (22), August (9), September (20). Three days of August in Term 2 are not included.
[36] Per capita P2 Kinyarwanda teacher absence rate at the school level was calculated as the sum of all absences of all P2 Kinyarwanda teachers, divided by the number of P2 teachers, divided by the total number of school days in that period.

FIGURE 14: P2 KINYARWANDA TEACHER ABSENCE RATE, SCHOOL LEVEL



In addition, we collected administrative data on student absences for Term 2 and August and September of Term 3 for the class observed. When students' absence rate for each class is plotted against the absence rate calculated for the P2 teacher observed (see Figure 15 below), we find a positive and significant correlation of 0.504, which suggests classes with high teacher absenteeism tend to have high student absenteeism.

FIGURE 15: CORRELATION BETWEEN TEACHER AND STUDENT ABSENTEEISM[37]



[37] Schools with hollow circle with higher rates (potential outliers) are excluded from the estimation of the correlation.

This relationship indicates systemic issues at schools with high absenteeism and may be associated with school leadership or perhaps challenging contextual factors.

Next, we look at the time findings for lessons that we observed.

FIGURE 16: TOTAL TIME COMPOSITION/LOST

OBSERVED TIME (100%)

(TEACHER ABSENCE – 8.6%)

(STUDENT ABSENCE – 14%)

(TEACHER LATE – 2.1%)

(STUDENT LATE – 3.8%)

(loss of 28.5%)

IN-CLASSROOM TIME (71.5%)

(STUDENTS ENGAGED IN NON-INSTRUCTIONAL ACTIVITIES – 11.0%)

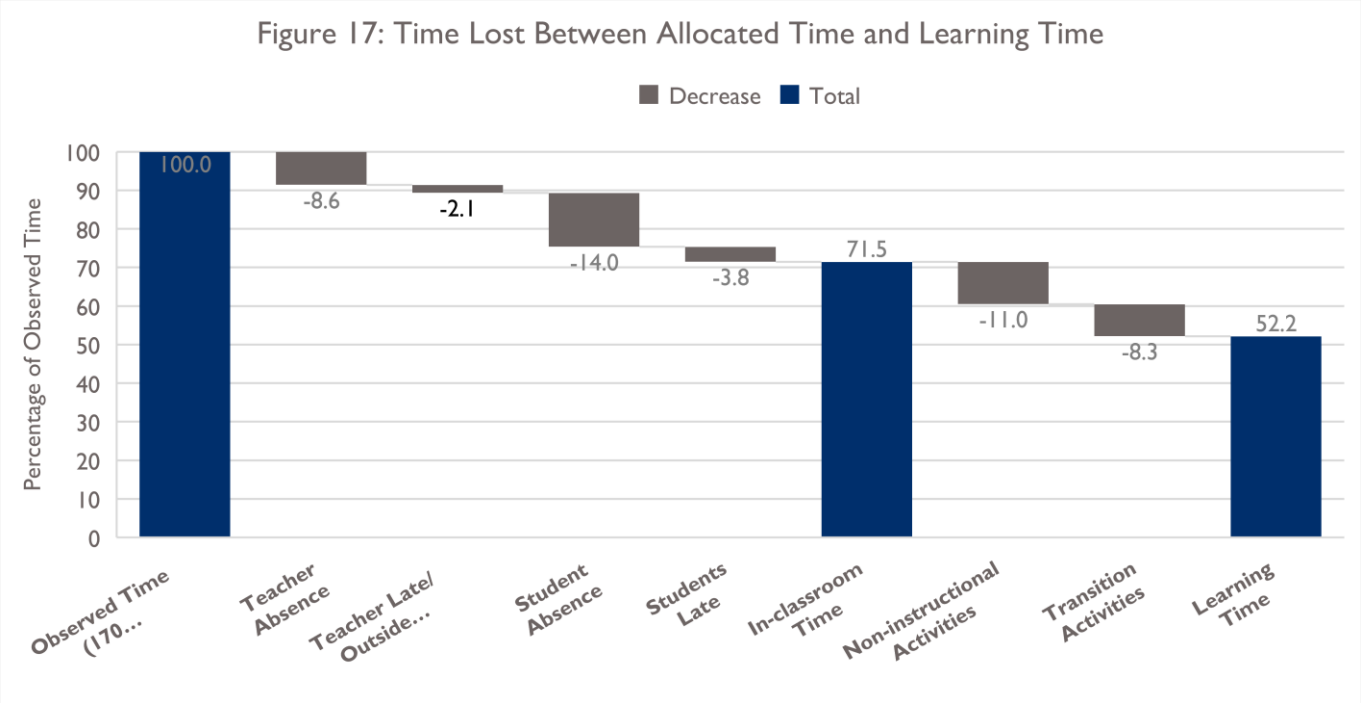
(STUDENTS ENGAGED IN TRANSITION ACTIVITIES – 8.3%)

(loss of 19.3%)

LEARNING TIME (52.2%)

Figure 16 above and 17 below illustrate instructional time loss due to all factors including teacher/student absence, teacher/student tardiness and non-instructional activities when the teacher and students are in the classroom together.[38]

FIGURE 17: TIME LOST BETWEEN ALLOCATED TIME AND LEARNING TIME[39]



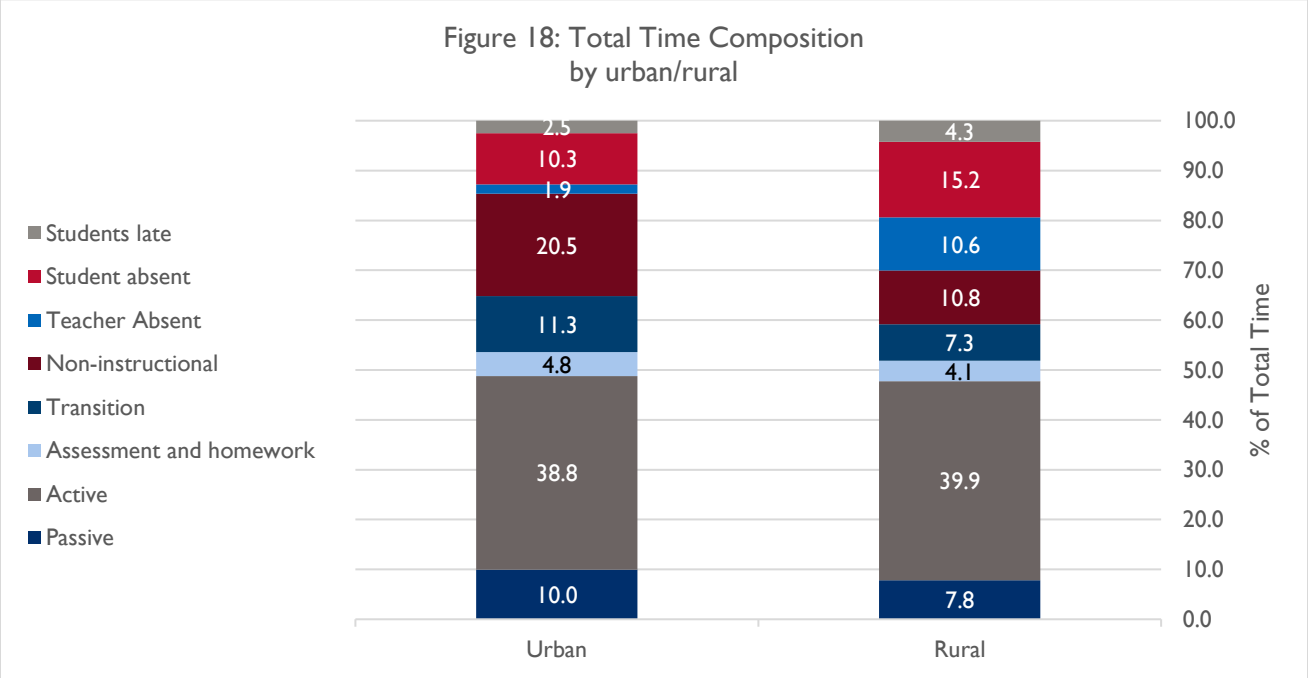
Teacher/student absence and tardiness represent 28.5% loss of instructional time,[40] and results in teachers and students being together in the classroom only 71.5% of time. Beyond this loss due to absences and tardiness, 19.3% of in-classroom time was utilized for activities other than direct reading instruction (non-instructional and transition activities), resulting in a total of 52.2% of the intended/allocated Kinyarwanda reading instruction time being dedicated to learning time. Figure F.1 in Annex F shows the distribution of % of in-classroom and learning time and frequency all across classes observed. In the subsequent sections, we will discuss in more detail time lost due to teacher and student absences and the composition of in-classroom time.

[38] The ratios presented below are based on percentage of time observed. Note that the time observed was actually 108% of REB allocated time due to lesson time overruns.
[39] The analysis is restricted to the schools in our random sample — 85 schools observed twice (which totals 170 class observations).
[40] See Annex D for a detailed description of how time is calculated.

LOST TIME DUE TO TEACHER AND STUDENT ABSENCE AND TARDINESS

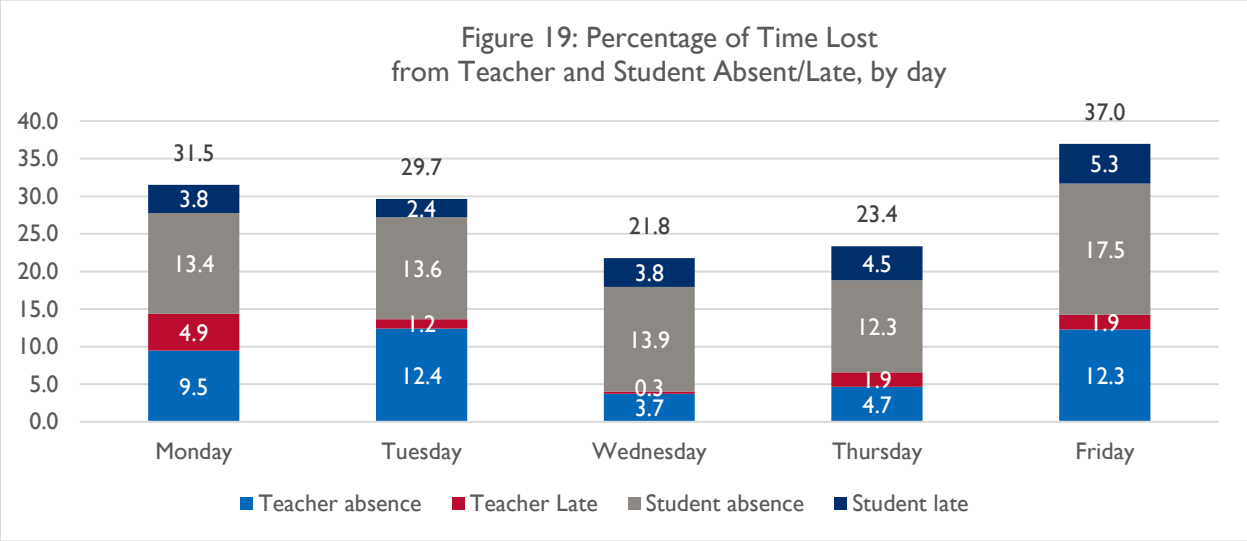
Teacher and student absence and tardiness constitutes a loss of nearly 30% of time and as such deserves further exploration in terms of variation (e.g., day of week, class period, urban/rural) and underlying reasons. First, we look at time lost due to teacher and student absence and tardiness disaggregated by urban and rural schools. The most pronounced variations are in teacher/student absence and the amount of non-instructional time (this can be attributed to teacher not in classroom, students off-task, etc.). In urban schools, there was much less time lost due to teacher and student absence. Teacher absence in urban schools was only 1.9% on average as compared to 10.6% in rural schools. Student absence averaged 10.3% in urban schools and 15.2% in rural schools.

FIGURE 18: TOTAL TIME COMPOSITION, URBAN/RURAL



We also analyzed teacher/study absence and tardiness by day of the week. Student absence and tardiness rates are similar across days of the week, with some increase in Fridays. However, the percentage of time lost due to teacher absence on Mondays, Tuesdays and Fridays is striking, with a high of 12.4% on Tuesdays and a low of 3.7% on Wednesdays.

FIGURE 19: PERCENTAGE OF TIME LOST FROM TEACHER AND STUDENT ABSENT/LATE, BY DAY

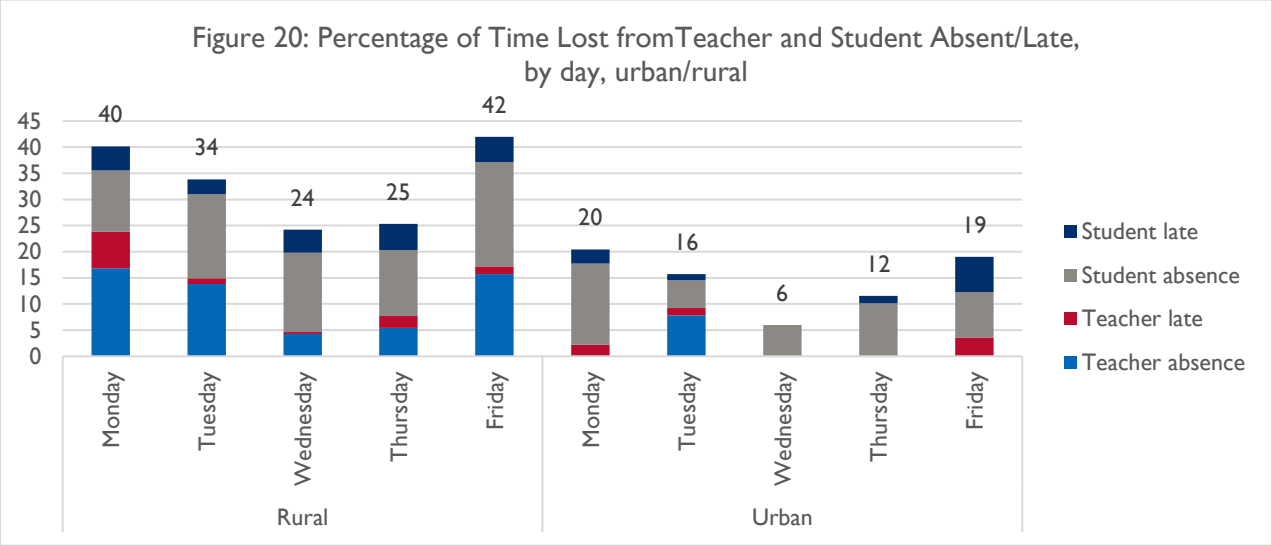


Note: The table above shows the proportion of lost time relative to the number of minutes and hours observed. The number of classes observed by day of week can be found in Annex C.

The data show that teacher absence is predominantly a rural phenomenon as illustrated in Figure 20. This may be driven by factors such as market days[41] or long travel distance on the weekends to visit family or home community. We did analyze whether teachers in rural areas more frequently moved from their home community for their teaching position, and that was not the case. We also analyzed differences in commuting time reported by urban and rural teachers and did not find a difference, on average, between urban and rural areas.

[41] Further exploration of teacher absence by school/sector and events in the local community such as market days could help explain these trends in teacher absence.

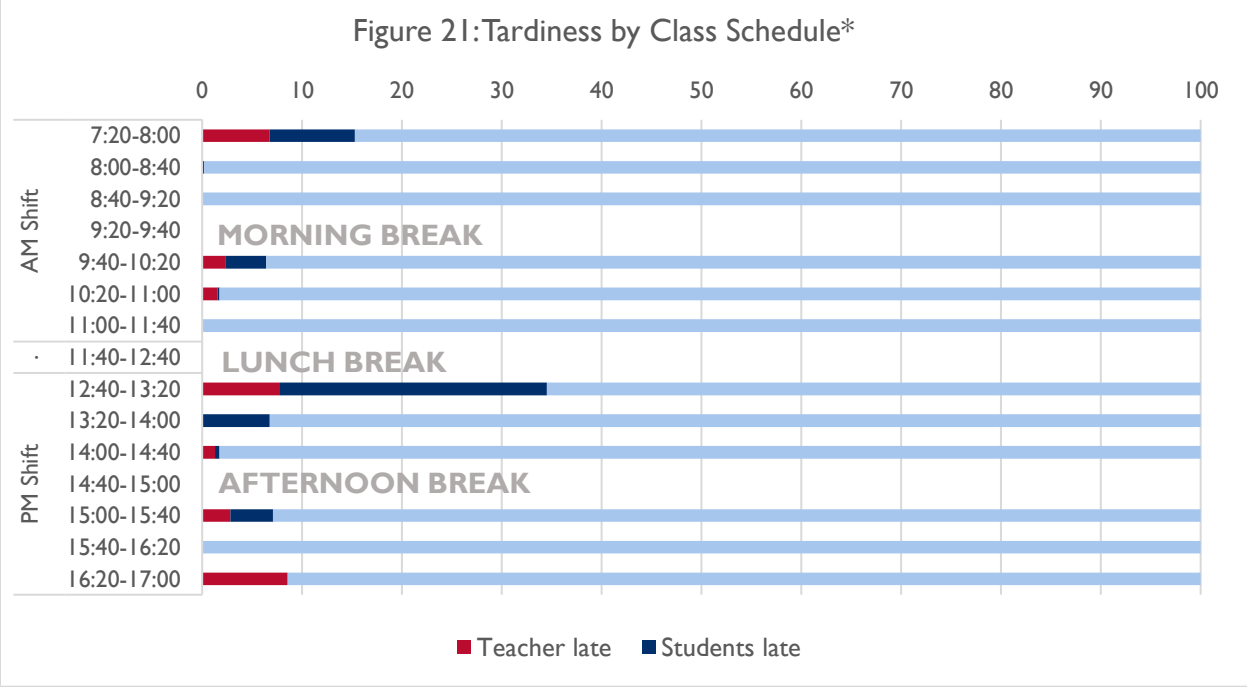
FIGURE 20: PERCENTAGE OF TIME LOST FROM TEACHER AND STUDENT ABSENT/LATE, BY DAY, URBAN/RURAL



The overall pattern of time lost by day due to teacher/student absence and tardiness is similar across urban and rural schools with more time lost on Mondays, Tuesdays and Fridays, however the overall proportion of time lost is much higher in rural areas with a maximum of 42% on Fridays as compared to a maximum of 20% on Mondays in urban schools.

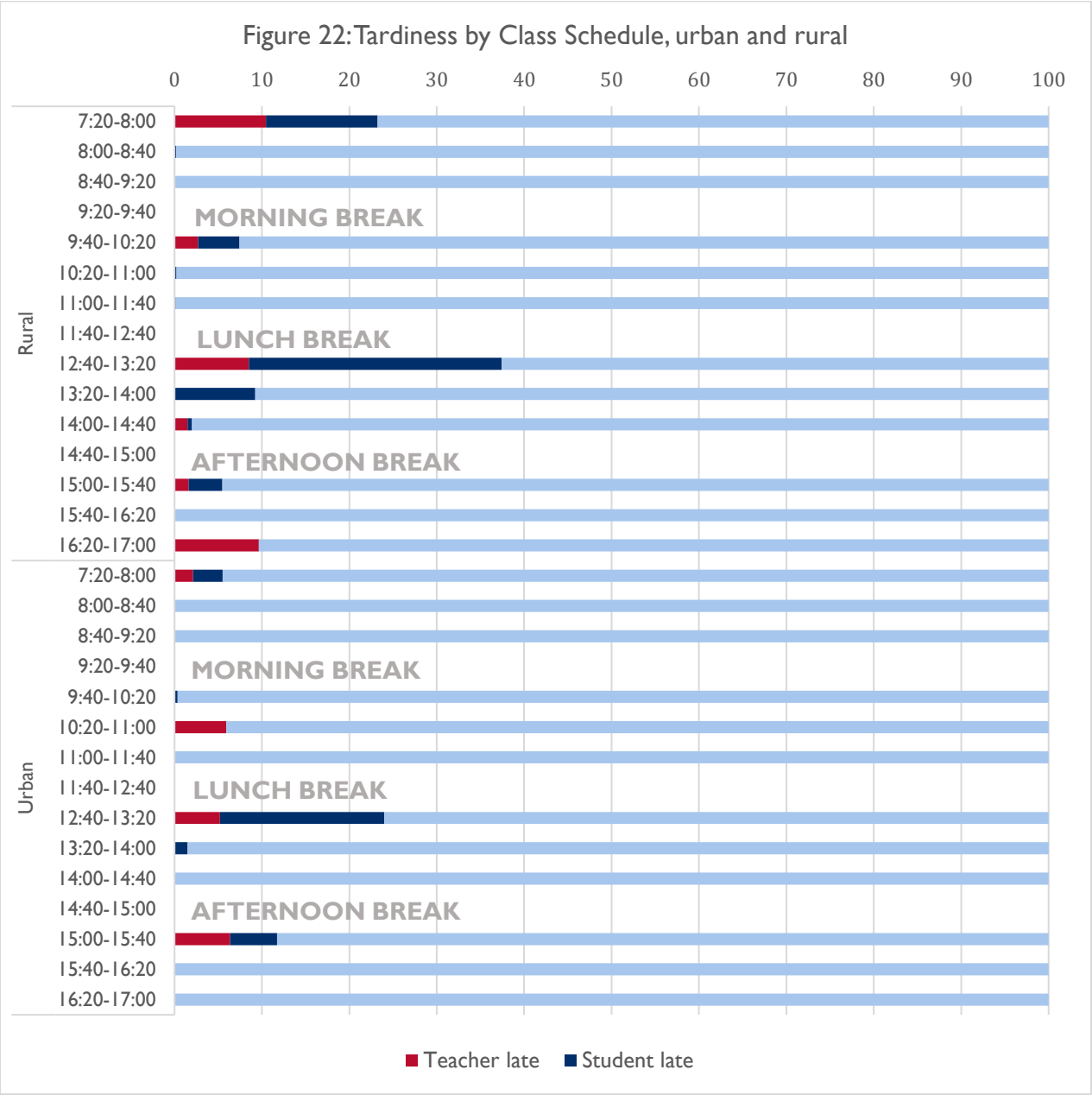
In addition to day of the week, we disaggregated the tardiness data by class schedule. Tardiness is calculated as proportionate to the total number of hours observed.

FIGURE 21: TARDINESS BY CLASS SCHEDULE*



**Classes where teacher was not present are not included. Four classes with irregular schedules were not included. See Annex C for table of classes observed by hour and day.*

FIGURE 22: TARDINESS BY CLASS SCHEDULE, URBAN AND RURAL

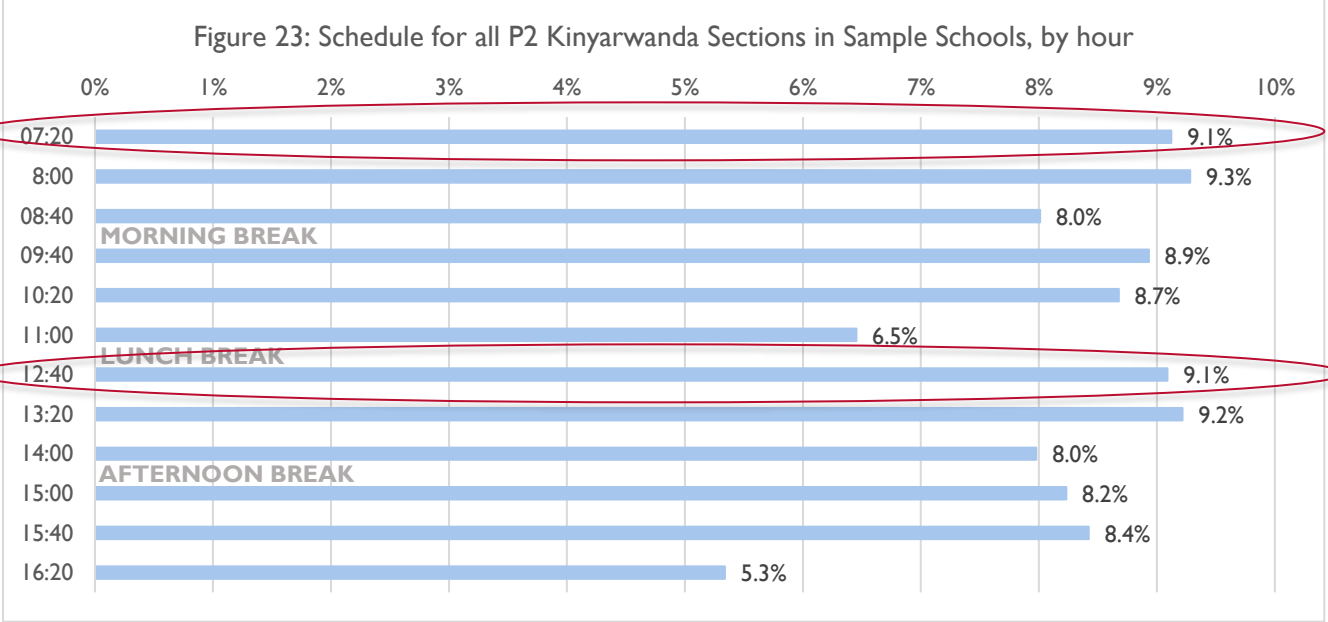


Note: Classes where teacher was not present are not included. Four classes with irregular schedules are not included. See Annex C for table of classes observed by hour and day.

To further understand the impact of the absence and tardiness rates discussed above, we compiled the schedules (day of week and time of day) for all P2 Kinyarwanda class sections across the 100 schools in our sample (not just for our observed classes). There are a total of 3,146 scheduled sessions of P2 Kinyarwanda classes across the 100 schools, with an average of just under four P2 Kinyarwanda sections per school that meet eight times per week.

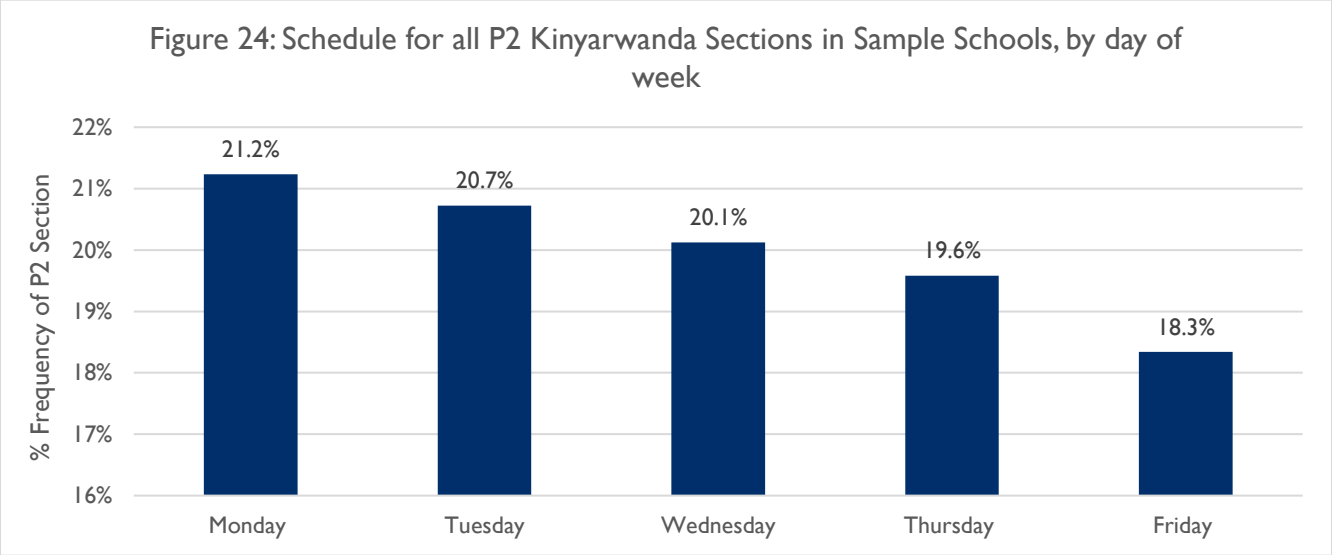
When we compare the tardiness data by hour with the distribution of P2 Kinyarwanda classes by hour in Figure 23 (below), we see that the two times with the highest combined teacher and student tardiness rates, 7:20 and 12:40, are high frequency times for P2 Kinyarwanda classes. This scheduling compounds the amount of lost time.

FIGURE 23: SCHEDULE FOR ALL P2 KINYARWANDA SECTIONS IN SAMPLE SCHOOLS, BY HOUR



Additionally, when we look at the schedule data across days of the week and hours of the day, we see the frequency of P2 Kinyarwanda classes are highest in the early part of the week and then decrease by day, albeit the total spread is minimal with a high of 21.2% of P2 Kinyarwanda classes occurring on Mondays and 18.3% of classes taking place on Fridays (see Figure 24, next page). If we reflect on the rates of teacher and student absence and tardiness by day of the week (see Figure 19), 21.2% of classes are losing 31.5% of time just from teacher and student absence and tardiness. For rural schools, the loss is even more significant, with teacher and student absence and tardiness totaling 40% on Mondays (see Figure 20) when the highest percentage of P2 Kinyarwanda classes are scheduled.

FIGURE 24: SCHEDULE FOR ALL P2 KINYARWANDA SECTIONS IN SAMPLE SCHOOLS, BY DAY OF WEEK



Please see the student and teacher findings sections for further discussion about the reported reasons for teacher and student absences and tardiness.

IN-CLASSROOM TIME

Having discussed time lost associated with teacher/student absence and tardiness (28.5%), we now turn to the findings about what happens during Kinyarwanda classes when the teacher and students are in the classroom together (71.5% of the allocated time). Throughout, we refer to this as "in-classroom time." It is important to recall from Section IV that *the total time calculation for in-classroom is based on student* activities and engagement level in those activities.[42]

Recalling from the study design discussion, there are five primary categories of in-classroom time (with illustrative examples below):

- Active learning time
  - Teacher asks students to make predictions about a text/story/poem from illustrations
  - Teacher asks student(s) to model blending letters to form a word
  - Student(s) identify letters and sounds, verbally, contained in syllables or words the teacher is saying
- Passive learning time
  - Students are listening to the teacher blend letters or break a word into letters
  - Students choral repeat letter names or sounds out loud from the board, charts, cards, etc.
  - Teacher models the writing of letters while students watch

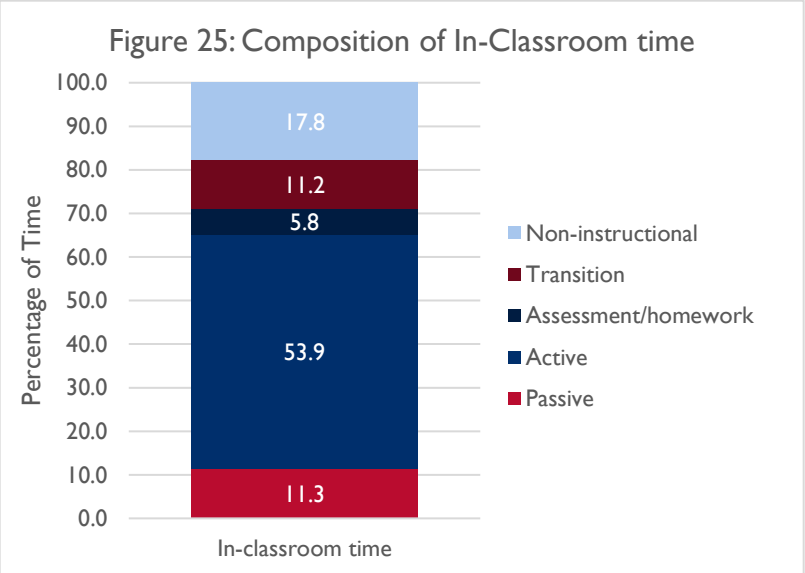
[42] A minor adjustment was performed to separate out the time a teacher arrives late or leaves the classroom for a certain period during the observation and attribute that amount of time lost to teacher tardiness/not in the classroom, rather than students' engagement in non-instructional activities (since there is no instruction happening) during teacher late arrival/missed period of class.

- Assessment/Homework activities
  - Teacher is assessing individual students or checking classroom work (e.g., walking around, asking children to come to her desk)
  - Teacher is assigning homework (verbally, writing on board, handing out papers)
- Transition time
  - Teacher is restating/asking questions about a previous lesson
  - Teacher is setting up classroom/handing out materials/preparing exercise
- Non-instructional time
  - Teacher is in the classroom, but distracted and not engaged
  - Teacher is doing administrative paperwork in classroom
  - Teacher is disciplining students

We first look at the composition of in-classroom time for the entire random sample in Figure 25 to the right. We see that the majority of in-classroom time is spent on direct reading instruction (active and passive), with 65.2% of in-classroom time.

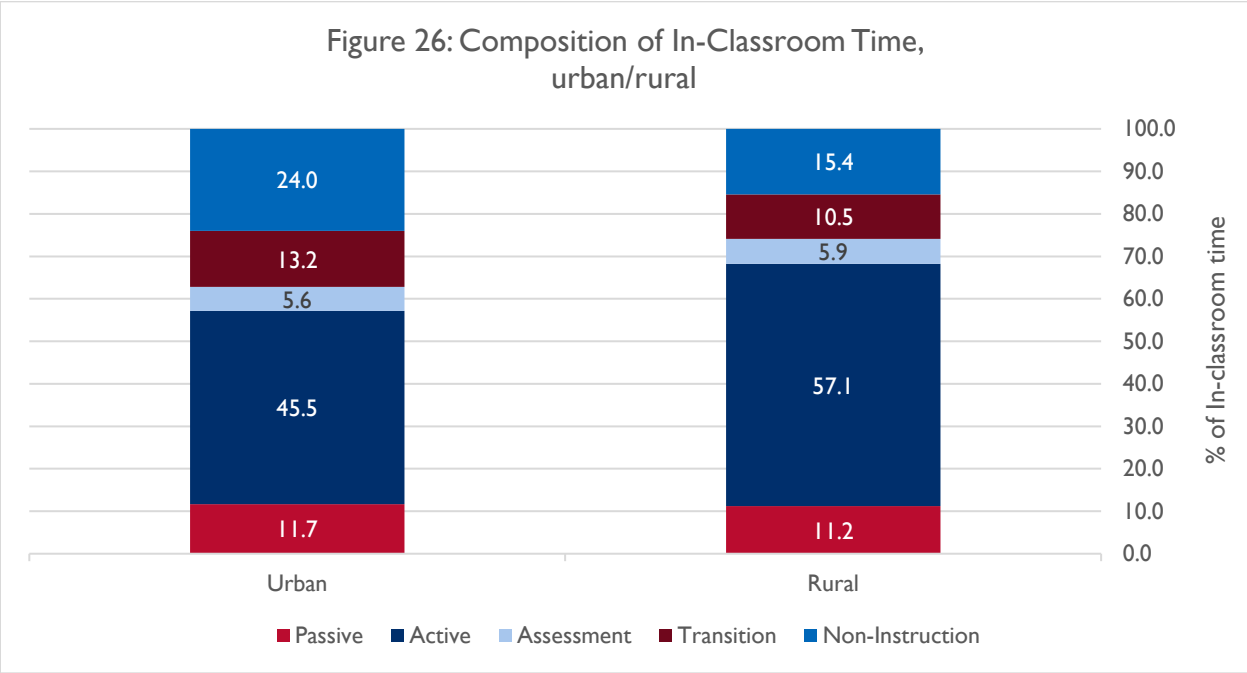
When we disaggregate the composition of in-classroom time, we do see variances between urban and rural schools with learning time constituting 57.2% of in-classroom time for urban schools on average and 68.3% for rural schools on average. Rural schools spend more time on active instruction (57.1%) than urban schools do (45.5%). There is almost no variation between urban and rural schools for time spent on assessment/homework. Non-instructional time takes up a far greater percentage of time in urban schools.

FIGURE 25: COMPOSITION OF IN-CLASSROOM TIME



Note: In-classroom time represents 71.5% of the allocated/observed time (see Figure 17).

FIGURE 26: COMPOSITION OF IN-CLASSROOM TIME, URBAN/RURAL



These next sections disaggregate how time is spent across urban and rural schools, within the following categories of in-classroom time: learning (active and passive), assessment/homework time, transition time and non-instructional time. It is important to remember that the denominator for all of these analyses is shown in Figure 25 (previous page). Annex F shows the distribution of active and passive as a percent of in-classroom time, by urban and rural.

COMPOSITION OF LEARNING TIME (BY READING SKILL)

As shown in Figure 17, when we move from in-classroom time (71.5% of allocated time) to learning time (52.2% of allocated time), the further loss of instructional time is significant and driven by the amount of transition and non-instruction time (totaling 19.3%) taking place in classrooms.

If we model these percentages across the 320 minutes allocated each week for P2 Kinyarwanda lessons, we arrive at 228.80 minutes of in-classroom time and only a total of 149.18 minutes of learning time.

In the next stage we capture an even more granular level: the type of reading skill being focused on in the classroom (e.g., syllable sounding, familiar word reading), again based on student activities and their level of engagement in those activities.

We find that phonics syllables/words is by far the dominant reading skill focused on during P2 Kinyarwanda lessons with nearly 31% of learning time being dedicated to this skill. Learning to read words and learning to read sentences the next most frequent. The difference between these latter two skills in practice can be quite minimal. This brings the total time focusing on learning to read words/sentences to 22.3% of in-classroom time. This allocation of time across reading skills could in part be a function of the distribution of observations by day given that the Kinyarwanda curriculum progresses through reading skills.

FIGURE 27: PERCENTAGE OF LEARNING TIME PER READING SKILL (EXCLUDING NON-INSTRUCTION AND TRANSITION TIME)[43]

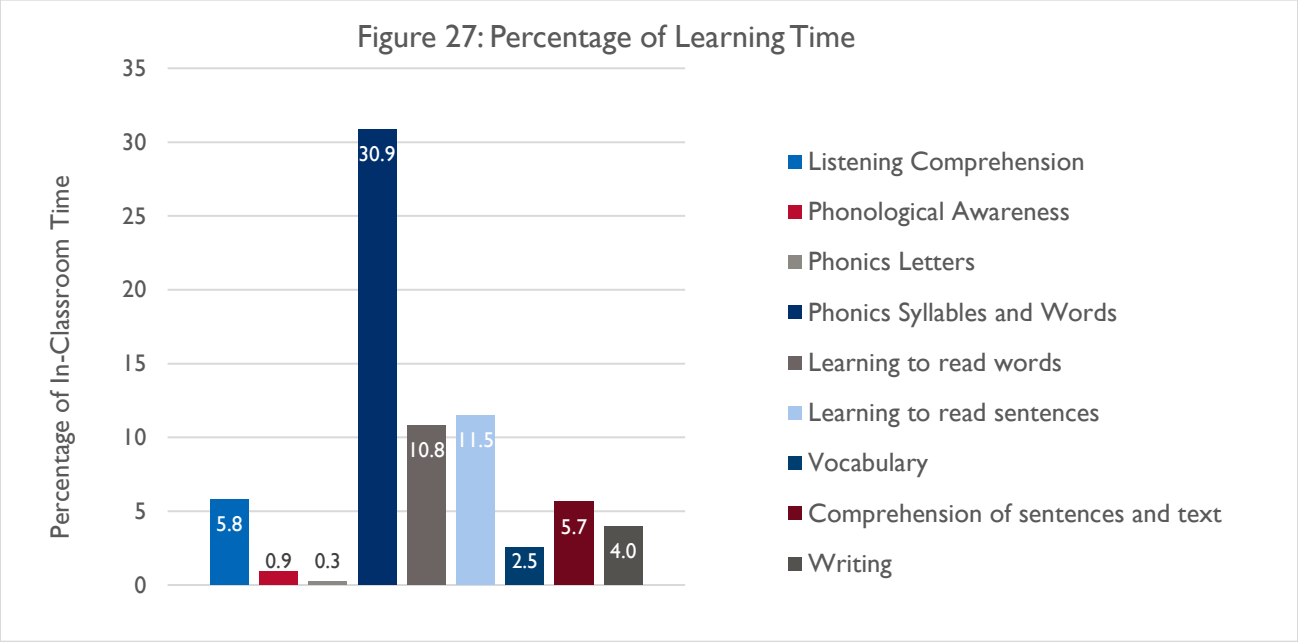
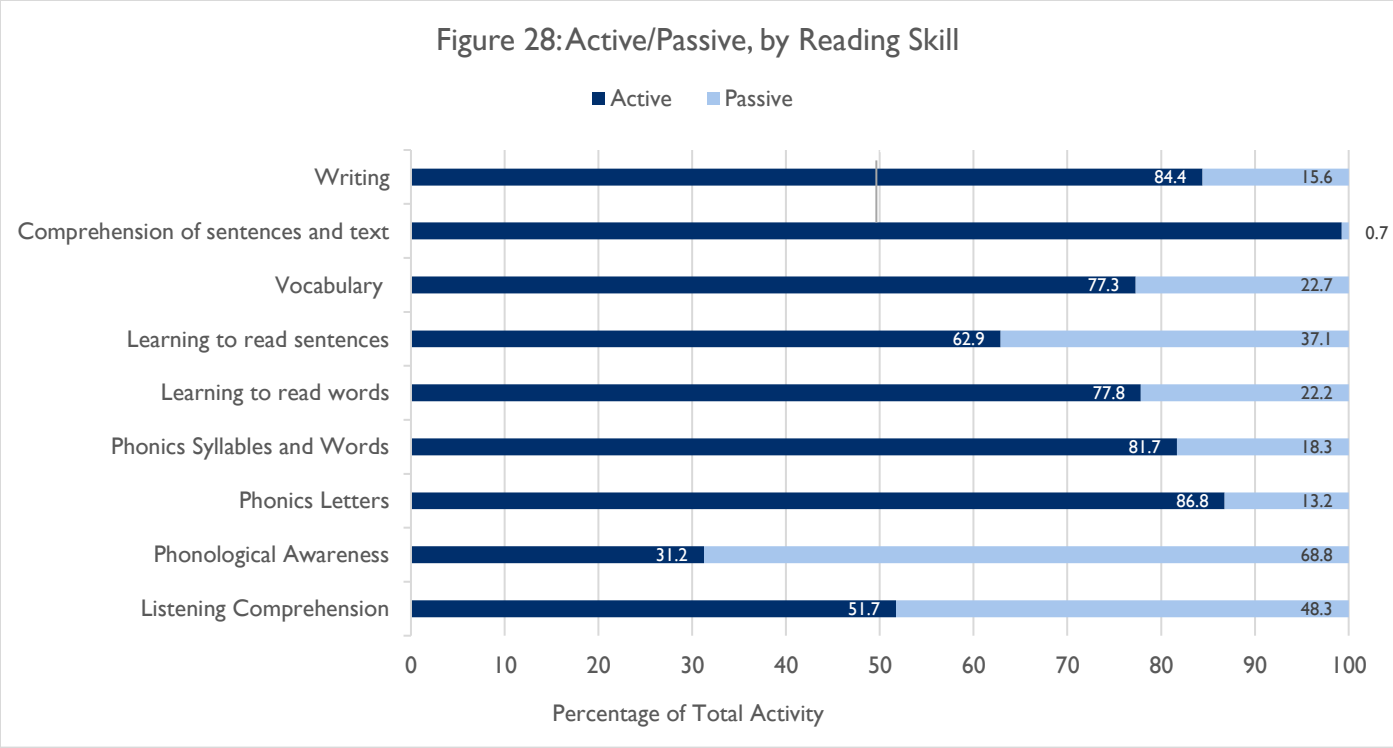


FIGURE 28: ACTIVE/PASSIVE, BY READING SKILL

As noted in the study design section, each reading-related activity was coded as active or passive. Passive examples include: "students choral repeat the syllable of the day," "students listen to teacher blend syllables," or "students choral read words fluently (from board, etc.)." Active examples include: "students are writing original syllables or words on the board," "students guess the meaning of a word," or "students are summarizing a text or story the teacher has read."[44] This distinction between active and passive reading activities is critical given the positive and significant relationship found in this study between time spent on active reading activities and student reading scores. See the latter part of this section for a full discussion.

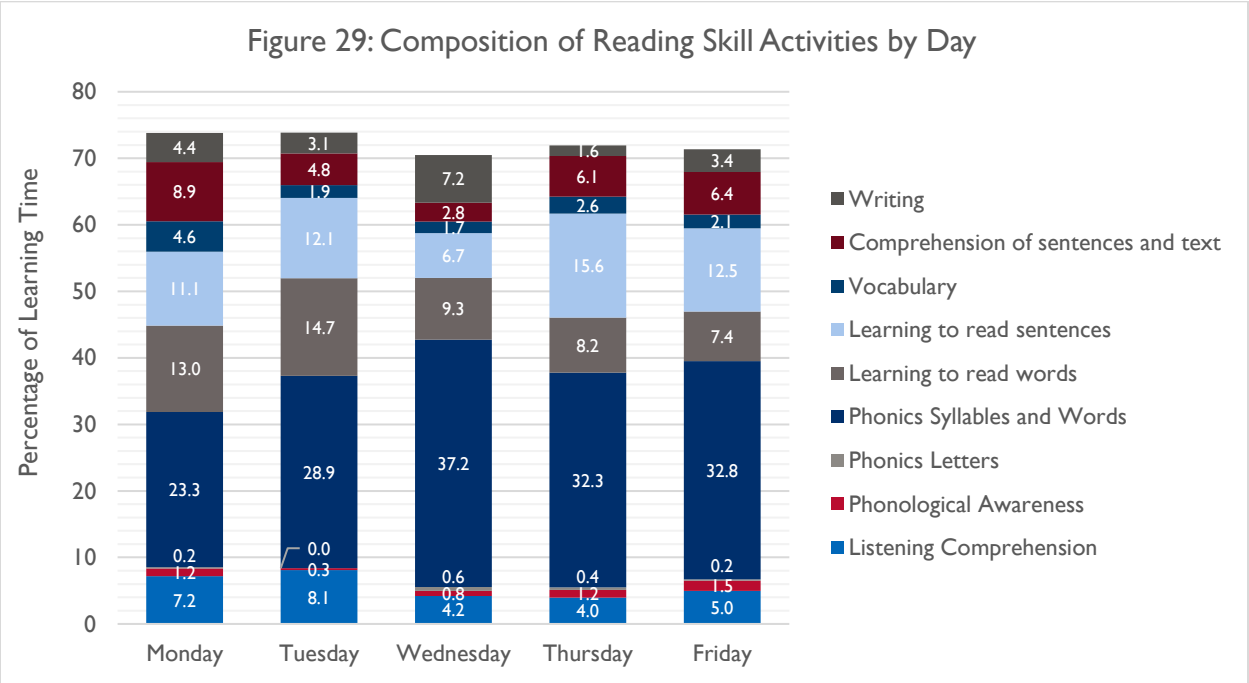
[43] Excludes transition and non-instructional activities/time.

[44] Note the designation between active and passive reading activities does not always infer one is more effective than the other as some instruction coded as passive is necessary for introducing lessons, modeling skills, implementing gradual release pedagogy strategies, etc. That said, the great majority of activities coded as passive utilize rote learning strategies and may have alternatives that engage students in more active and participatory learning.



For future reviews of reading instruction as it relates to the curriculum, schemes of work, and lesson planning, the following data on reading skill focus by day of week might be useful. Based on the current curriculum, it is expected that listening comprehension would be higher in the beginning of the week, but we do not see that in the observation findings. We can also consider phonics letters and phonics syllables/words as a single activity given the similarities and difficulty in categorization across these two skills. The same collapsing of reading words and reading sentences may also be appropriate given the similarity of these activities in the classroom.

FIGURE 29: COMPOSITION OF READING SKILL ACTIVITIES BY DAY[45]



COMPOSITION OF TRANSITION TIME

A significant proportion of time is spent on transition activities, with variations across urban and rural schools. We explored the composition of transition time, disaggregated by urban and rural schools, to further understand how this time was being spent. In both urban and rural schools, "preparing for the lesson" constitutes the largest amount of transition time with 36.6% and 32.2% respectively. The most striking difference is the amount of time (23%) spent "singing, stretching at the direction of the teacher" in rural schools. It is important to recognize that some level of transition time is necessary for instruction and provides useful connections between different activities and lesson content. Notwithstanding this, use of instructional time for these activities should be further explored with respect to efficacy.

[45]We might consider phonics letters and phonics syllables/words as a single activity given the similarities and difficulty in categorization across these two skills. The same collapsing of reading words and reading sentences may also be appropriate given the similarity of these activities in the classroom. See Section II and Annex H for further discussion of reading skills activities and categorization.

FIGURE 30: COMPOSITION OF TRANSITION TIME, URBAN

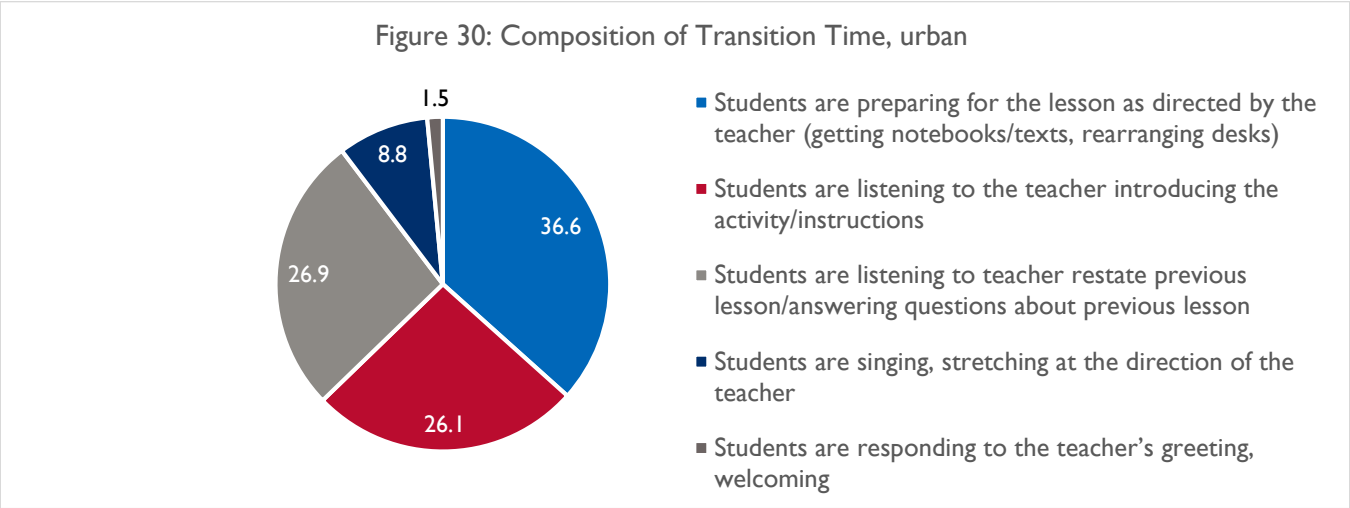
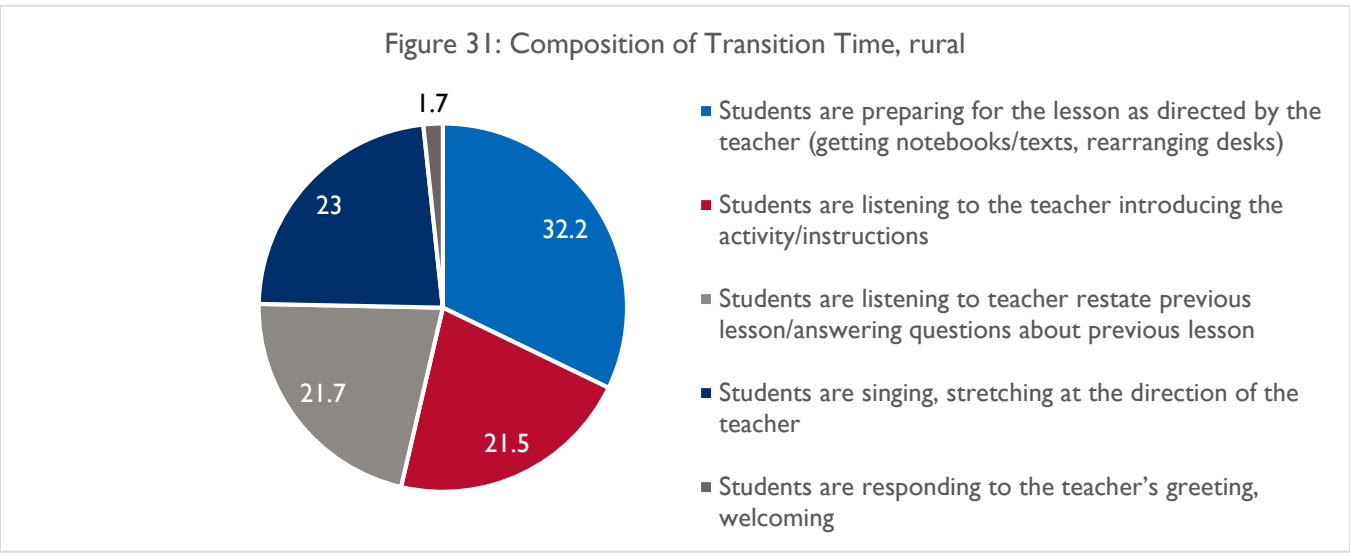


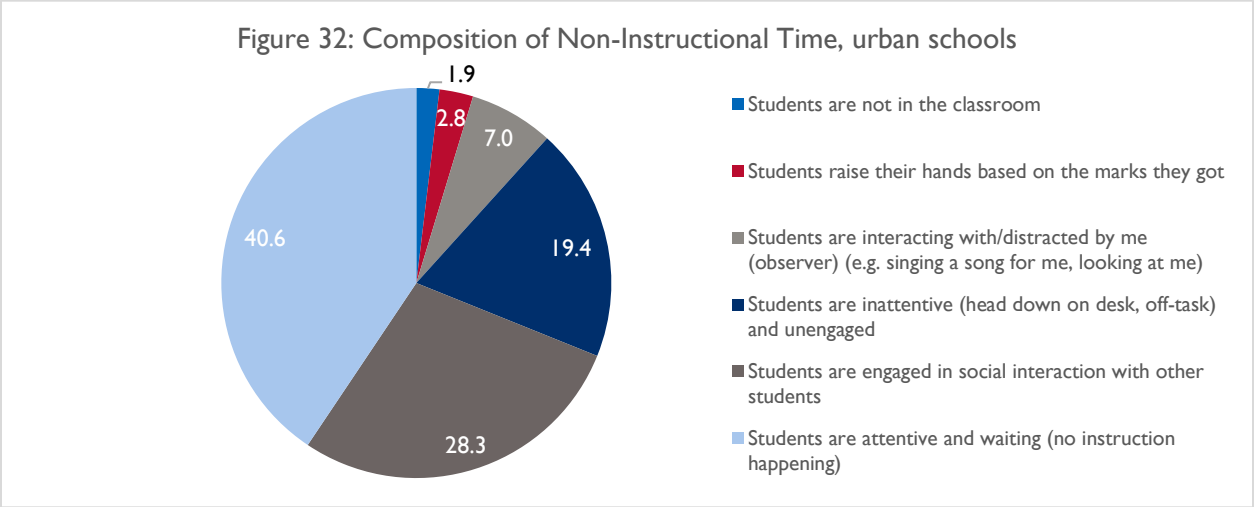
FIGURE 31: COMPOSITION OF TRANSITION TIME, RURAL



COMPOSITION OF NON-INSTRUCTIONAL TIME

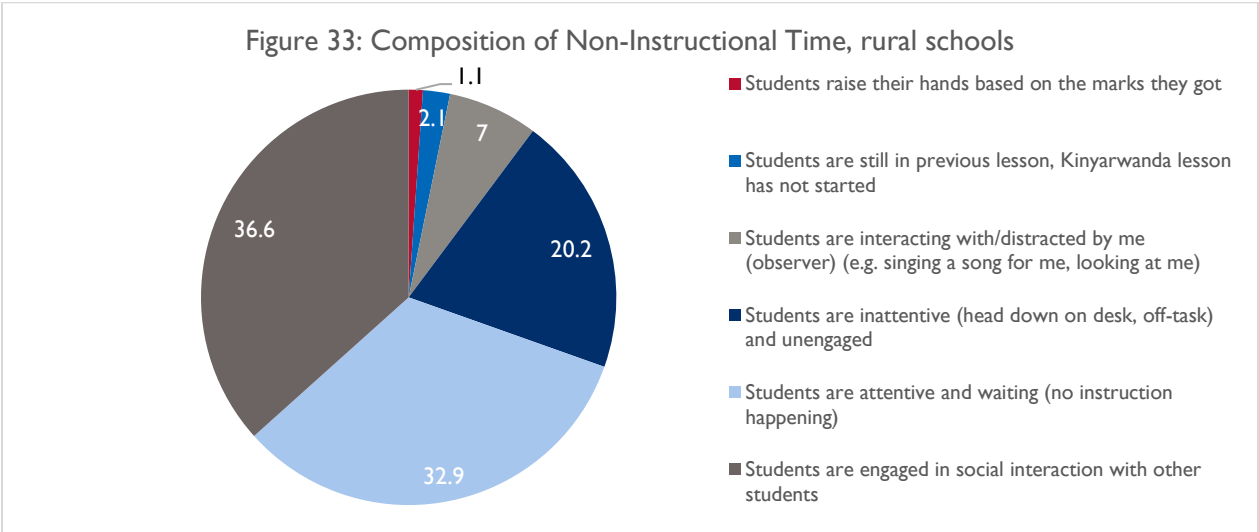
The largest allocation of non-instructional time is when "students are attentive and waiting for instruction," making up 40.6% of non-instructional time in urban schools and 36.6% of non-instructional time in rural schools. Note that non-instructional time, on average, is only 17.8% of the overall allocated time. This time represents a lost opportunity to fully leverage in-classroom time for learning and is largely in control of teachers. Of the 24% of total in-classroom time that is non-instructional time in urban schools, we see that over 50% of non-instructional time is attributed to students being off-task, either engaged socially with other students or inattentive. The ratios are similar in rural schools in terms of students being off-task. These data are further reinforced by data from observations that roughly 25% of teachers made efforts to engage students only some of the time or less often/none of the time, so we know that there is room to improve teacher efforts to decrease student off-task time.

FIGURE 32: COMPOSITION OF NON-INSTRUCTIONAL TIME, URBAN SCHOOLS[46]



In rural schools, 15.4% of in-classroom time is non-instructional, and there is a lower proportion of non-instructional time due to students waiting for instruction, but it is still 32.9% of non-instructional time. Nearly 57% of non-instructional time is attributed to students' non-engagement while instruction is happening, highlighting the need for more effective strategies to keep students engaged.

FIGURE 33: COMPOSITION OF NON-INSTRUCTIONAL TIME, RURAL SCHOOLS[47]



[46] The graph and key colors are based on rank order, which is different for rural schools (Figure 33). "Students raise their hands based on the marks they received" refers to when a teacher asks students with certain scores or score ranges on an assignment or test to raise their hands.

[47] The graph and key colors are based on rank order, which is different for urban schools (Figure 32). "Students are still in previous lesson" is coded when students are in the classroom but the lesson for another subject is still on-going. This was only observed in rural areas.

FINDINGS: READING SKILLS

As noted in the methodology section, reading skills were assessed using a range of sub-tasks with differing levels of difficulty to avoid floor and ceiling effects. The following section explores the reading scores of students included in the random sample.

MEAN READING SCORES

TABLE 8: READING ASSESSMENT SCORES

The following table summarizes the zero scores, average reading scores with zero scores and without zero scores for each sub-task.

Table 8: Reading Assessment Scores[48]	% Zero Scores	Average % Correct with Zero Scores	Average % Correct without Zero Scores
Letter Naming	2%	85%	86%
P1 Syllable Sounding	15%	67%	80%
P2 Syllable Sounding	30%	49%	59%
P1 Word Reading	32%	60%	88%
P2 Word Reading	33%	59%	87%
P3 Word Reading	35%	55%	82%
P1 Passage Reading (1-minute timed assessment)	35%	16.2 cwpm	25.1 cwpm
P1 Passage Reading (3-minute timed assessment)	N/A	15.6 cwpm	24.1 cwpm
P1 Comprehension Questions Correct	40%	2/5	3.1/5

In the next section, we look at the distribution of both oral reading fluency and comprehension scores, disaggregated by urban/rural and male/female. We see the high zero scores in rural schools. Taking the zero scores out of the picture, the distribution of scores between urban and rural schools is not that dissimilar. With respect to the distributions comparing male and female scores, we see higher zero scores for males and a general skew to the left, indicating a greater proportion of lower scores for male students.

FIGURE 34: ORAL READING FLUENCY DISTRIBUTION, URBAN/RURAL

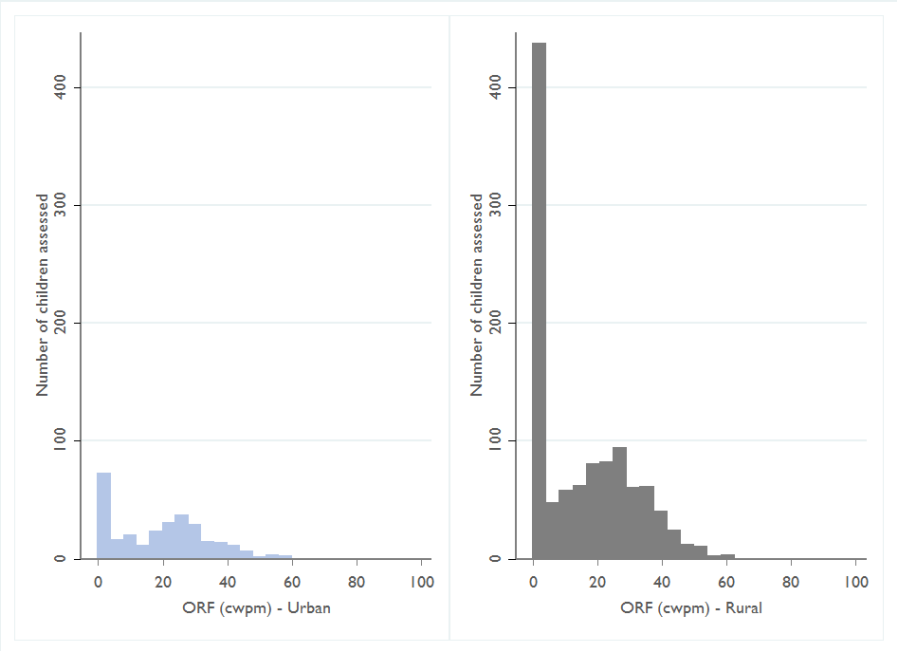


FIGURE 35: READING COMPREHENSION QUESTIONS CORRECT DISTRIBUTION, URBAN/RURAL

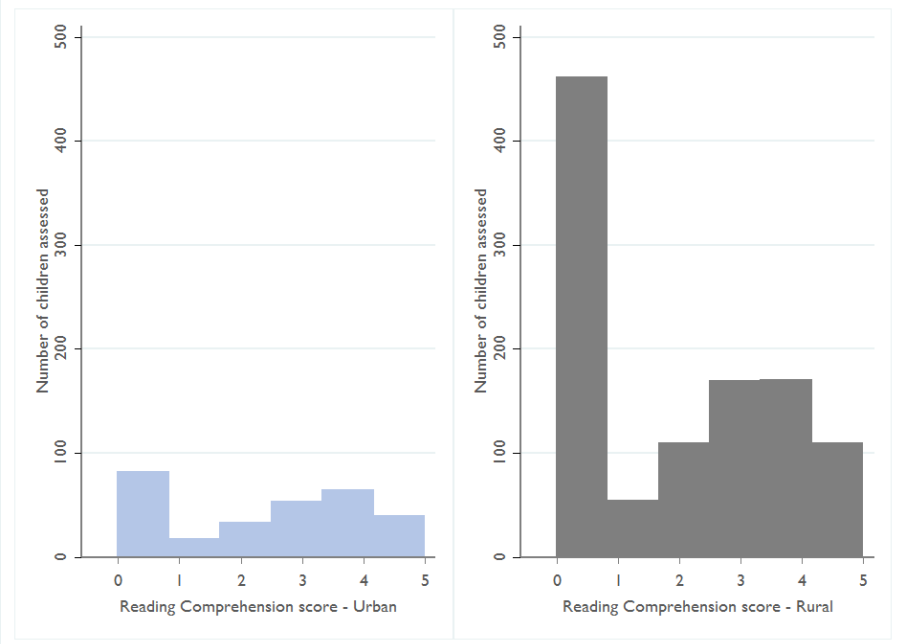


FIGURE 36: ORAL READING FLUENCY DISTRIBUTION, MALE/FEMALE

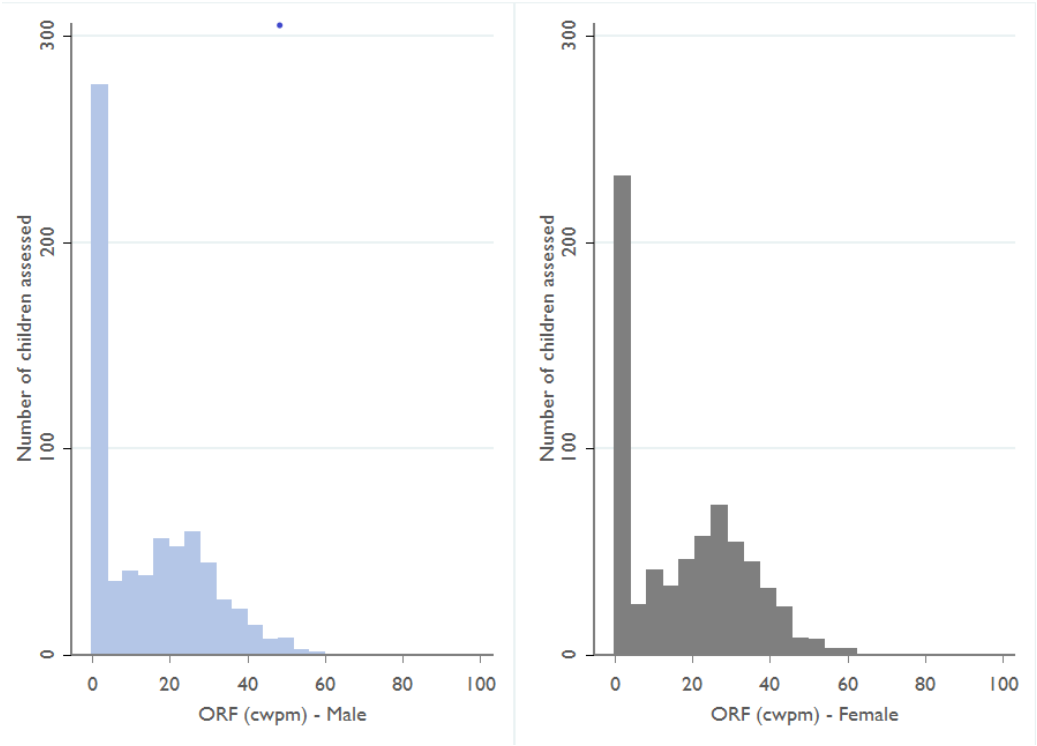
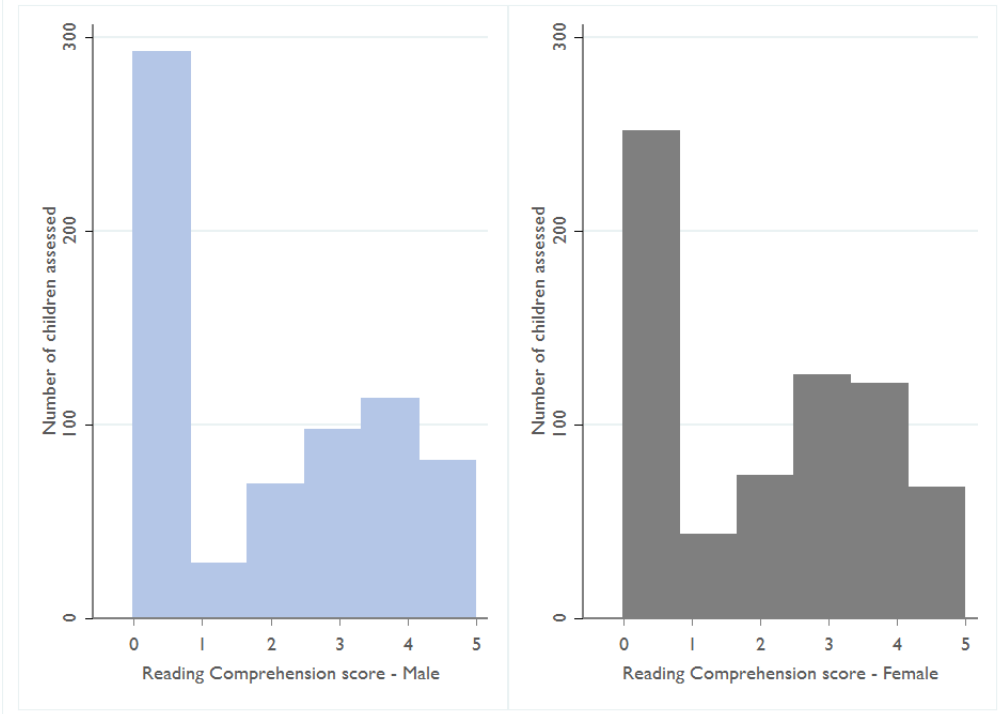


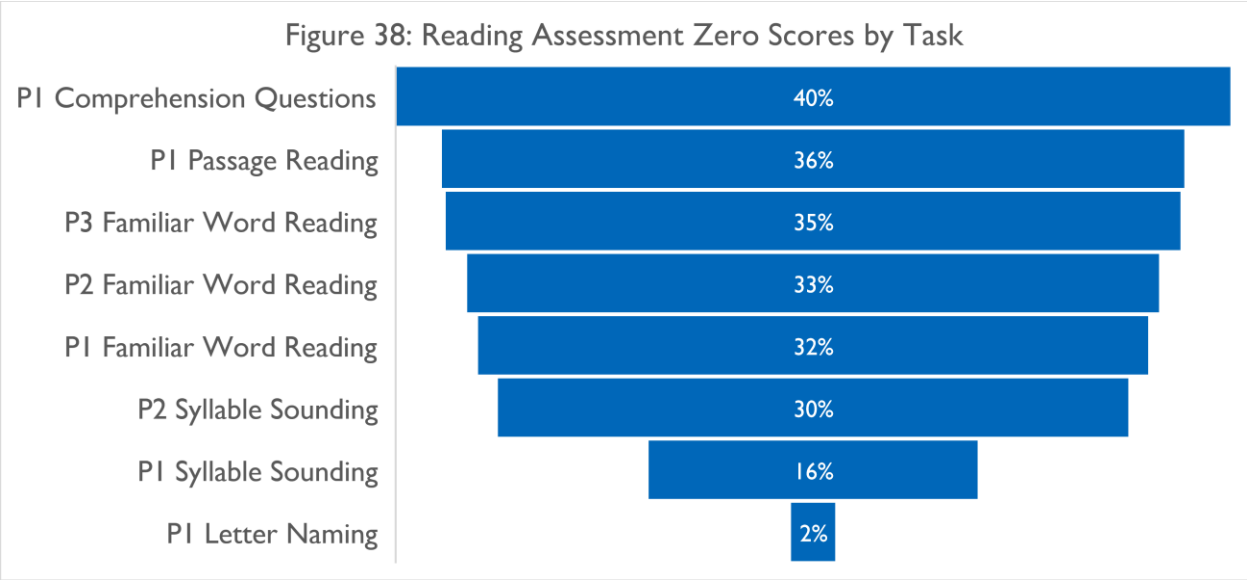
FIGURE 37: READING COMPREHENSION QUESTIONS CORRECT DISTRIBUTION, MALE/FEMALE



ZERO SCORES

Zero scores by sub-task and grade level show a marked drop in student skills between P1 and P2 syllable sounding, with zero scores jumping from 16% for P1 items to 30% for P2 items. Given the assessment was administered at the end of the P2 school year, all P2 syllables included in the assessment should have been covered in instruction. Interestingly, zero scores increase as the sub-tasks get more difficult, indicating that students who have not learned level-appropriate syllables are not able to transition to word and passage reading. For example, P2 students with P2 syllable sounding zero scores are also unable to sound out even P1-leveled familiar words. This is aligned with theories of reading instruction that assert sufficient mastery of phonics is required to progress to word reading and reading connected text.

FIGURE 38: READING ASSESSMENT ZERO SCORES BY TASK



In Figure 39 at right, we rank ordered the classes/schools from the random sample by zero scores for ORF, with designation of urban and rural schools. We find that there are five schools with no zero scores at the top of the performance spectrum. The two worst performing schools have over 80% of students with zero scores on ORF. There are 23 schools with 50% or greater zero scores within the 85-school random sample.

FIGURE 39: PERCENTAGE OF ZERO SCORES BY SCHOOL, URBAN/RURAL

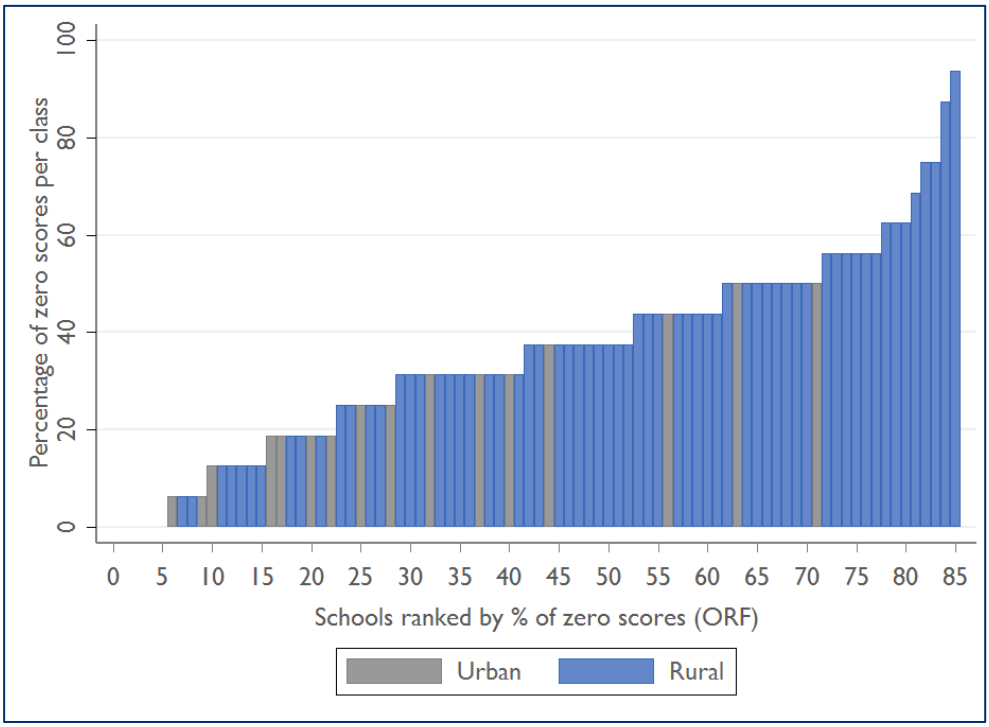
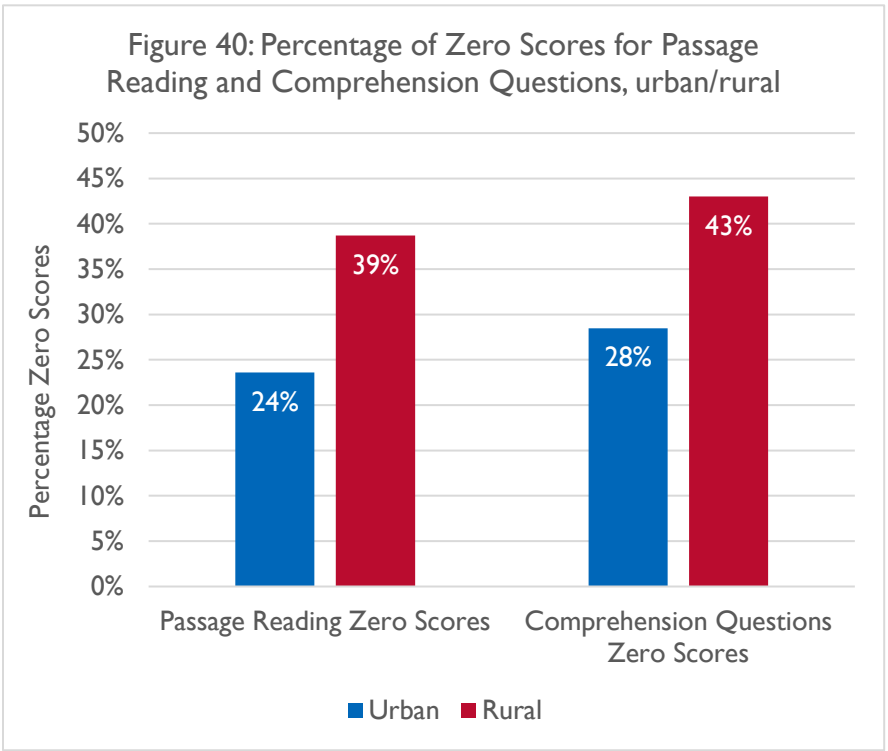


FIGURE 40: PERCENTAGE OF ZERO SCORES FOR PASSAGE READING AND COMPREHENSION QUESTIONS, URBAN/RURAL

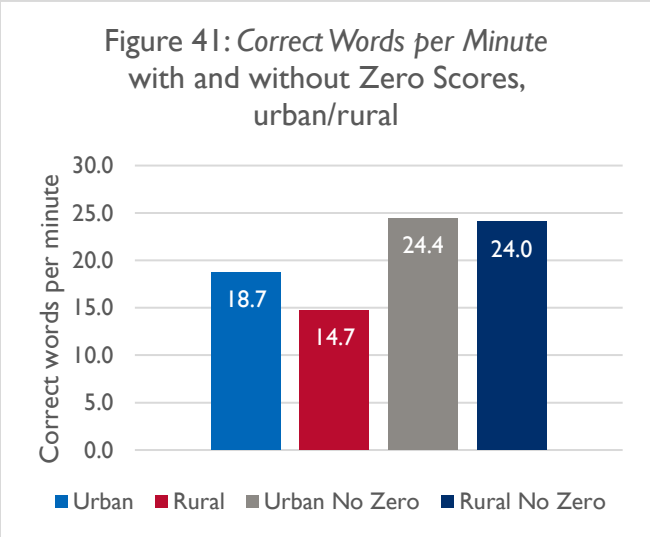


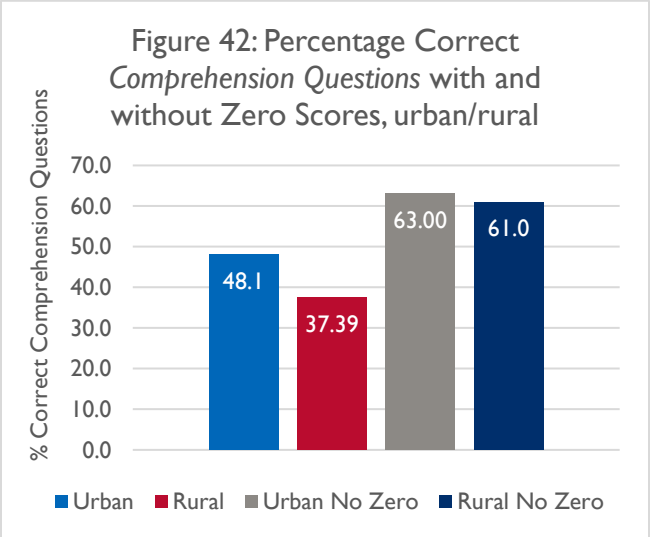
The difference in mean oral reading fluency and comprehension scores across urban and rural schools can be further explained by looking at the difference in zero scores and also how zero scores influence mean scores. Figure 40 illustrates the difference in zero scores while Figures 41 and 42 illustrate how the mean scores shift when zero scores are removed from the sample and the rural "disadvantage" in reading outcomes virtually disappears once zero scores are removed. These data indicate that for the children that can read at least one word in rural schools, their reading performance is nearly

equivalent to students in urban schools. These data also further underscore the substantial percentage of students in rural areas that being left behind as other students progress in their learning journey.

FIGURE 41: *CORRECT WORDS PER MINUTE* WITH AND WITHOUT ZERO SCORES, URBAN/RURAL

FIGURE 42: PERCENTAGE CORRECT *COMPREHENSION QUESTIONS* WITH AND WITHOUT ZERO SCORES, URBAN/RURAL

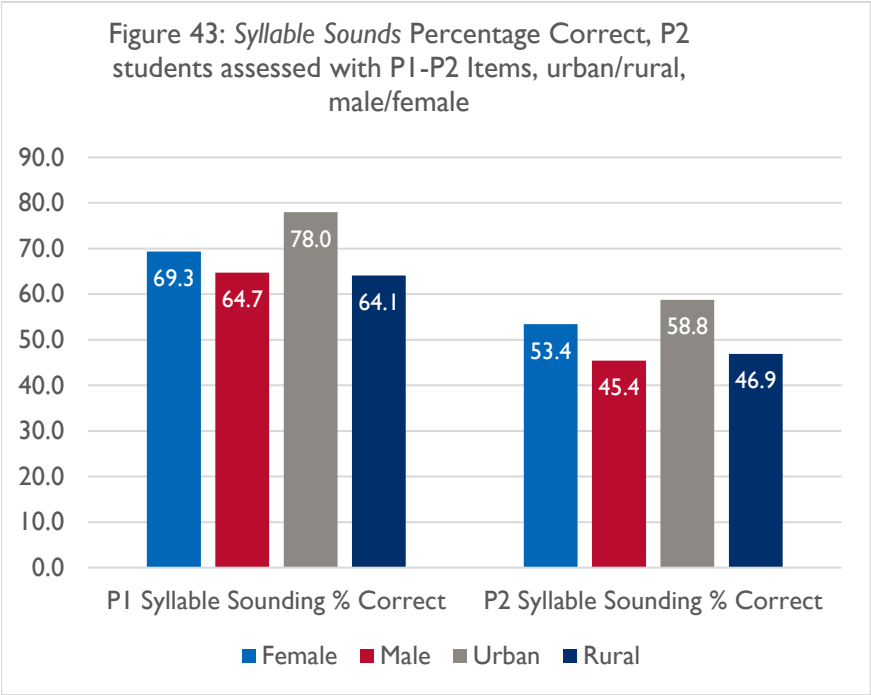




SUB-TASK SCORES, WITH DISAGGREGATIONS

FIGURE 43: SYLLABLE SOUNDS PERCENTAGE CORRECT, P2 STUDENTS ASSESSED WITH P1-P2 ITEMS, URBAN/RURAL, MALE/FEMALE

As we try to further understand how issues of equity are intersecting with reading outcomes, we analyzed scores across several sub-skills



disaggregated[49] by male/female and urban/rural. The data tell us that across the full sample, boys' average scores are equivalent to the average score of all rural students. This holds for syllable sounding and familiar word reading across different grade-level items. The disadvantage to boys and the average rural student persists across P1- P2 leveled syllable sounds and P1-P3 leveled familiar words. Boys slightly outperform the average rural student in oral reading fluency, but not in reading comprehension.

FIGURE 44: FAMILIAR WORD READING PERCENTAGE CORRECT, P2 STUDENTS ASSESSED WITH P1-P3 ITEMS, URBAN/RURAL, MALE/FEMALE

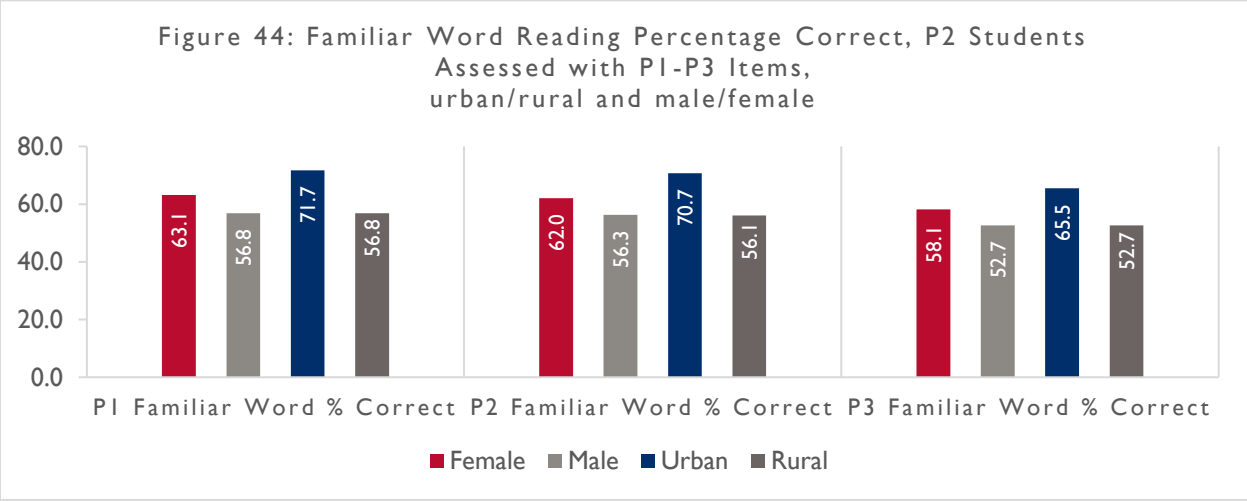
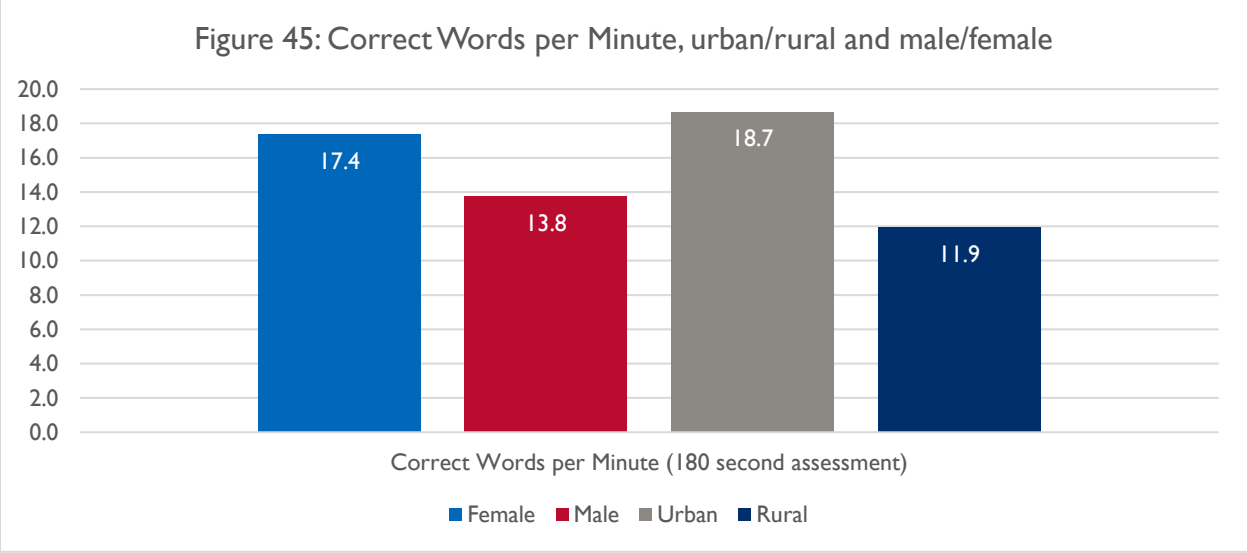
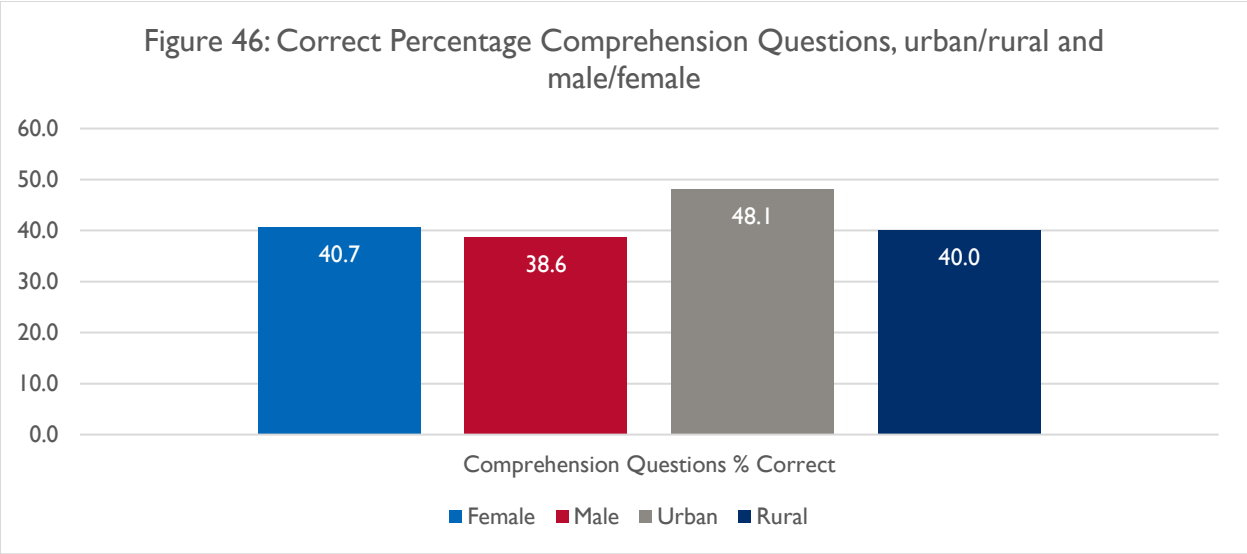


FIGURE 45: CORRECT WORDS PER MINUTE, URBAN/RURAL, MALE/FEMALE



[49] Note that the sample is not representative at the male/female or urban/rural disaggregated level. Care should be taken when interpreting the disaggregated findings.

FIGURE 46: CORRECT PERCENTAGE COMPREHENSION QUESTIONS, URBAN/RURAL, MALE/FEMALE



FINDINGS: RELATIONSHIP BETWEEN TIME AND READING SKILLS

A key part of this study is to identify if there is a relationship between time and reading outcomes.[50] We assessed this relationship by looking at absence rates and several different levels of time (in-classroom time, learning time, and active/passive learning time) as compared to tw9 student outcomes: oral reading fluency scores and reading comprehension scores.[51] The following is a summary of these analyses.

Overall in-classroom time and oral reading fluency scores: We investigated the relationship between the amount of time lost due to absences and tardiness (both teacher and student) and oral reading fluency scores and did not find a significant relationship. Put differently, time lost due to teacher and student absence[52] and tardiness was not found to be associated with lower oral reading fluency scores. We hypothesize that we did not find a relationship between in-classroom time and oral reading fluency scores because the composition of in-classroom time includes a variety of time, including non-instructional, transition and learning time, and some of this in-classroom time advances learning (learning time) and some of it does not support learning (e.g., non-instructional time).

Composition of in-classroom time and oral reading fluency scores: We then investigated the relationship between the composition of in-classroom time and oral reading fluency scores by including the type of instructional activities teachers directed during in-classroom time (transition, assessment/homework, active and passive reading activities) in the model. More specifically, we assessed the first three categories of in-classroom time against passive reading activities (our reference category).[53] We

[50] Additional information about the statistical model and assumptions can be found in Annex E.

[51] While the reading assessment items' level of difficulty are aligned with the Rwandan Kinyarwanda curriculum, this study does not endeavor to determine the degree to which P2 students have mastered the curriculum.

[53] Our four categories sum up to 100% of in-classroom time spent on teaching and are measured as percentages. Therefore, a unit change gives the effect of a one percentage point change.

controlled for several child and household characteristics (gender, age, urban or rural area, whether household has electricity, whether household head read) and show results with and without clustered standard errors at the school level.

With respect to active instruction, we found a significant relationship with oral reading fluency scores and reading comprehension scores. The association is significant at the 1% level without clustering the standard errors on both outcomes, and at the 10% level and 5% for oral reading fluency and comprehension scores, respectively, with clustered standard errors. In terms of effect sizes, when we model a 10% increase in active instruction (and a decrease of the same magnitude in passive instruction) we found this increase is associated with an increase of 0.936 in oral reading fluency (cwpm) and of 0.117 in reading comprehension score, both equivalent to a 0.06 standard deviation increase. In oral reading fluency, this means getting one extra word correct per minute.

Other expected results include that older children and boys perform worse than their counterparts and having a head of household who is able to read is associated with an increase in a student's scores. Table 9 below outlines the results of the regressions in full.

TABLE 9: RESULTS OF REGRESSION ANALYSIS

VARIABLES	(1) ORF (cwpm)	(2) ORF (cwpm)	(1) Total correct Reading Comprehension	(2) Total correct Reading Comprehension
Active instruction (proportion of in-classroom time)	0.0936***	0.0936*	0.0117***	0.0117**
	(0.0354)	(0.0528)	(0.00444)	(0.00573)
Transition activities (proportion of in-classroom time)	0.160***	0.160**	0.0132**	0.0132*
	(0.0494)	(0.0621)	(0.00619)	(0.00734)
Assessment/homework (proportion of in-classroom time)	0.107**	0.107	0.00967	0.00967
	(0.0520)	(0.102)	(0.00652)	(0.00956)
Proportion of time lost of REB allocated time	-0.559	-0.559	-0.472	-0.472
	(2.887)	(4.100)	(0.362)	(0.444)
Child lives in urban area	2.437**	2.437	0.391***	0.391*
	(1.091)	(1.831)	(0.137)	(0.217)
Child's age	-1.047***	-1.047***	-0.119***	-0.119***
	(0.263)	(0.319)	(0.0330)	(0.0397)
Child is a boy	-3.313***	-3.313***	-0.0620	-0.0620
	(0.800)	(0.861)	(0.100)	(0.0988)
Head of household is not the father	-0.322	-0.322	0.0121	0.0121
	(0.933)	(0.979)	(0.117)	(0.125)
Head of household reads	4.046***	4.046***	0.542***	0.542***
	(1.031)	(1.024)	(0.129)	(0.134)
Household has electricity	1.192	1.192	0.0346	0.0346
	(0.846)	(0.930)	(0.106)	(0.114)
Constant	13.71***	13.71**	1.687***	1.687***
	(3.763)	(5.417)	(0.472)	(0.637)
Observations	1,356	1,356	1,356	1,356
R-squared	0.064	0.064	0.046	0.046

Notes: Standard errors in parentheses. Models (1) and (2) show results with and without clustered standard errors at the school level, respectively. Significance levels *** p<0.01, ** p<0.05, * p<0.1

DISCUSSION

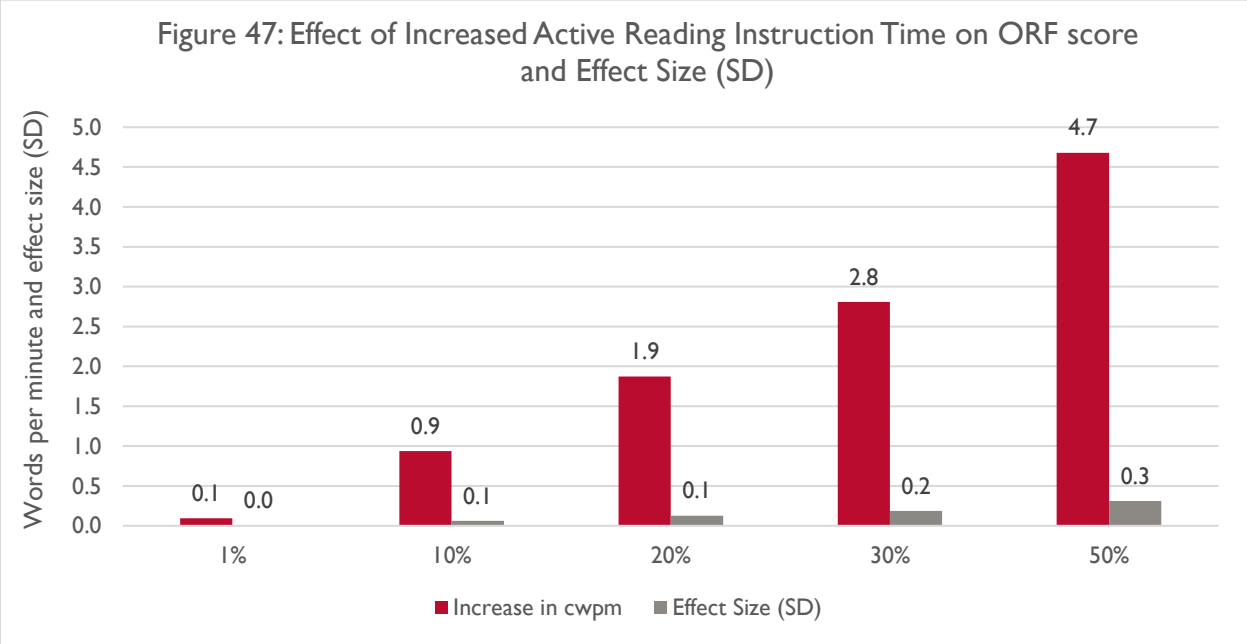
To place the results into context, we converted the increase in reading scores associated with increased active reading instruction into standard deviations. This helps us understand how instructional time changes could translate into improved outcomes in the same language used by evaluations and other research assessing the effectiveness of early grade reading interventions. We also modeled multiple levels of increased active reading instruction to understand at what level improvements in reading scores become meaningful in Table 10 and Figure 47 below.

TABLE 10: MODELING EFFECT SIZE OF ACTIVE INSTRUCTIONAL TIME

Time increase in active reading instruction	Increase in cwpm	Effect Size (SD)	Increase in Comp. Score	Effect size (SD)
1%	0.09	0.01	0.01	0.01
10%	0.94	0.06	0.12	0.06
20%	1.87	0.13	0.23	0.13
30%	2.81	**0.19**	0.35	**0.19**
50%	4.68	0.31	0.59	0.31

The 1 cwpm increase in oral reading fluency (equivalent to .06 SD) associated with a 10% increase in active reading instruction time does not seem meaningful in the context of policy objectives, standards for reading outcomes, and effectiveness of early grade reading interventions in general. It is only with a 30% increase in active instruction that the improvements in reading scores become meaningful, with an increase in nearly 3 cwpm in reading fluency and an increase of .35 in comprehension scores and an effect size of .19 SD. Figure 47 below represents the same data in Table 10 in a different format.

FIGURE 47: EFFECT OF INCREASED ACTIVE READING INSTRUCTIONAL TIME ON ORF SCORE AND EFFECT SIZE (SD)



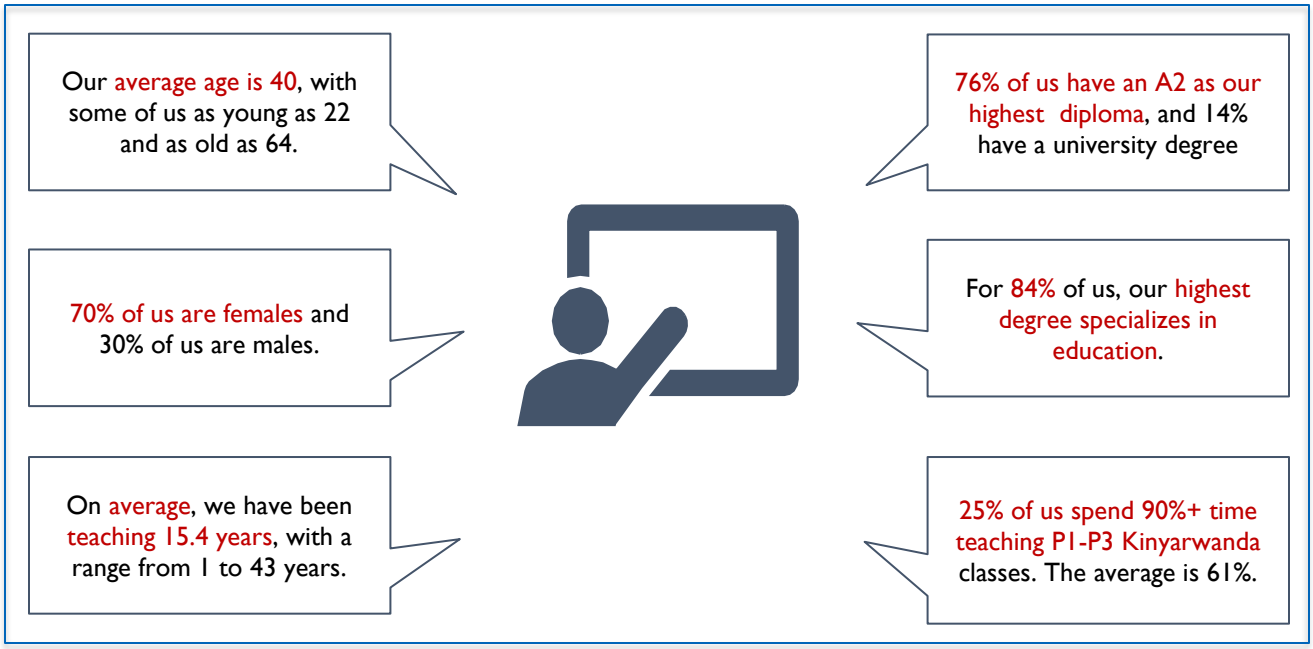
FINDINGS: TEACHER PROFILE

To deepen the analysis and connect classroom observation data with teacher characteristics and attitudes, all teachers whose instruction was observed participated in a survey that included the following topics:

- Background profile (e.g., gender, age, number of years teaching, academic qualifications, non-teaching jobs, transportation)
- Teaching profile (e.g., number of subjects taught, number of hours taught)
- Perceptions of student reading skills
- Perceptions of teaching
- Lesson preparation and administrative tasks
- Motivation related to Kinyarwanda lesson planning
- Individual self-efficacy related to reading instruction and class management
- Support and feedback received
- Expectations of student learning

Figure 48 below provides summary information about the teachers included in the 85-school random sample. There is a broad range in the number of years teaching Kinyarwanda, from 1 to 43 years, with an average of 15.4 years across the sample. 84% of teachers' highest degree specializes in education, and teachers are generally involved in teaching multiple subjects, with only 25% spending 90% of their teaching time exclusively focused on early grade Kinyarwanda.

FIGURE 48: TEACHER PROFILE



TEACHER WORK LOAD

HOURS. Our P2 Kinyarwanda teachers teach an average of 30.2 hours per week. Around 42% teach 30 hours or less per week, and 58% teach more than 30 hours per week.

SUBJECTS. They teach 3.1 subjects on average. Only 14% teach one subject (Kinyarwanda), 25% teach two subjects, 22% teach three subjects, and 39% teach four subjects or more. About 40% of Kinyarwanda teachers also teach social religious or math classes, and 24% of Kinyarwanda teachers also teach English.

GRADES. They teach 1.7 grades on average. 44% only teach one grade, 45% teach two grades, and 11% teach three or more grades. All teachers in the sample teach P2; 32% also teach P1, and 26% teach P3. 14% or fewer teach P4-P6 in addition to P2.

SECTIONS. They teach 8.6 sections[54], across all subjects, on average. Twenty-nine percent teach five to six sections, 34% teach seven to eight sections, and 36% teach 10-16 sections. Teachers teach an average of 3.5 Kinyarwanda sections and 5.1 sections from other subjects. Note some subjects have less hours allocated, so the latter does not necessarily translate in more workload.

TEACHER SELF-REPORTED REASONS FOR ABSENCE AND LATENESS

Teacher who arrived late for their observed lesson were asked the reason for their lateness during a brief interview following the observation. 50% of the teachers reported that they were gathering students who were late to class, 33% reported that they were late because they were having lunch or had very little time for lunch, and the balance reported varied reasons including meetings or personal errands.

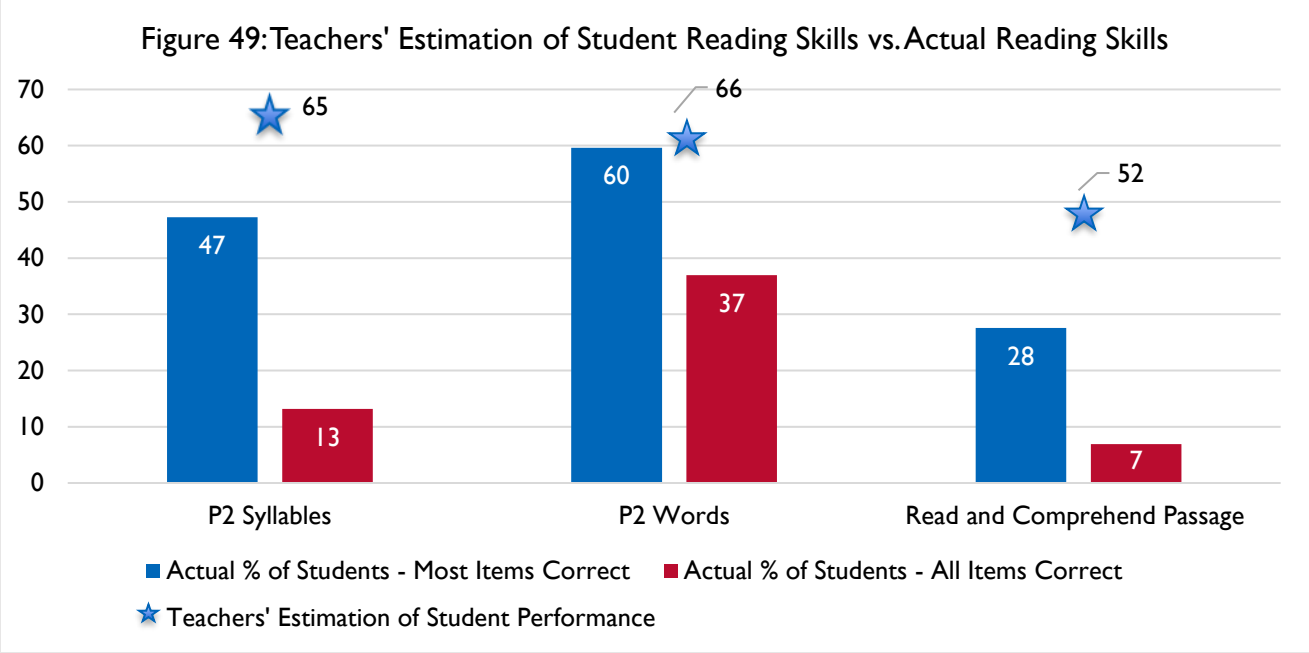
[54] A section is a 40-minute class period on average.

In terms of teacher absence, the predominant reason provided by administrators was that the teacher was ill, or he or she had a family member that was ill.

PERCEPTIONS OF STUDENT READING SKILLS AND CAPACITY TO READ

As discussed in a section above, we assessed the reading skills of students in the classrooms that we observed. We also asked the teachers who conducted those classes to estimate the reading skills of their students. The results are depicted in Figure 49 below and shows that teachers consistently over-estimate their students' reading skills, and in some cases, substantially. With respect to estimates of passage reading and comprehension, the teachers estimated that 52% of their students could perform this task with a P2-leveled passage and set of questions. In reality, only 28% of students scored 80% on this task and only 7% scored 100% using a P1-leveled passage and set of questions.

FIGURE 49: TEACHERS' ESTIMATION OF STUDENT READING SKILLS VS. ACTUAL READING SKILLS



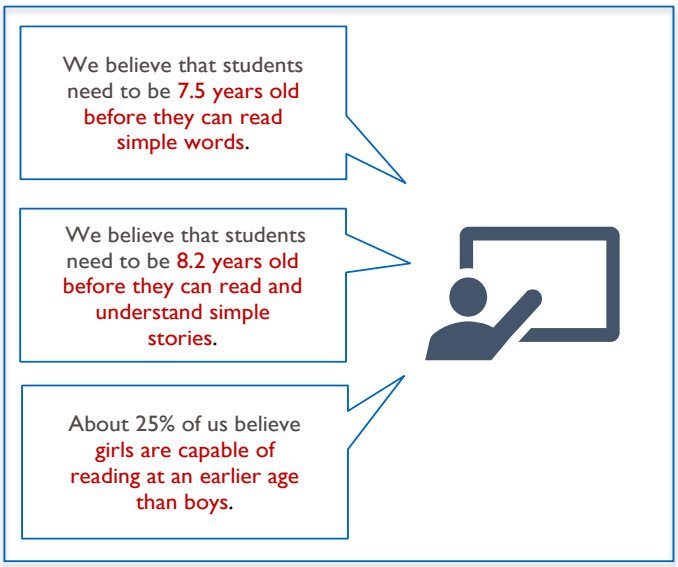
55

[55]Teachers were asked to estimate the percentage of their P2 students (in observed class) that could: a) read most P2 syllables, b) read P2 words and c) read and comprehend a P2-leveled passage. Students were assessed on P2-leveled syllables and words and a P1-leveled passage and comprehension questions.

We also asked teachers their thoughts about what age children had *the capacity* to perform certain reading skills (using P2-leveled examples). These estimations are based on average teachers' perceptions of children's potential rather than the current performance of students in their class.

Given the average age of P2 students in our sample is 9, and the student mean scores listed above in Figure 49, we see that teachers do believe that students' capacity is above the level at which students are currently performing.

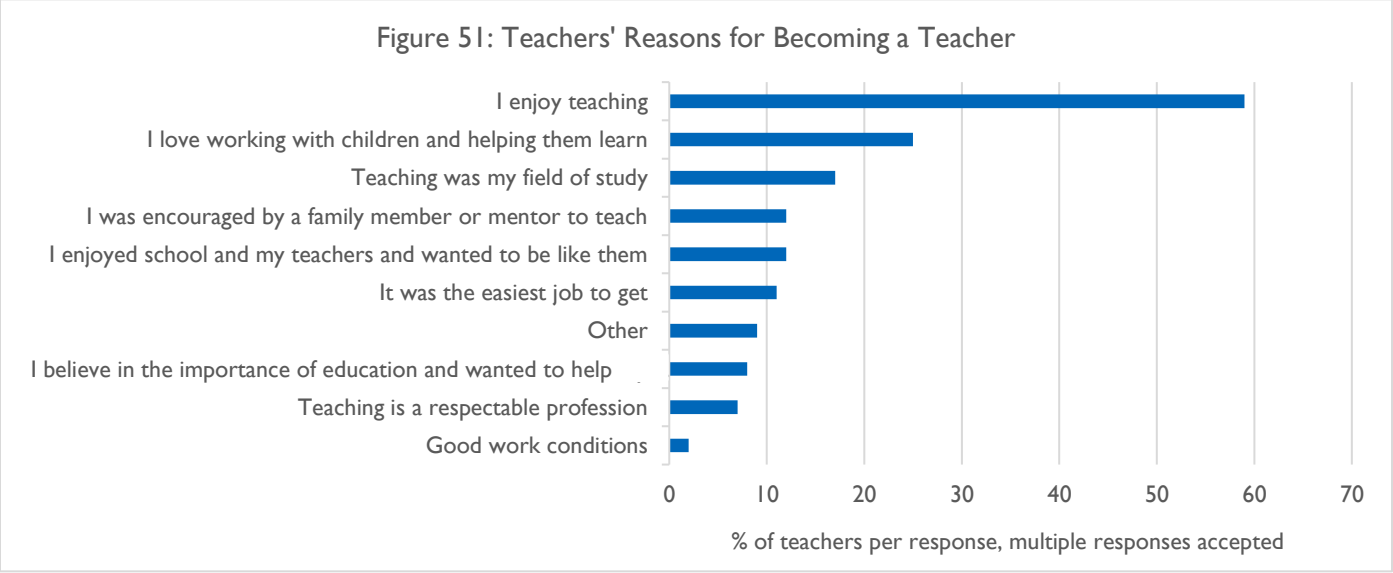
FIGURE 50: TEACHERS' BELIEFS ABOUT LEARNING



PERCEPTIONS OF TEACHING

71% of teachers stated that teaching was their first choice of a career and that they did not pursue other job options. This aligns with the 84% of teachers who report that their highest academic qualification specializes in education. We wanted a deeper understanding of teachers' perceptions of their careers and what originally motived them to become a teacher and Figure 51 below reflects high frequency of reported intrinsic motivation to become a teacher.

FIGURE 51: TEACHERS' REASONS FOR BECOMING A TEACHER



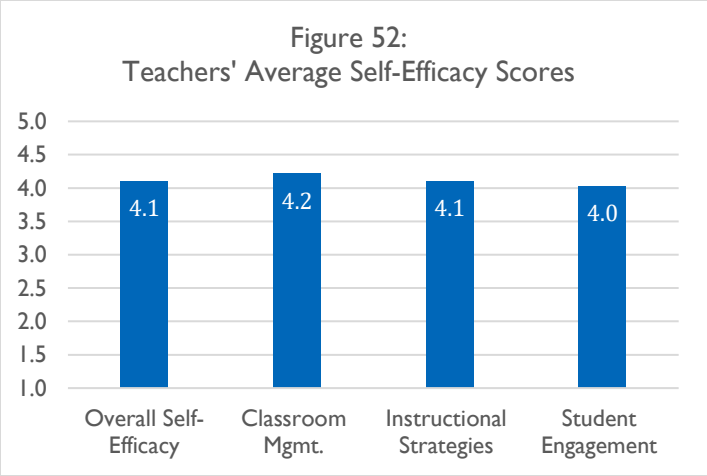
Notwithstanding this initial intention to become a teacher reported by many respondents, roughly 50% of teachers state they would leave the teaching profession if they had another opportunity. This might be related to teaching salary given that 21% of teachers noted increased salary as a strategy that could

make them more effective in supporting students' reading development.[56] Only 5% of teachers reported having a paid job outside of their teaching position and on average, teachers spend 5 hours working on a family farm or business per week.

TEACHER SELF-EFFICACY: GENERAL AND READING INSTRUCTION

FIGURE 52: TEACHERS' AVERAGE SELF-EFFICACY SCORE

The self-efficacy scale used in this study was adapted from the Teachers' Sense of Efficacy scale[57] and includes three subscales: classroom management, instructional strategies, and student engagement. The scale was piloted with 78 respondents and assessed for reliability using the Cronbach alpha statistic. This statistic is a measure of the correlation among responses to the set of twelve questions. The alpha for the scale was .86 for the full sample, substantially higher than the .7 threshold commonly used to determine acceptable reliability.



Teachers responded to a set of 12 questions on a scale from 1 "nothing at all" to 5 "a lot." Illustrative questions include:

- To what extent can you implement effective reading strategies in your classroom? (IS)
- How much can you do to meet the needs of struggling readers? (SE)
- How much can you do to get children to follow classroom rules? (CM)
- How much can you motivate students who show low interest in reading? (SE)

We see that teachers' reported self-efficacy scores are consistently high across the sub-scales with an overall mean score of 4.1 out of 5, indicating that teachers feeling relatively empowered as they manage their classrooms, deliver instruction and engage with students.

TEACHER MOTIVATION: LESSON PLANNING

The teacher survey also included a teacher motivation scale, which was adapted from the Work Tasks Motivation Scale for Teachers.[58] This scale includes five subscales: intrinsic motivation, introjected regulation, identified regulation, external regulation and amotivation. Respondents were asked to focus on lesson planning and preparation specifically as they considered the questions. The scale was piloted in three rounds with a total of 78 respondents and assessed for reliability using the Cronbach alpha statistic. This statistic is a measure of the correlation among responses to the set of fifteen questions.

---

[56] A pay increase for teachers was in-process during data collection for this study and these responses are not reflective of that pay increase.

[57] https://cpb-us-w2.wpmucdn.com/u.osu.edu/dist/2/5604/files/2018/04/TSES-scoring-zted8m-1s63pv8.pdf

[58] Definitions of motivation sub-elements and more information on the scale can be found at https://selfdeterminationtheory.org/SDT/documents/2008_FernetEtAl_WTMST-JCA.pdf

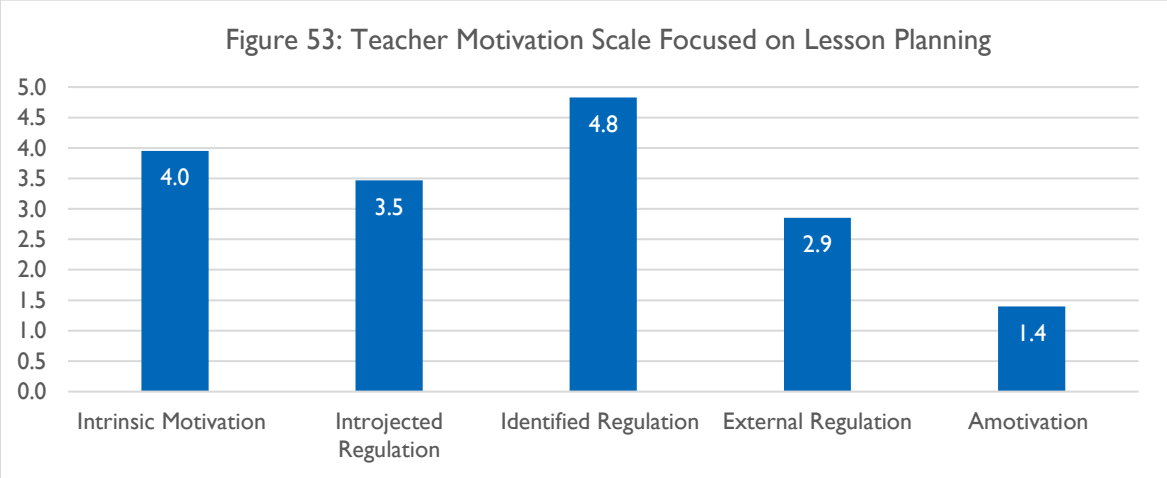
The alpha for the scale was .66 for the full sample, slightly below the .7 threshold commonly used to determine acceptable reliability. The reliability across sub-scales varied from .73 (intrinsic motivation) to .48 (identified regulation). These low alphas could be a function of unclear translation, sub-constructs that are too divergent, poor comprehension of the questions, etc.

Examples of questions for each sub-scale include:

- Because I find this task important for the academic success of my students. (IR)
- Because the school system requires me to do it. (ER)
- I don't know, sometimes I don't see its purpose. (AM)
- Because I find this task interesting to do. (IM)
- Because I would feel guilty not doing it. (IR)

Figure 53 illustrates that teachers surveyed have fairly high levels of motivation overall with regards to lesson planning and preparation, with identified regulation (see examples above) scoring highest and very low levels of amotivation. Based on these findings, teachers' motivations seem to be driven by proactive attitudes and a recognition that their actions are tightly linked with student outcomes, rather than drivers stemming from external requirements or because they are required to perform these tasks only because their supervision or school system requires them to do so.

FIGURE 53: TEACHER MOTIVATION SCALE FOCUSED ON LESSON PLANNING

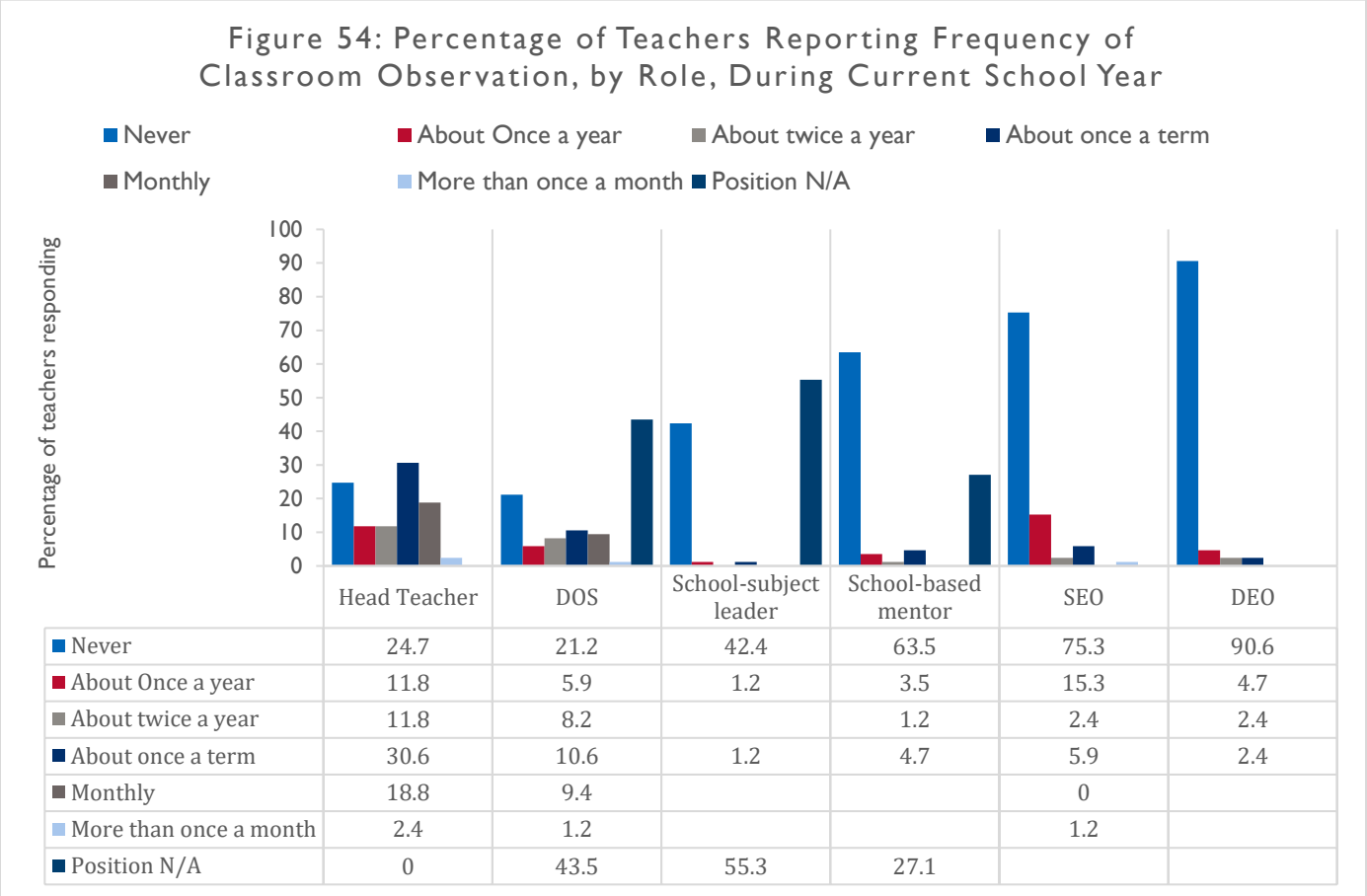


SUPPORT AND FEEDBACK RECEIVED

Teachers were surveyed about who observes their lessons and how frequently, as well as the support and feedback they receive based on these observations. 21.2% of teachers reported that their head teacher has observed their class monthly or more frequently in the past school year, and 10.6% of teachers reported the director of studies has observed monthly or more frequently in the past school year.[59] There were no reports of school-subject leaders or school-based mentors observing classes monthly or more frequently.

[59] Note that Director of Studies are only present in 56.5% of schools.

FIGURE 54: PERCENTAGE OF TEACHERS REPORTING FREQUENCY OF CLASSROOM OBSERVATION, BY ROLE, DURING CURRENT SCHOOL YEAR



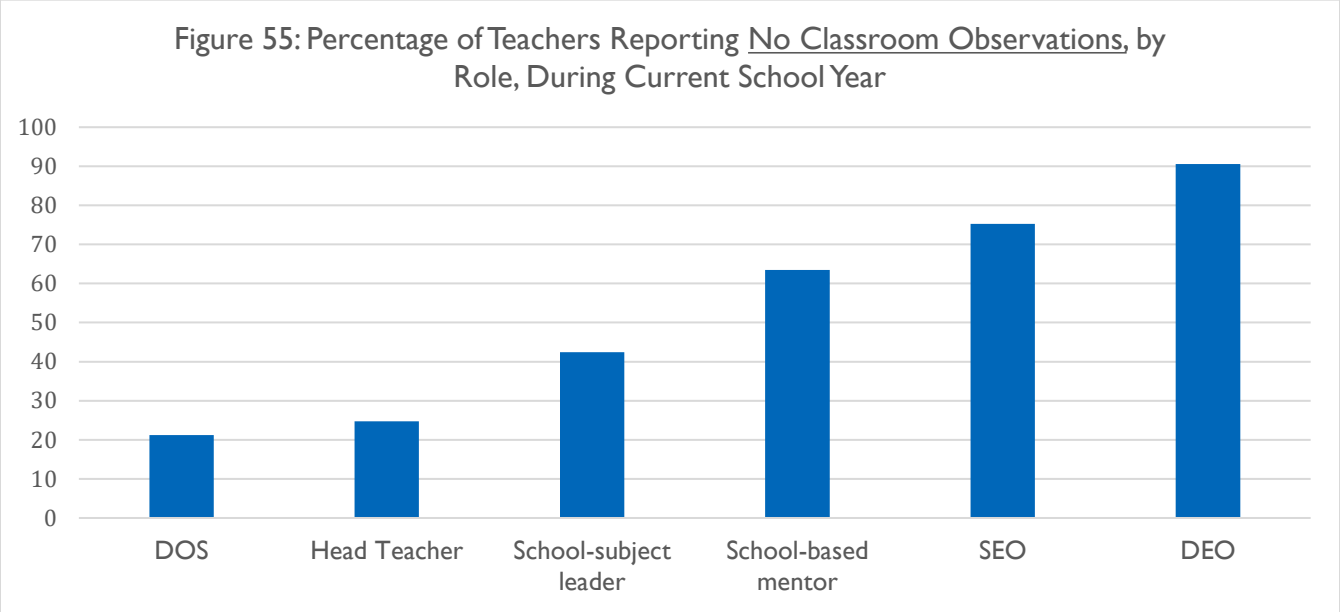
	Head Teacher	DOS	School-subject leader	School-based mentor	SEO	DEO
Never	24.7	21.2	42.4	63.5	75.3	90.6
About Once a year	11.8	5.9	1.2	3.5	15.3	4.7
About twice a year	11.8	8.2		1.2	2.4	2.4
About once a term	30.6	10.6	1.2	4.7	5.9	2.4
Monthly	18.8	9.4			0	
More than once a month	2.4	1.2			1.2	
Position N/A	0	43.5	55.3	27.1		

60

In addition, 90.6% and 75.3% of teachers reported DEOs and SEOs (respectively) never observed one of their lessons in the past school year, and 15.3% and 4.7% of teachers reported that DEOs and SEOs (respectively) observed their instruction about once a year. Strikingly, 24.7% of teachers report their head teacher had not observed one of their lessons in the past year. There remains a substantial percentage of teachers, 22%, who report not having anyone observe their classroom within the school year. Figure 55 breaks this down by role.

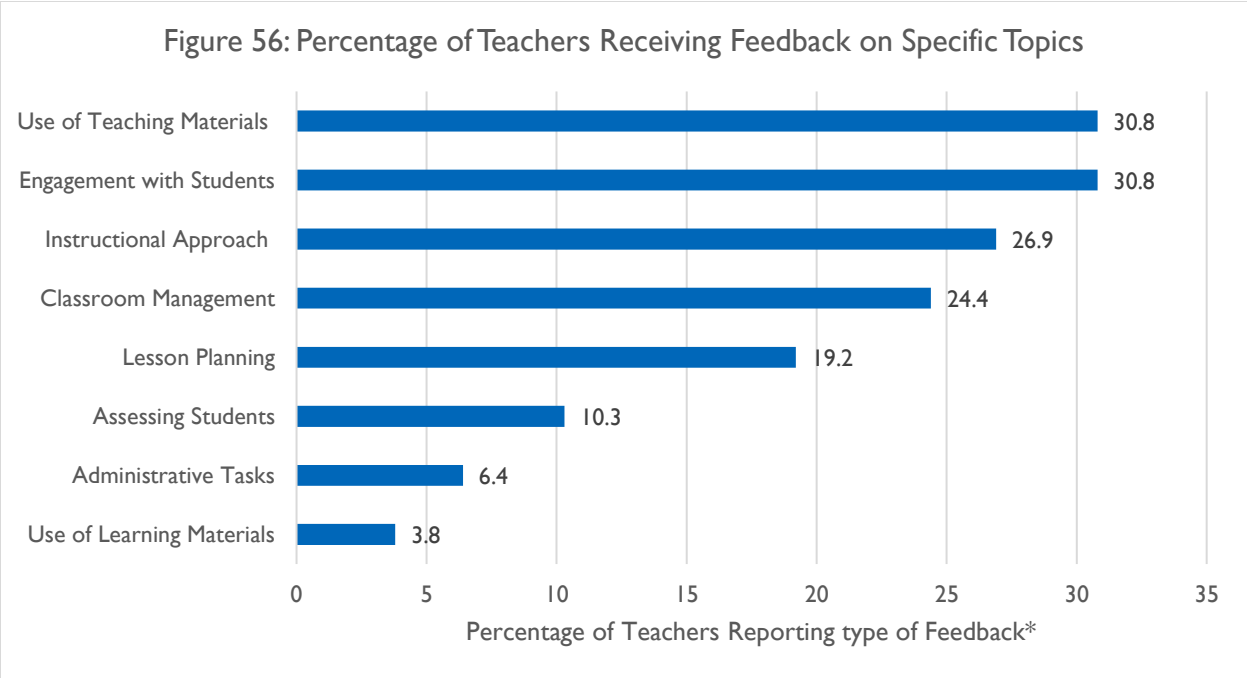
[60] "Position N/A" indicates schools for which that position is not applicable. For example, directors of studies are only found in schools with secondary level classes.

FIGURE 55: PERCENTAGE OF TEACHERS REPORTING NO CLASSROOM OBSERVATIONS, BY ROLE, DURING CURRENT SCHOOL YEAR



78% of teachers in our sample were observed by someone at least once in the past school year and of these 78% a very high percentage (98.7%) reported receiving feedback after the observation. The focus of the feedback is outlined in the graph below, showing that both the use of teaching materials and engagement with students were the most prevalent topics for feedback. Instructional approach and classroom management both had relatively high frequency rates as well.

FIGURE 56: PERCENTAGE OF TEACHERS RECEIVING FEEDBACK ON SPECIFIC TOPICS

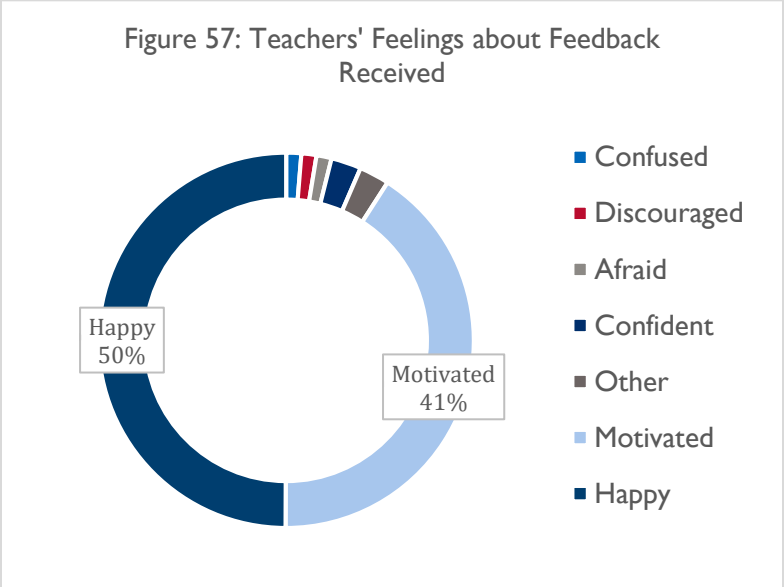


**more than one response accepted.*

For 96.2% of teachers, the feedback they received from the person who observed their class most often was helpful. 100% of teachers report taking action as a result of the feedback.

Teachers had overwhelmingly positive feelings associated with the feedback they received, with 91% of teachers reporting feeling motivated or happy after receiving feedback.

FIGURE 57: TEACHERS' FEELINGS ABOUT FEEDBACK RECEIVED



RESULTS OF REGRESSIONS: TEACHER CHARACTERISTICS AND STUDENT READING SCORES

Reflecting on the range of data about teachers collected in this study, we wanted to further understand how teacher characteristics might be associated with student reading scores. To investigate this further, we conducted multiple regressions, exploring different teacher characteristics, while controlling for student and household factors.[61] We tested these relationships across P2 syllable sounding percent correct, P2 familiar word percent correct, oral reading fluency, and percent correct of reading comprehension (we did so with and without clustering standard errors at the school/class level). The findings discussed below are with clustering standard errors at the school level.[62]

Teachers' gender and number of years teaching was not found to be associated with student reading scores. Interestingly, the teacher variable we did find most strongly associated (negatively) with student reading scores was whether the teacher exclusively teaches Kinyarwanda lessons (rather than Kinyarwanda and other subjects). We observe that students being taught by teachers who exclusively teach Kinyarwanda, scored 19.5% fewer correctly sounded syllables, read 22% fewer familiar words correctly, and read on average 6.8 words less per minute.[63] To put these findings in context, only 13% of teachers in our sample exclusively teach Kinyarwanda. Teachers who only teach Kinyarwanda also are

[61] The common student and household controls used for all regressions in this study include: age, sex, urban/rural, wealth index, and head of household literate.
[62] See Annex E for a description of adjusted standard errors.
[63] All of these are significant at the >.1% level. Reading 6 more words per minute is equivalent to the average annual gain in oral reading fluency in Rwanda, so the effect/coefficient is meaningful.

less likely to have teaching qualifications (81% vs. 94%), are younger on average (36 years vs 40 years) and have been teaching for fewer years (12 years vs. 15.7 years), but none of these differences are statistically significant. Further studies should explore who these teachers are that only teach Kinyarwanda and how their preparation and instructional quality differs from teachers who teach more than once subject.

Teachers whose qualification/degree specialized in education had students with an average of 10% higher scores on percentage correct syllable sounding and more than 10.5% higher scores on familiar word reading. They also read over 3 words per minute[64] more than students whose teachers did not specialize in education. The significance of the association of the latter two variables is observed in our sample but should be taken with caution as significance disappears when we adjust the standard errors for the clustering we observe.[65]

We also looked at the relationship between our measures of teacher motivation and self-efficacy and student reading scores. There were some significant findings for two of the motivation sub-scales (external regulation and amotivation), but we are not reporting those as meaningful because the measurement reliability for those sub-scales is not sufficiently high.

For the teacher self-efficacy measure, we do find a significant positive relationship (i.e., 7.267* on syllable sounding and 2.728** on ORF, see Annex E for other outcomes) between teachers' self-reported instructional strategy strengths (one of the self-efficacy sub-scales) and student reading skills. Instructional strategy items included:

- How much can you gauge student comprehension of what you have taught?
- To what extent can you implement effective reading strategies in your classroom?
- How well can you determine the reading skills of all students in your classroom?
- How much can you adjust reading strategies based on ongoing informal assessments such as quizzes, tests or homework of your students?

We also found a significant negative relationship (i.e., -5.764**on syllable sounding, see Annex E for other outcomes) with another self-efficacy sub-scale: classroom management. This sub-scale measures teachers' confidence in their ability to control disruptive behaviors or get students to follow classroom rules. These findings indicate that the strategies that teachers perceive as being effective in classroom management are potentially having a negative effect on student learning outcomes.

[64] Significant at the .5% level. The significance disappears when we add number of hours worked by the teacher and whether the teacher exclusively teaches Kinyarwanda.
[65] Co-linearity variance inflation factor was 16.6 and the threshold is 10, so findings should be taken with caution.

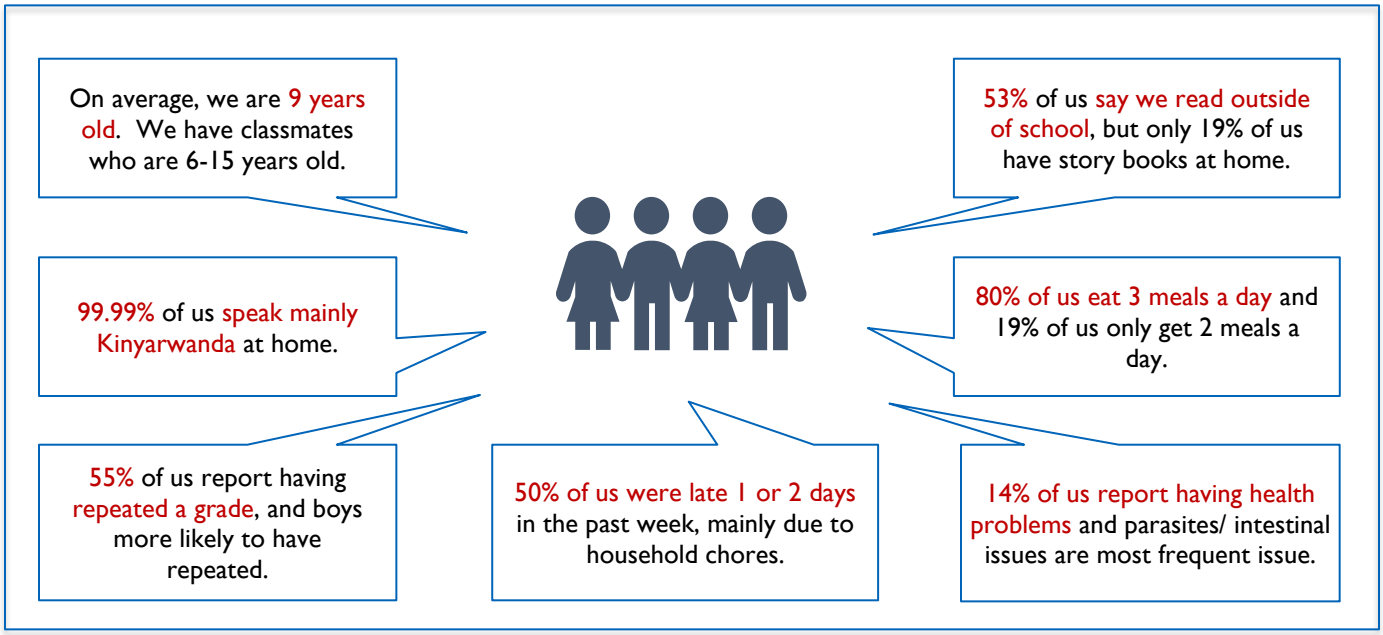
FINDINGS: STUDENT PROFILE

The sample included 1,360 Grade 2 students. 16 students (8 male/8 female) were randomly sampled from each of the classrooms that we observed. After providing assent for the survey and reading assessment, students answered a set of profile questions in advance of taking the reading assessment. These questions included:

- Basic demographic data such as sex and age
- Household information including: primary language spoken, head of household, household assets, etc.
- Health issues and number of meals per day
- Availability of reading material in the home and support for schooling/reading at home
- Grade repetition, tardiness and reasons for tardiness

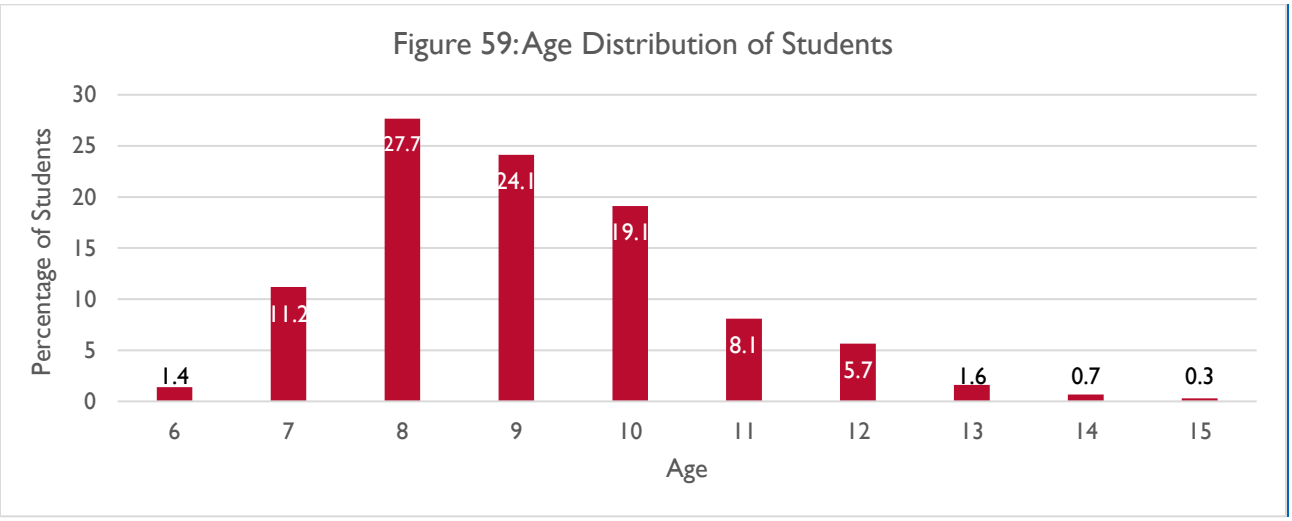
Who are the students in our sample and what are their households like?

FIGURE 58: STUDENT PROFILE



When surveying children in this age range, we expect some inconsistencies due to question comprehension challenges and desirability biases. As such, the findings should be interpreted with these factors in mind. Through the piloting process we did eliminate some items that seemed to elicit unreliable responses, such as asking students to estimate the amount of time they spent traveling to and from school each day. The range of responses for this question indicated that estimates of time for students this age were not consistently reliable. Other questions required additional definitions/examples and more detailed enumerator instructions, such as clarity about what constitutes a meal and what types of reading materials they have in their homes.

FIGURE 59: AGE DISTRIBUTION OF STUDENTS

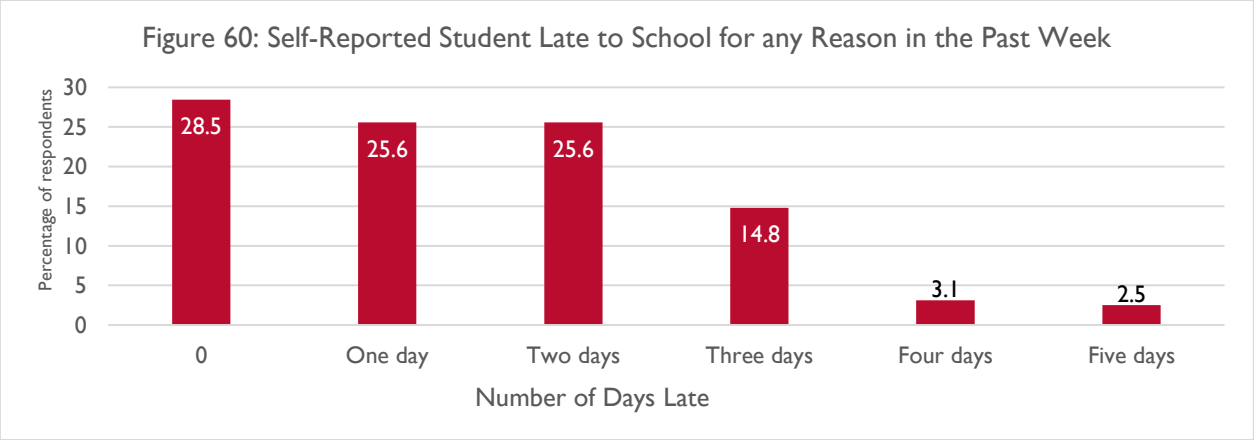


The average student age of students is 9 years, with some students reporting being as young as 6 years old and as old as 15 years. There were 3 students who responded that they did not know their age.

REPORTED REASONS FOR STUDENT ABSENCE AND TARDINESS

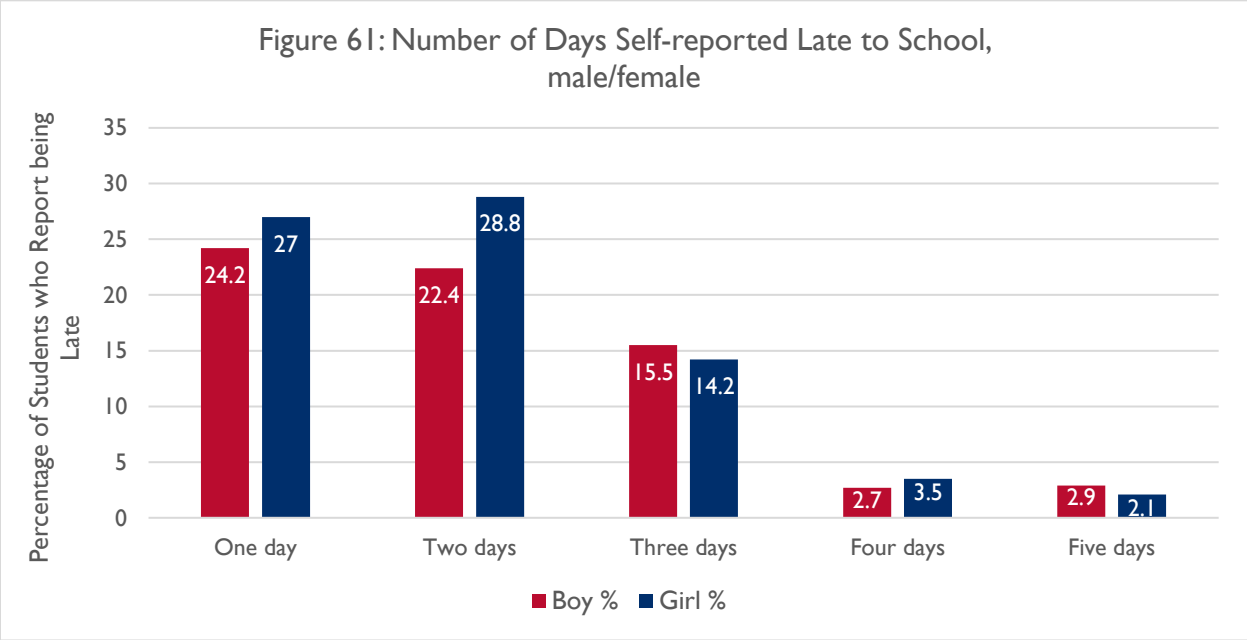
Students who participated in the reading assessment provided information about how frequently they were tardy in the past week and the reason for tardiness. 28% of students reported no tardiness, followed by a combined total of 50% of children reporting being tardy one to two days in the past week.

FIGURE 60: SELF-REPORTED STUDENT LATE TO SCHOOL FOR ANY REASON IN THE PAST WEEK



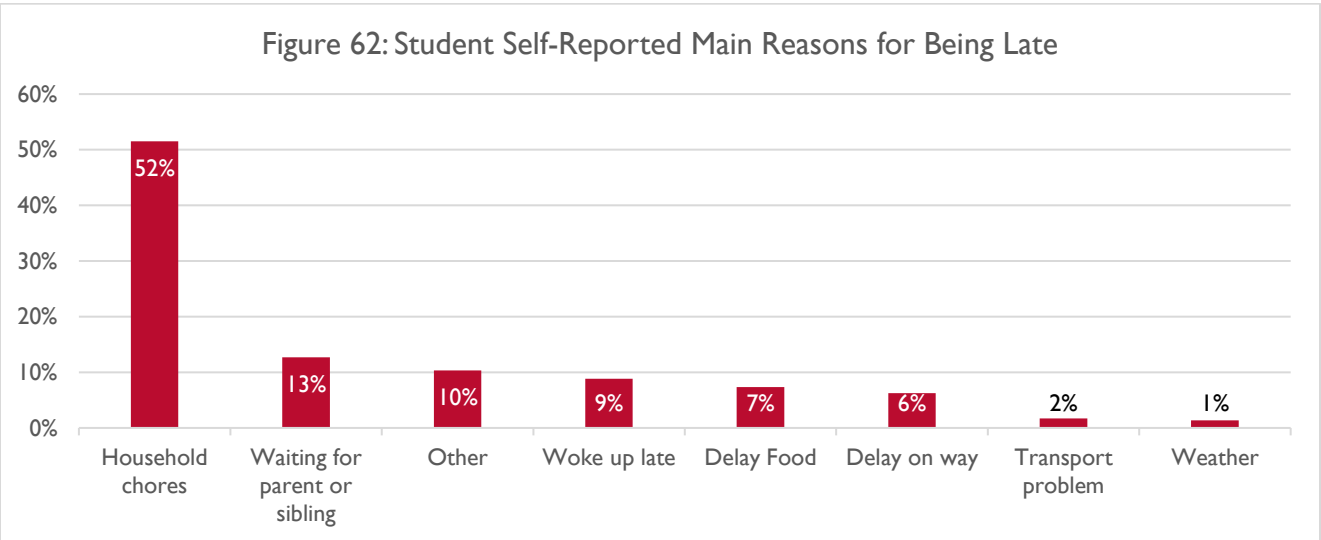
Female students report being late to school in the past week more often than their male counterparts with 76% of females reporting being late as compared to 68% of male students. 32.4% of male students and 24.5% of female students reported they were not late on any day in the past week.

FIGURE 61: NUMBER OF DAYS SELF-REPORTED LATE TO SCHOOL, MALE/FEMALE



The majority of students cited household chores as the reason for their tardiness. Reasons in "other" category included school uniform issue, illness of student or family member, no school materials, looking after livestock, etc.

FIGURE 62: STUDENT SELF-REPORTED MAIN REASONS FOR BEING LATE



12% of children gave a reason other than those listed above for being late. 19% of the "other" reasons were associated with uniforms (e.g., missing, waiting to dry) or other personal care reasons for being delayed.

FIGURE 63: PERCENTAGE OF STUDENTS REPORTING TYPE OF HOUSEHOLD CHORES ON A SCHOOL DAY

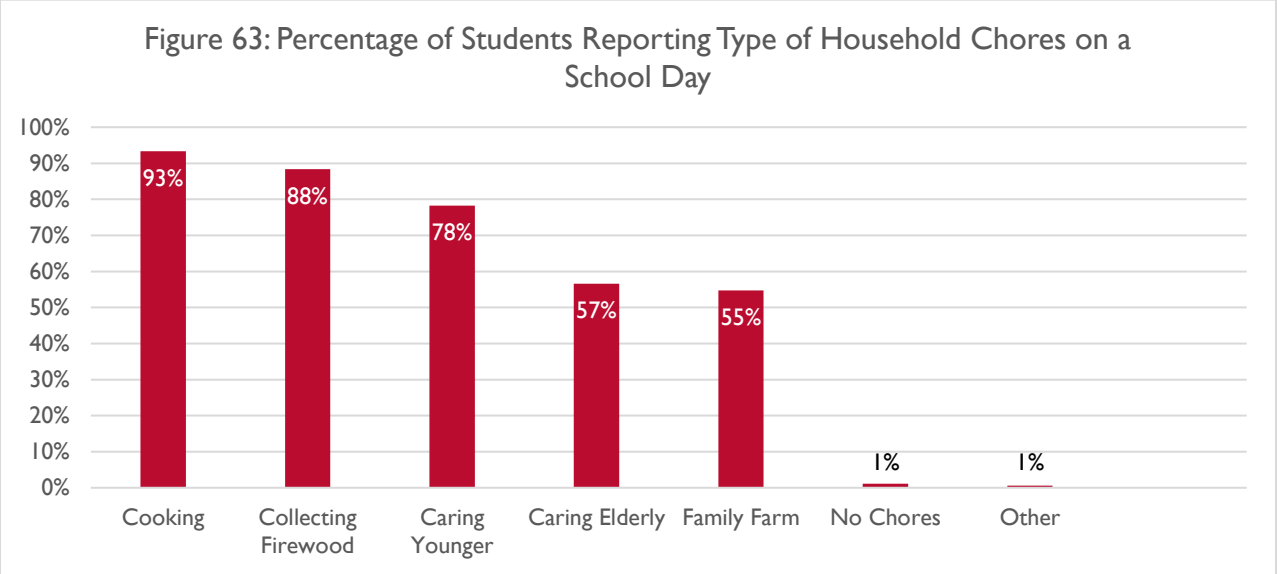
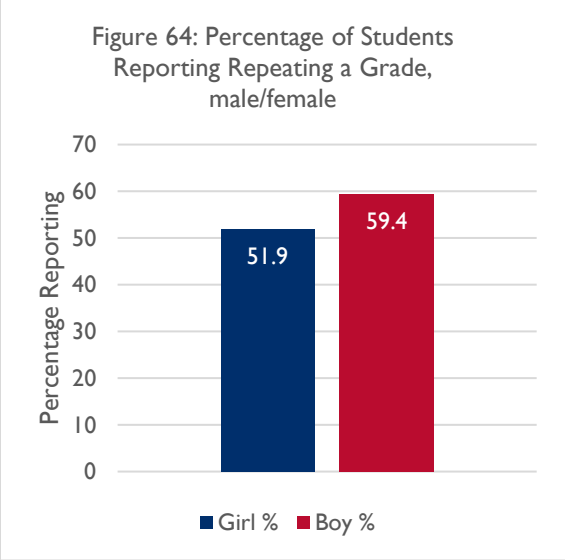


FIGURE 64: PERCENTAGE OF STUDENTS REPORTING REPEATING A GRADE, MALE/FEMALE

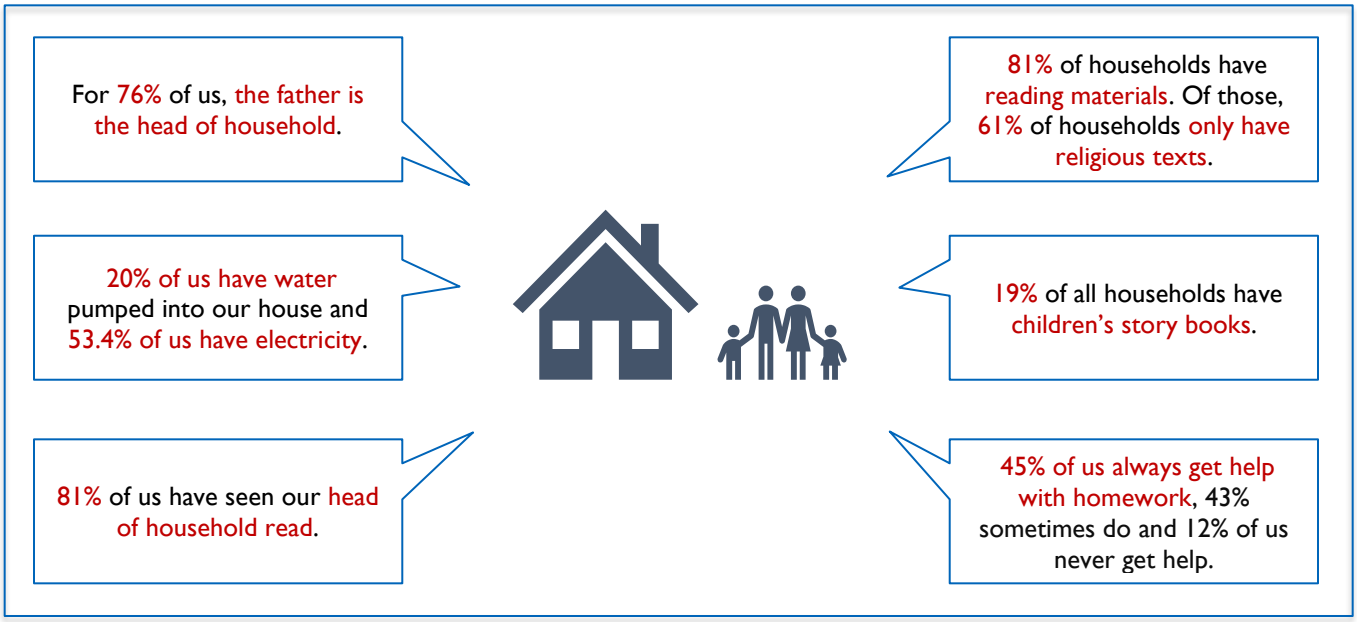


One of the more striking findings of the student profile is the high number of self-reported grade repetitions, with 55% of students reporting having repeated a grade. A higher percentage of male students reported repeating a grade as shown in Figure 64 to the left. Given the average age in our sample is 9 and 35.44% of the sample reports being 10 years or older, this repetition rate is possible. Referencing other related studies in Rwanda, we found reports of up to 44% annual student repetition rates in P1 to P3.[66]

[66] Friedlander, E. & Goldenberg, C. (eds.). (2016). Literacy Boost in Rwanda: Impact Evaluation of a 2-year Randomized Control Trial. Stanford, CA: Stanford University.

FIGURE 65: STUDENT HOUSEHOLD PROFILE

In terms of our students' households, we find that for the great majority (76%), the father is the head of household.[67] 81% of students reporting having seen their head of household read, which aligns with the percentage of students reporting their household has reading materials. Of the 81% of students that report having any type of reading materials (other than textbooks) in their house, 61% of households only have religious texts. Only 19% of all households have children's story books, compared to 78% with religious texts.

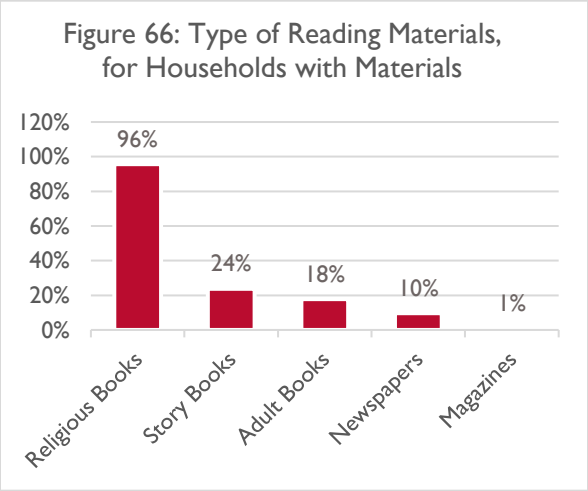


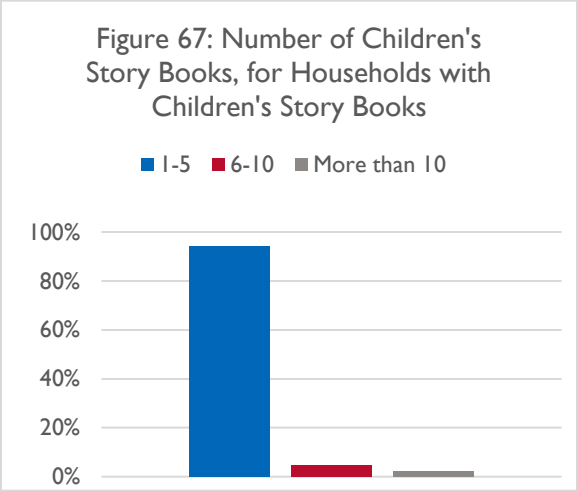
As noted above, religious texts are the primary reading material found in homes. The chart below depicts the types of reading materials in households that have reading materials (81% of the total sample). Only 24% of the 81% of households that have reading materials have children's storybooks. Of those, 94% only have 1-5 children's storybooks.

[67] This is defined as the perceived household decision-maker.

FIGURE 66: TYPE OF READING MATERIALS, FOR HOUSEHOLDS WITH MATERIALS

FIGURE 67: NUMBER OF CHILDREN'S STORY BOOKS, FOR HOUSEHOLDS WITH CHILDREN'S STORY BOOKS





RESULTS OF REGRESSIONS: STUDENT AND HOUSEHOLD FACTORS AND READING SCORES

In Table 11 below, we show the results of regressions exploring the relationship between student and household factors and outcomes across four of our reading assessment sub-tasks.[68] Please note that the co-efficients across P2 syllables, P2 familiar words, and reading comprehension are comparable (all based on percentage correct) in terms of interpreting the relative effect sizes. However, the co-efficients for ORF should be interpreted differently in terms of their effect relative to the other three sub-tasks since ORF is based on additional correct words per minute, rather than percentage correct. The factors with significant findings are listed below, in general order of the magnitude of the findings:

- Whether the student's head of household reads is by far the most important factor based on the size of the co-efficients and the significance level (1% across all sub-tasks).
- Whether a child has repeated a grade also has fairly high co-efficients across sub-tasks (all significant at the 1% level except one).
- Being in an urban area has fairly high co-efficients but the significance levels vary (this could be a function of sample size for urban schools).
- Across three of the four sub-tasks, being a male student is associated with lower reading scores, though the level of significance varies across sub-tasks.
- Students who reported being late at least once in the previous week were also more likely to have lower scores across most sub-tasks, though with varying levels of significance.

[68] Standard errors clustered for all regressions shown. Given the high scores on letter naming and therefore lack of variation, letter naming was excluded from the analysis.

- Lastly, across of the sub-tasks, being over-age is associated with lower scores. This is consistent with other studies that indicate over-age children perform worse on average.[69]

TABLE 11: STUDENT AND HOUSEHOLD FACTORS AND READING SCORES, REGRESSION RESULTS

Variable	P2 Syllables		P2 Familiar Words		ORF		Reading Comp.	
		Cluster Fixed Effects		Cluster Fixed Effects		Cluster Fixed Effects		Cluster Fixed Effects
Age of child	-1.792*	-2.102**	-2.052**	-2.434***	-0.667**	-0.753**	-1.546**	-1.975***
	(0.907)	(0.870)	(0.909)	(0.910)	(0.301)	(0.304)	(0.766)	(0.731)
Child is boy	-7.369***	-6.973***	-5.244**	-4.731*	-3.387***	-3.287***	-1.385	-0.949
	(2.356)	(2.384)	(2.344)	(2.382)	(0.862)	(0.876)	(1.929)	(1.962)
Urban location	8.317*		10.96**		2.576		9.373**	
	(4.534)		(4.430)		(1.737)		(4.089)	
Wealth Index (standardized)	2.261	0.301	2.278*	-0.0666	0.867	0.00455	-0.144	-1.340
	(1.415)	(1.194)	(1.342)	(1.212)	(0.526)	(0.445)	(1.228)	(1.080)
Household head reads	12.49***	10.33***	11.53***	8.957***	4.084***	3.397***	11.76***	10.52***
	(2.814)	(2.815)	(3.118)	(3.029)	(1.072)	(1.070)	(2.830)	(2.730)
Child reports repeating a grade	-8.393***	-8.054***	-6.791**	-7.545***	-2.924***	-2.558***	-7.677***	-7.293***
	(2.669)	(2.438)	(2.719)	(2.478)	(0.989)	(0.862)	(2.328)	(2.181)
Child reports being late at least once the previous week	-2.758***	-1.734*	-2.782**	-1.768*	-0.858***	-0.443	-2.062**	-1.375
	(0.936)	(0.888)	(1.064)	(1.043)	(0.315)	(0.300)	(0.796)	(0.831)

[69] Over-age students typically have either started school late or repeated grades. Given the high percentage of students reporting grade repetition, we can assume that is a substantial driver of over-age students in P2.

We also explored what student and household factors were associated with the probability of having a zero score across four sub-tasks. We again find that whether or not the head of household reads is the most strongly and consistently associated with the probability of scoring zero across all subtasks, significant at the 1% level. Being a student in a rural location also increases the likelihood of scoring zero across all sub-tasks, in keeping with the finding that zero scores exist predominantly in rural areas earlier in this report.

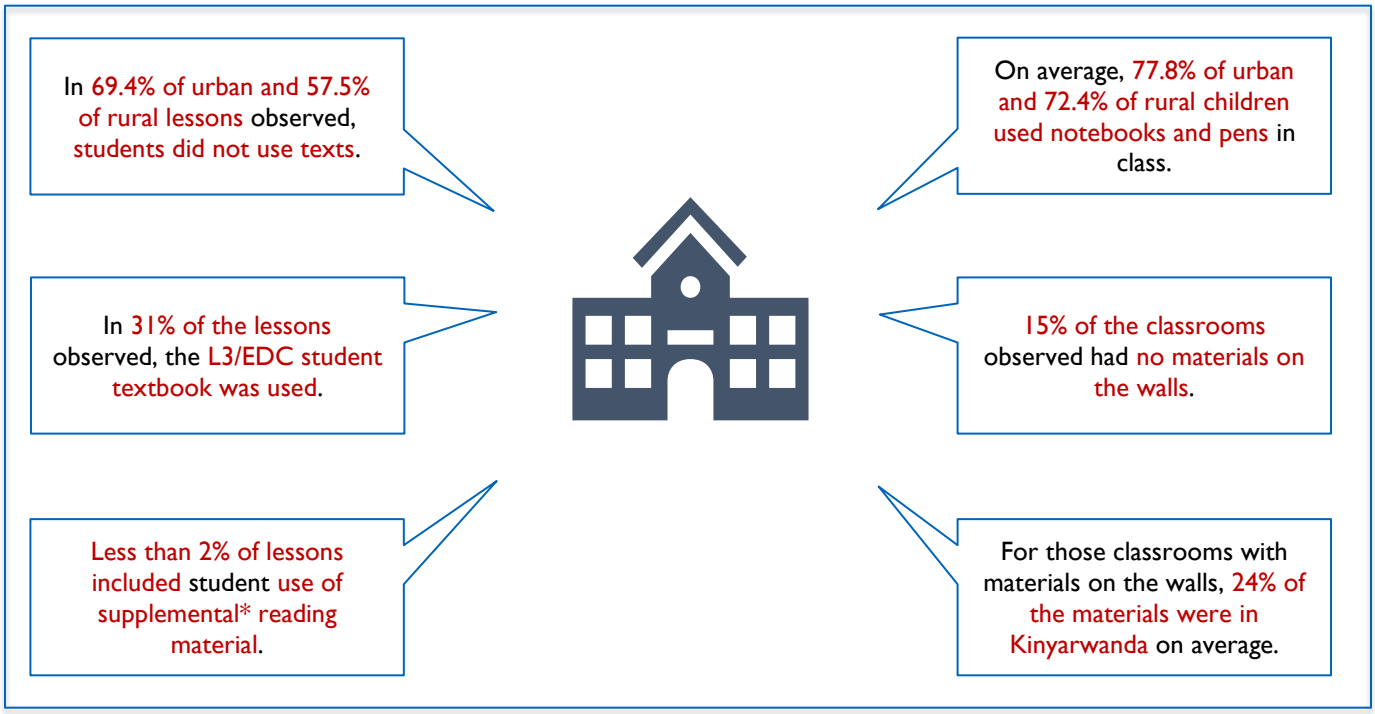
TABLE 12: STUDENT AND HOUSEHOLD FACTORS AND PROBABILITY OF ZERO SCORE

Probability of Zero Score	P2 Syllable		P2 words		ORF		Reading Comp	
Variable		Cluster Fixed Effects		Cluster Fixed Effects		Cluster Fixed Effects		Cluster Fixed Effects
Age of child	0.0202*	0.0214**	0.0260**	0.0338***	0.0267**	0.0331***	0.0271***	0.0318***
	(0.0105)	(0.00995)	(0.00998)	(0.00962)	(0.0103)	(0.00981)	(0.00993)	(0.00914)
Child is boy	0.0480*	0.0467*	0.0460*	0.0418	0.0532**	0.0495*	0.0526**	0.0495*
	(0.0267)	(0.0268)	(0.0259)	(0.0261)	(0.0258)	(0.0260)	(0.0253)	(0.0257)
Urban location	-0.0889**		-0.108**		-0.112**		-0.112**	
	(0.0379)		(0.0419)		(0.0429)		(0.0479)	
Wealth Index (standardized)	-0.0194	-0.000636	-0.0217	0.00150	-0.0269*	-0.00295	-0.0206	0.00361
	(0.0142)	(0.0141)	(0.0146)	(0.0136)	(0.0146)	(0.0135)	(0.0159)	(0.0141)
Household head reads	-0.109***	-0.0842***	-0.109***	-0.0824**	-0.110***	-0.0869**	-0.141***	-0.120***
	(0.0285)	(0.0301)	(0.0353)	(0.0351)	(0.0330)	(0.0334)	(0.0369)	(0.0363)

FINDINGS: CLASSROOM PROFILE

In addition to collecting data on how time was spent in the classroom, we also collected information about the general classroom climate and material use during lessons. The figure below provides highlights of materials utilization by students as well as print-materials in the classroom.[70] Notably, in 69.4% of urban and 57.5% of rural lessons observed, students did not use texts. The urban/rural proportions are unexpected and certainly the overall use of texts by students during reading instruction is quite low. The most prevalent student textbook in use was the L3/EDC textbook. A surprisingly low percentage (24%) of classroom wall materials were in Kinyarwanda.

FIGURE 68: CLASSROOM AND STUDENT MATERIALS

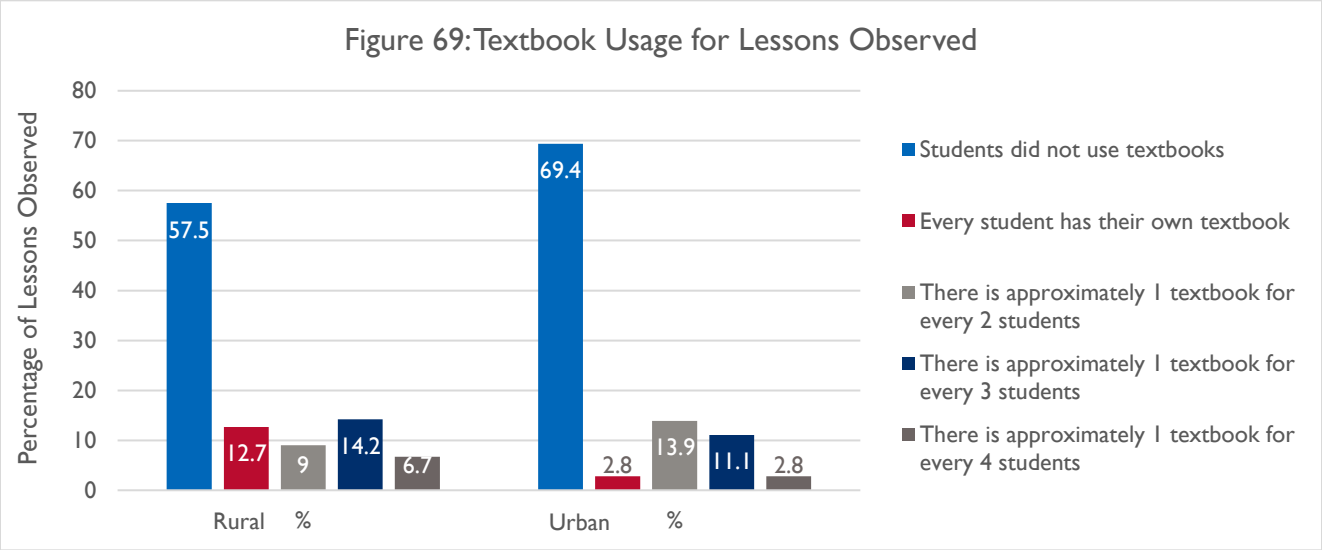


*Reading materials other than student textbooks such as storybooks.

Recognizing that access to reading materials is critical for the delivery of reading instruction, we estimated the student : textbook ratio for observed classes and found that when textbooks were used by students, the most prevalent student : textbook ratio was 3:1, with 2:1 a close second. Urban schools had fewer textbooks in use by students and when the urban students did use textbooks, their student : textbook ratio was lower than rural schools overall.

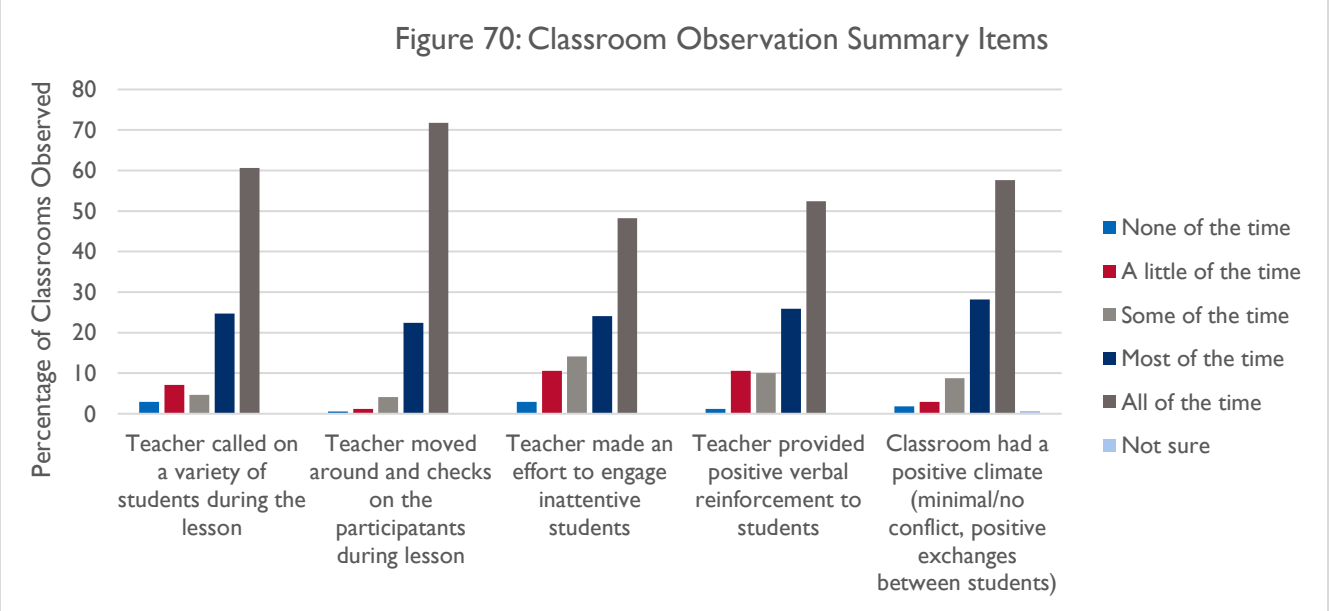
[70] For data in the classroom profile section, mean scores from the first observation (by lead observer) is used in analysis. Data from the second observation of the same classrooms is not included in the analysis.

FIGURE 69: TEXTBOOK USAGE FOR LESSONS OBSERVED



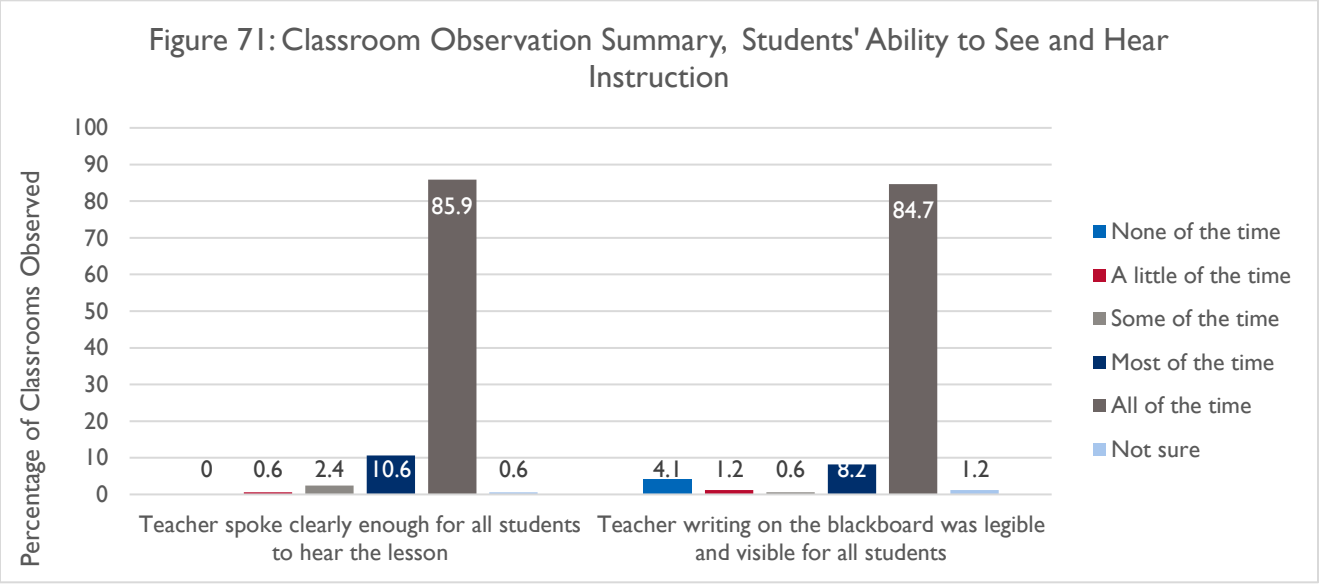
Beyond materials, observers also recorded summary information about the lesson, teachers' engagement of students, and the classroom environment to capture a more qualitative understanding of the observed classrooms. As outlined in Figure 70 below, a majority of teachers make consistent efforts across the observation items all of the time or most of the time. That said, there is room for improvement particularly with respect to student engagement strategies given what an important contributor this is to the to the amount and quality of instructional time. 27.6% of teachers are partially or not at all engaging inattentive students, and 14.7% of teachers are partially or not at all calling on a variety of students.

FIGURE 70: CLASSROOM OBSERVATION SUMMARY ITEMS



A consistent concern about classrooms in developing countries is the ability for students to hear the teacher (due to large class sizes, open walls/windows, etc.) and see what the teacher is writing on the board (due to low contrast on blackboards, large classrooms, low lighting) so we collected data about these aspects of the classroom. We find that in the roughly 85% of lessons observed, students could both hear the lesson and see the blackboard all of the time.

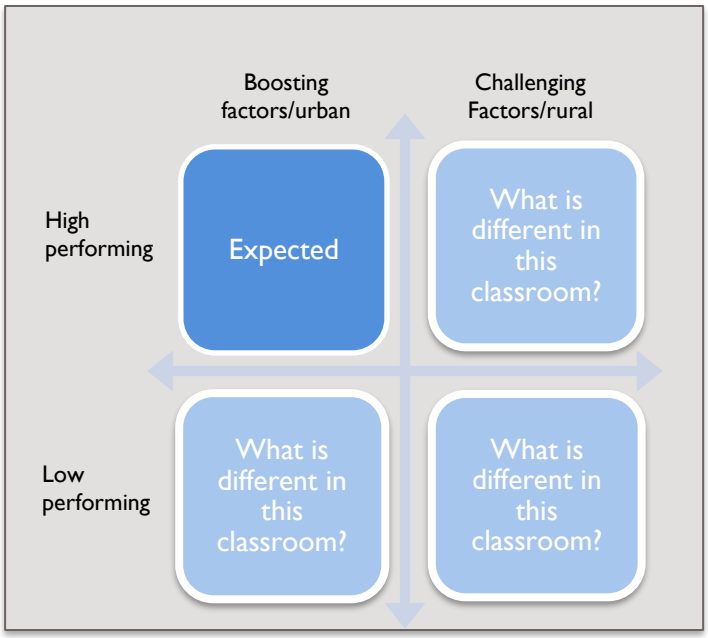
FIGURE 71: CLASSROOM OBSERVATION SUMMARY, STUDENTS' ABILITY TO SEE AND HEAR INSTRUCTION



FINDINGS: CASE STUDY SCHOOLS– COMMON TRENDS IN TIME, STUDENT PROFILE, TEACHER CHARACTERISTICS.

Analysis of the 2016 Learning Assessment of Rwandan Students (LARS) data tell us that there is a range of school performance in terms of learning outcomes and that there are schools in Rwanda performing well, even in rural areas. There are also schools in urban areas that have lower learning outcomes than some schools in rural areas. To deepen our analysis of how different factors, especially time, translate to reading outcomes, we created case studies to examine schools with differing levels of performance.

FIGURE 72: CASE STUDY APPROACH

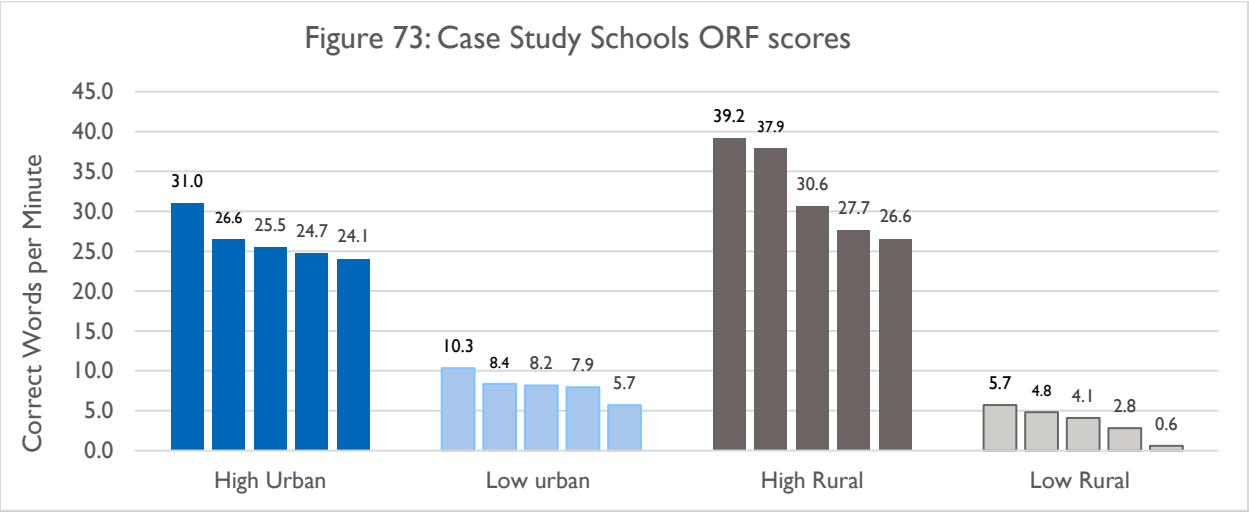


Originally, we allocated 15 schools out of the 100-school sample to create 3 case studies of 5 schools each. The three case studies were 1) high-performing schools in urban areas (expected), (2) high-performing school in rural areas (positive deviance), and (3) low-performing schools in urban areas (unexpected). We added a fourth case study of 5 low-performing schools in rural areas following data collection, totally 20 schools included in our case studies.

In the original design, these schools were selected using LARS reading scores, with the highest/lowest performing school per province (urban/rural) selected for each of the 3 case studies. Following data collection, we reviewed the selected case study schools using reading assessment data from this study and found that the cluster of schools identified for each case study did not effectively represent the highest/lowest performing schools in our sample. The primary reason for this was that the initial selection of case study schools was based on highest/lowest performing school (per LARS data) per province. Given the diversity in average school performance across provinces in Rwanda, selecting the highest urban/rural and lowest urban *per province*, resulted in a wide spread of performance within each case study. For example, using LARS reading performance data to select the top rural school per province resulted in a spread of 38 cwpm to 12 cwpm in oral reading fluency and a range from 2nd to 74th in rank within the overall study sample.

To address this issue, utilizing the reading assessment results from the study, we re-selected the case study schools based on oral reading fluency rank within the full sample (rather than selecting highest/lowest per province). This selection process resulted in a much stronger clustering of schools into case studies to allow us to explore why these clusters have such different student reading outcomes. Figure 73 below shows the average oral reading fluency for each school and set of schools in a case study.

FIGURE 73: CASE STUDY SCHOOLS ORF SCORES



The following table shows mean scores for each case study cluster across all reading skills, illustrating that the designation of case study schools holds across sub-tasks scores.

TABLE 13: READING SCORES FOR CASE STUDY SCHOOLS

Table 13: Reading Scores for Case Study Schools	Full Random Sample	High-Performing Urban	High-Performing Rural	Low-Performing Urban	Low-Performing Rural
Mean Percentage Correct					
Letter Naming	84.5	98.3	98.9	83.9	65.0
P1 Syllable Sounding	67.0	93.4	95.2	59.9	26.3
P2 Syllable Sounding	49.4	80.4	86.5	32.0	14.4
P1 Familiar Word Reading	60.0	92.1	93.4	47.3	18.4
P2 Familiar Word Reading	59.2	89.5	92.0	45.6	17.6
P3 Familiar Word Reading	55.4	88.6	92.5	39.4	15.3
Mean Correct Words per Minute					
P1 Oral Reading Fluency	15.6	26.4	32.4	8.1	3.6
Mean Score on Comprehension Questions					
P1 Comprehension	2.0/5	3.0/5	3.3/5	1.4/5	0.7/5

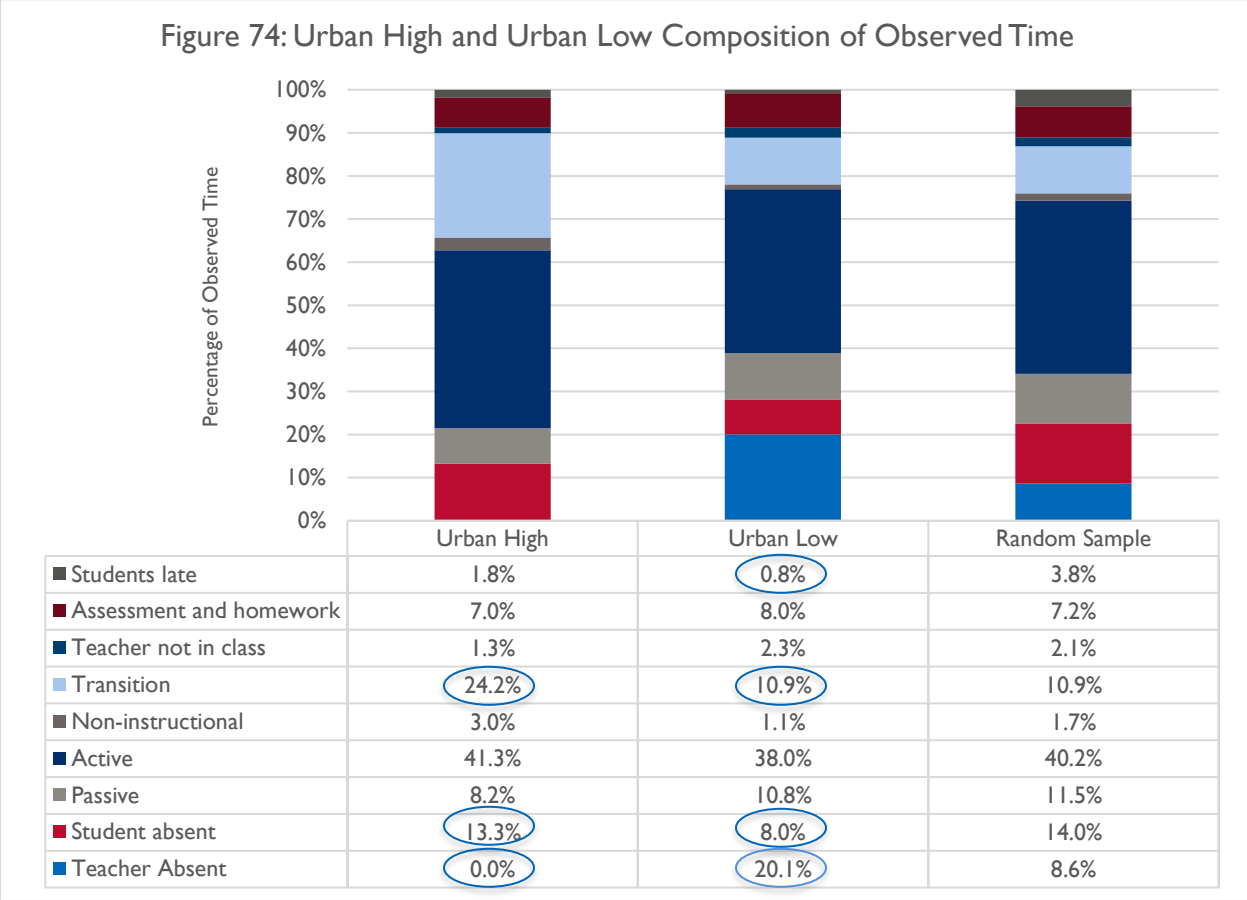
COMPOSITION OF TIME, CASE STUDY SCHOOLS

For each case study cluster, we assessed time using the same approach and calculations as for the random sample. We have presented the results with a focus on identifying variations across case study

clusters, especially between Urban High and Urban Low and Rural High and Rural Low, following the case study objective of identifying what factors drive school performance within similar contexts.

First, we look at the composition of total time between Urban High and Urban Low schools. We find data that both support and run counter to theories about characteristics of high-performing schools. Unexpectedly, students in Urban Low performing schools have substantially less time lost due to tardiness. This might be expected in even low-performing urban schools as compared to the random sample average, given the high proportion of rural schools in the overall sample where children typically have to travel greater distances on foot to get to school. We also see fewer student absences in Urban Low performing schools than in Urban High performing schools, which again, runs counter to theories about the relationship between time and learning outcomes. However, the substantial amount of time lost in Urban Low performing schools due to teacher absence is notable and supports the hypothesized relationship between time and learning outcomes.

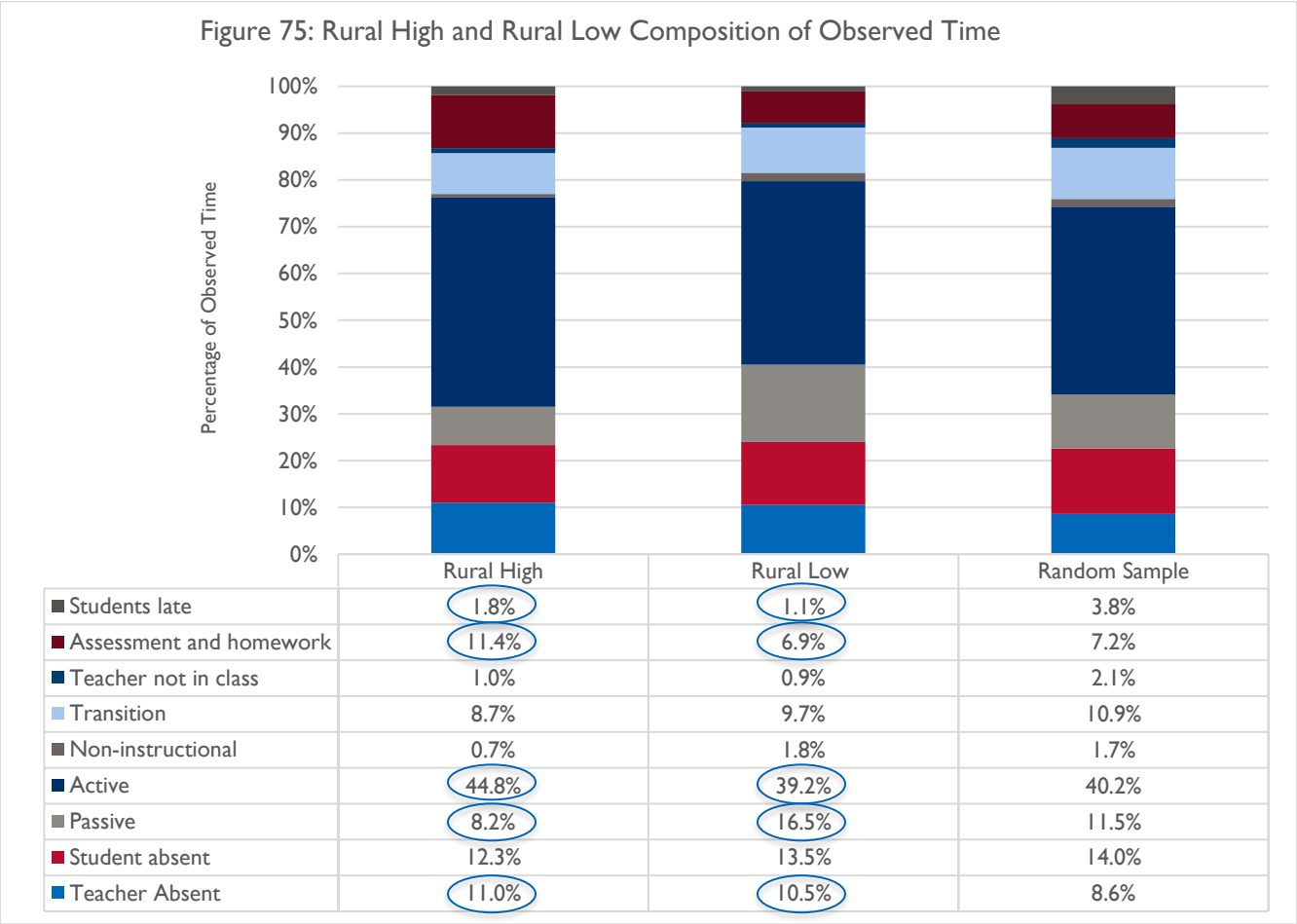
FIGURE 74: URBAN HIGH AND URBAN LOW COMPOSITION OF OBSERVED TIME



	Urban High	Urban Low	Random Sample
Students late	1.8%	0.8%	3.8%
Assessment and homework	7.0%	8.0%	7.2%
Teacher not in class	1.3%	2.3%	2.1%
Transition	24.2%	10.9%	10.9%
Non-instructional	3.0%	1.1%	1.7%
Active	41.3%	38.0%	40.2%
Passive	8.2%	10.8%	11.5%
Student absent	13.3%	8.0%	14.0%
Teacher Absent	0.0%	20.1%	8.6%

The difference in how in-classroom time is spent across Urban High and Urban Low schools is instructive and could be a contributing factor in the gap between reading outcomes. There is a marked difference in the amount of time spent in transition between activities, with Urban High schools spending considerably more time on transition activities. One would expect to see less transition time and more active or passive learning time in higher-performing schools, but that is not what these data reveal. There is additional discussion about the composition of in-classroom time across case study clusters in the following sections.

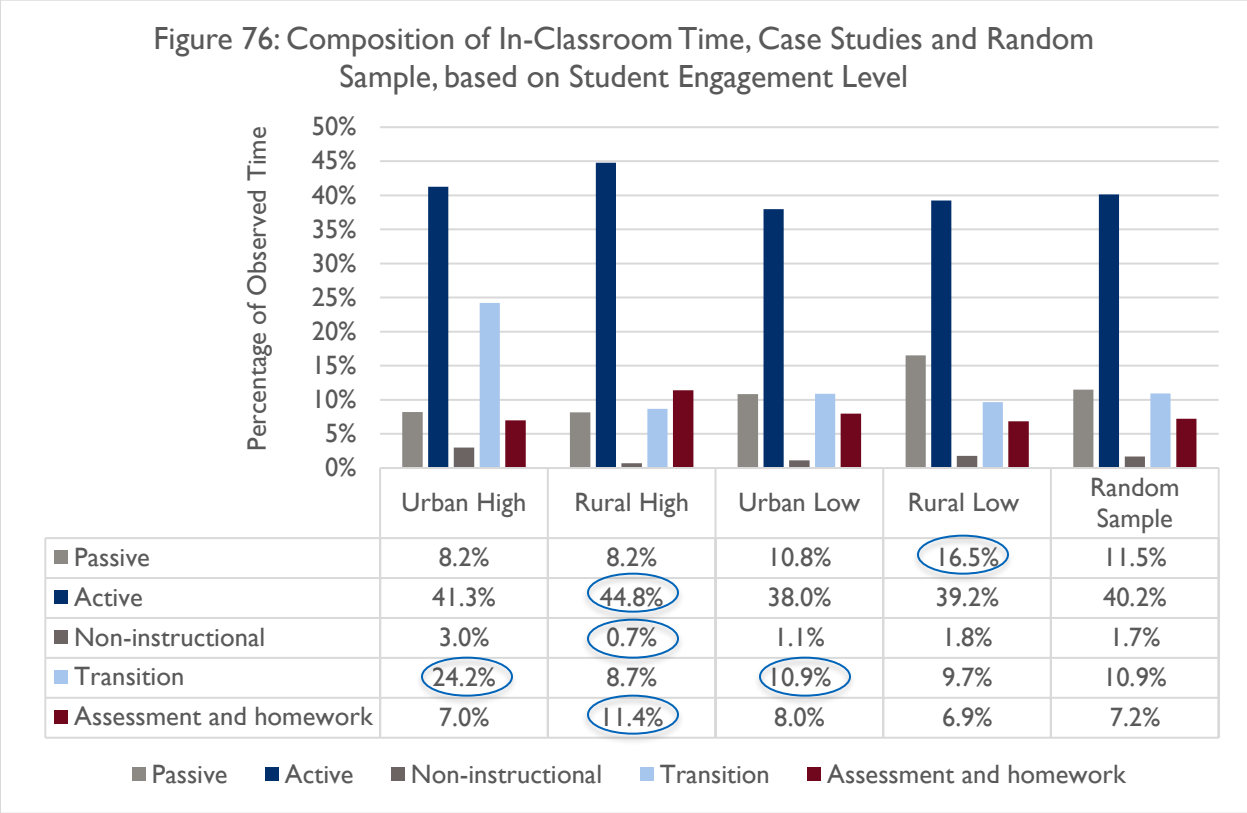
FIGURE 75: RURAL HIGH AND RURAL LOW COMPOSITION OF OBSERVED TIME



	Rural High	Rural Low	Random Sample
Students late	1.8%	1.1%	3.8%
Assessment and homework	11.4%	6.9%	7.2%
Teacher not in class	1.0%	0.9%	2.1%
Transition	8.7%	9.7%	10.9%
Non-instructional	0.7%	1.8%	1.7%
Active	44.8%	39.2%	40.2%
Passive	8.2%	16.5%	11.5%
Student absent	12.3%	13.5%	14.0%
Teacher Absent	11.0%	10.5%	8.6%

Comparing the composition of total observed time for Rural High and Rural Low performing schools, we see several important differences. The Rural High performing schools spend considerably more time on assessment and homework activities. Deeper analysis on the specific nature of these activities will further inform how this time might be contributing to the gap in reading outcomes. The ratio of active : passive reading instructional activities is also substantial between Rural High and Rural Low performing schools, with Rural High performing schools spending around 50% less time on passive reading activities and about 14% more time on active reading activities. Interestingly, and not expectedly, the teacher absent/late/not in-classroom rate is roughly the same between the two case studies, as are the student absent/late rates.

FIGURE 76: COMPOSITION OF IN-CLASSROOM TIME, CASE STUDIES AND RANDOM SAMPLE, BASED ON STUDENT ENGAGEMENT LEVEL



	Urban High	Rural High	Urban Low	Rural Low	Random Sample
Passive	8.2%	8.2%	10.8%	16.5%	11.5%
Active	41.3%	44.8%	38.0%	39.2%	40.2%
Non-instructional	3.0%	0.7%	1.1%	1.8%	1.7%
Transition	24.2%	8.7%	10.9%	9.7%	10.9%
Assessment and homework	7.0%	11.4%	8.0%	6.9%	7.2%

There are some notable differences in how in-classroom time is spent across these case studies that might offer some explanation for the gap reading outcomes. First, we see that the proportion of passive time in Rural Low Performing schools is substantially higher than in Rural High Performing schools. This could be a function of pedagogical approach, teacher orientation, materials to support more active instructional practice, etc. This hypothesis aligns with the proportion of active learning time in Rural High performing schools, which outpaces all other case study groups, including the Urban High performing schools. One explanation could be that more active pedagogies and classroom management practices compensate for student and other factors that make achieving reading outcomes in rural schools more challenging. We also see Rural High Performing schools spending more time on assessment and homework-related activities.

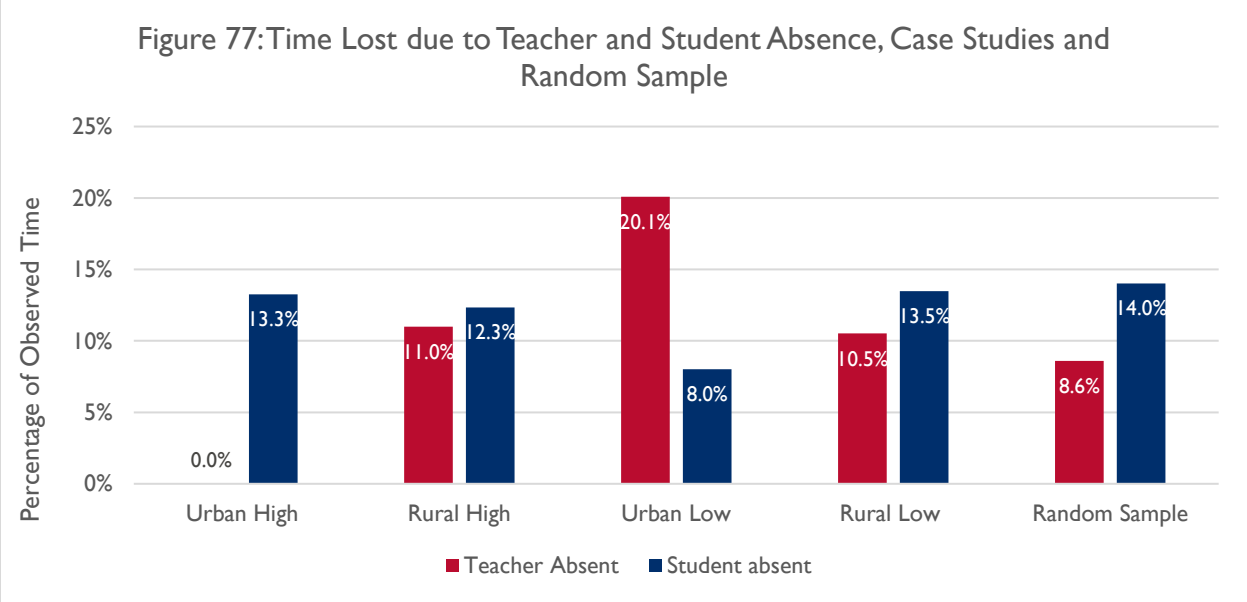
Interestingly, Urban High performing schools have a substantially higher proportion of in-classroom time dedicated to both non-instructional and transition activities. One underlying hypothesis of this study is that transition time activities, while necessary, do not contribute as strongly toward reading outcomes as passive or active reading instruction activities. The 24.2% of in-classroom time spent on transition time in Urban High performing schools challenges this assumption, but the low 8.7% of in-classroom time spend on transition time in Rural High performing schools provides support for this assumption. And of course, none of these factors are operating in isolation, with the overall balance of time allocation, teacher characteristics, student characteristics, school administration leadership and materials all intersect with time to drive outcomes. An increase in transition activities can also explain a corresponding higher proportion on non-instructional activities given the tendency for students to

disengage or be off-task during/around transition activities. The nature of the transition activities across schools can be explored in more detail using this study's data set.

TEACHER AND STUDENT ABSENCE/TARDINESS, CASE STUDIES

The graph below shows the percentage of time lost due to both teacher and student absence for case study schools and the random sample. It is important to recall that when teacher absence occurs, it contributes greatly to the overall loss in time given that all students enrolled in the class lose that learning time. When students are absent, they only lose their own personal learning time. Nonetheless, we still see student absence totaling a higher percentage of lost time than teacher absence in three of the four case study clusters. Only in Urban Low performing schools does the lost time due to teacher absence outstrip the other case study clusters and the average for random sample.

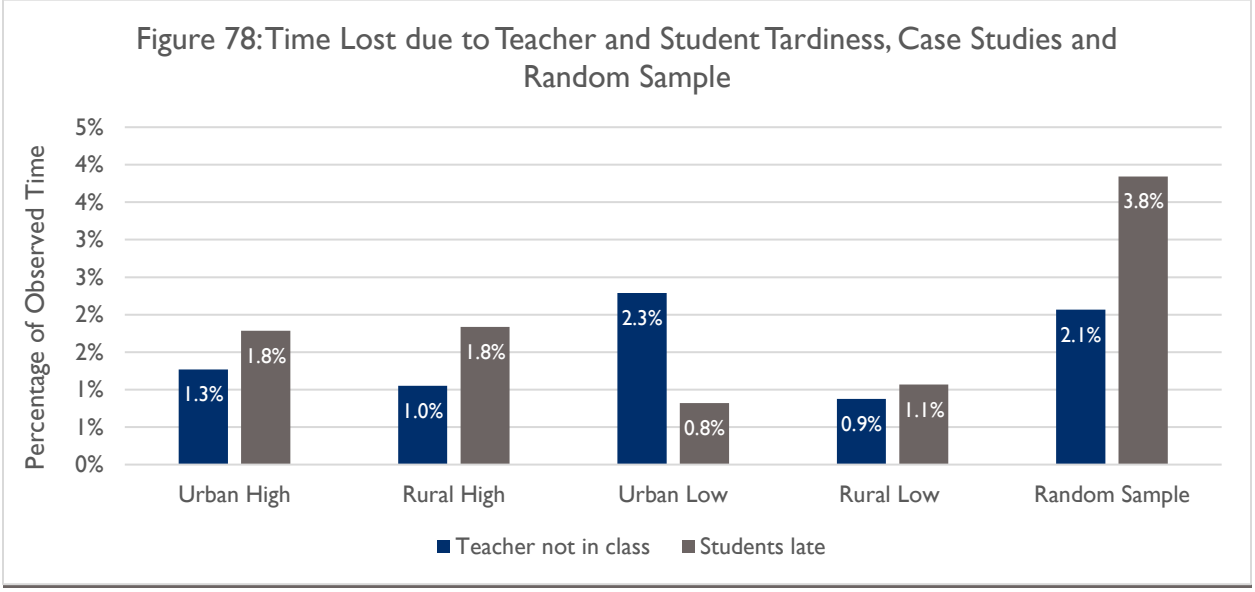
FIGURE 77: TIME LOST DUE TO TEACHER AND STUDENT ABSENCE, CASE STUDIES AND RANDOM SAMPLE



As with teacher absence, when teachers are late or not present in the classroom the effect on in-classroom time is multiplied by the number of students present. As such, teacher tardiness or time outside of the classroom during lessons[71] should be looked at carefully.

[71] It is important to recall that if the teacher has assigned an assessment or other activity for students and is outside of the classroom while the activity is underway, that time will be counted toward learning time, rather than teacher not present.

FIGURE 78: TIME LOST DUE TO TEACHER AND STUDENT TARDINESS, CASE STUDIES AND RANDOM SAMPLE



Following this discussion about variations in time across case study clusters, we turn to the characteristics of the schools, teachers, students and students' households to delve more deeply into what factors might be driving these differences. Are the most important factors at the classroom level or do other factors dominate?

Across case study clusters we see variation in teacher and student characteristics, some of which might offer insights into why there are such substantial differences in performance between the urban high and urban low schools and the rural high and rural low schools. With respect to teachers, we see a substantial difference in average teacher age between the Rural Low performing schools, all other case study schools and the mean for the full random sample. There is a corresponding difference in the number of years teaching, with Rural Low performing schools having an experience deficit ranging from 9-13 years as compared to the other case study schools. We also see a much lower per capita teacher absence rate in High Performing Urban schools as compared to the average for the random sample and all other case studies.

TABLE 14: TEACHER PROFILE IN CASE STUDY SCHOOLS

Table 14: Teacher Profile in Case Study Schools							
	Urban High	Urban Low		Rural High	Rural Low		Mean for Random Sample
% Female	100%	100%		60%	80%		71%
Age	44	44		41	31		40
% of teachers' whose highest degree specializes in education	100%	80%		100%	80%		84%

Table 14: Teacher Profile in Case Study Schools							
	Urban High	Urban Low		Rural High	Rural Low		Mean for Random Sample
Number of years teaching	20	16		17	7		15.4
Years teaching P1-P3 Kinyarwanda	14	8		7	5		8.5
Years teaching P1-P3 Kinyarwanda in current school	13	6		5	4		6.3
Time in minutes: round trip home to school	71	52		62	42		66
Hours spent/week: administrative tasks	8	3		2	6		4
Hours spent/week: P2 Kinyarwanda lesson planning	3	3		2	2		3
Total P2 Per Capita Teacher Absence Rate for school[72]	3.6	7.0		8.9	7.9		7.4

We ran t-tests on the teacher factors listed in the table above to identify if the differences were statistically significant. It is important to recognize that the co-efficient value (the magnitude of the difference) has to be very large to be detected by our 5-teacher sample for each case study.

We found the following statistically significant differences between Urban High and Urban Low:

- The Urban Low teacher absence is rate is 3.4% (***) higher than the Urban High teacher absence rate. and it is significant at 1% level.
- Urban High teachers had 7.4 (*) more years of experience than Urban Low teachers.

STUDENT AND HOUSEHOLD PROFILE IN CASE STUDY SCHOOLS

We also see marked differences in the student profile across case study clusters as one might expect (outlined in table below). There are consistent indicators that students in Urban High performing schools come from households with higher parental literacy and have access to more resources than students across the other case study clusters. Across all the variables, students attending Rural Low performing schools have access to fewer resources and receive less support for schooling. Recalling that the reading assessment scores between Urban High and Rural High schools are roughly equivalent, we look to understand how different household factors are between these two groups. Notably, reported

[72] Absences for all P2 teachers at the school was collected from front office administrative records. The rate is calculated by taking the total number of P2 Kinyarwanda teacher absent days, divided by the total number of P2 Kinyarwanda days and the number of scheduled school days per month according to the official school calendar.

head of household literacy is the same across Urban High and Rural High schools, as is support for reading at home. Students from Rural High performing schools actually outpace students from Urban High performing schools in terms of reading materials at home and habit of reading outside of school. Yet, we have indicators that the resources available between these two groups is strikingly different, with only 55% of students in Rural High case study schools having electricity at home as compared to 83% of Urban High students (this is also a function of area infrastructure and not exclusively attributable to socio-economic status).

TABLE 15: STUDENT PROFILE FOR CASE STUDY SCHOOLS

Table 15: Student Profile for Case Study Schools					
	High Urban	Low Urban	High Rural	Low Rural	Mean for Random Sample
Father head of household	73%	61%	68%	81%	76%
Has seen head of household read	88%	74%	88%	70%	81%
Has mobile phone at home	91%	89%	94%	78%	85%
Has electricity at home	83%	75%	55%	24%	53%
Has reading materials in home	80%	73%	90%	73%	81%
Reads outside of school often/sometimes	53%	50%	76%	28%	53%
Reads with someone at home	83%	80%	81%	78%	81%
Sometimes or always gets helps with homework	91%	93%	84%	88%	88%
Has three meals or more/day	84%	86%	85%	70%	80%

We ran t-tests[73] on the student and household factors listed in the table above to identify if the differences were statistically significant. We found the following statistically significant differences between Urban High and Urban Low:

- 13.7% (***) more heads of household in Urban High schools read than in Urban Low schools.

[73] A t-test is implemented to determine if the difference between two groups' or variables' means (averages) are statistically significant e.g., not likely different because of chance).

- 40% of students in Urban High schools reported not being late to school at all in the past week, as compared to 20% (***) of students in Urban Low schools.

We found the following statistically significant differences between Rural High and Rural Low:

- Students in Rural High schools score .66% of a standard deviation higher (***) on the wealth index[74] than Rural Low students.
- Students in Rural High schools have three meals a day 15% (**) more than student in Rural Low schools.
- The absence rate for Rural Low students is 4.22% (***) higher than Rural High students.
- 39% of students in Rural High schools reported not being late to school at all in the past week, as compared to 20% of students (**) in Urban Low schools.
- Students in Rural High schools read books outside of school 48.7% (***) more than students in Rural Low schools.

SECTION VII: REMAINING RESEARCH QUESTIONS

In this study, we were able to address a number of questions necessary to inform policy and practice with the aim of improving early grade reading outcomes in Rwanda. However, like most studies, we were not able to collect data on every aspect we may have desired to and in some cases the data we have collected could benefit from further exploration. In addition, research studies tend to identify new research questions, and this study is no exception. We have listed some key research recommendations below:

- Complement the findings of this study with in-depth qualitative studies to further investigate the reasons behind high teacher and student absenteeism and identify mechanisms for change.
- Investigate what the current school administrative response is to teacher and student absence in schools where teacher and student absence is high.
- Investigate what enabling contextual factors or accountability/incentive mechanisms are in place in schools with low teacher and student absence, especially the high-performing rural case study schools.
- Review the literature about effective practices to reduce teacher and student absence and tardiness.
- Investigate the extent to which pre-service and in-service training as well as the current curriculum and learning materials for lower primary Kinyarwanda instruction supports "active" instructional methods (which are associated with higher student reading scores).
- Investigate the extent to which teachers are able to fully deliver the early grade Kinyarwanda curricula given time available, recognizing that the scope and sequence and teaching and learning materials have been updated since the time of the study data collection.
- Further refine measurement of teacher motivation that are relevant to the Rwandan context and can measure changes over time.
- Further investigate the finding that students taught by teachers who only teach Kinyrawanda have lower reading assessment scores (13% of teachers in the sample only teach Kinyarwanda). If this relationship is further verified, explore the characteristics of these teachers that might explain the phenonmenon.

[74] See Annex D for description of wealth index calculation.

- Investigate why 69.4% of urban and 57.5% of rural lessons observed did not use any textbooks and conduct follow-up studies to ensure the textbooks being delivered by Soma Umenye are in classrooms and being utilized.
- Conduct predictive modeling to understand the level at which increases in instructional time are estimated to improve student scores across all reading assessment sub-tasks

SECTION VIII: RECOMMENDATIONS: IMPLICATIONS FOR POLICY & PRACTICE IN RWANDA

This study has identified that nearly 30% of the instructional time allocated for P2 Kinyarwanda classes is lost due to teacher/student absence and tardiness. An additional 19.3% of allocated instructional time is focused on activities other than direct reading instruction. The result is that approximately 50% of the intended instructional time, and all the associated costs and resources, are not leveraged for learning. The scope of this research, and the attendant recommendations, reflect that improving early grade reading outcomes at scale and with sustainability is a complex task and requires a range of actors working together toward the same goal. The recommendations below are derived from the findings in the report, are organized by the different actors or parts of the education eco-system that influence an array of different mechanisms, policies and practices.

NATIONAL GOVERNMENT (MINEDUC/REB)

- The analysis of average reading scores with and without zero scores in this study identified that urban and rural school perform similarly, on average, when zero scores are excluded. This indicates that there are opportunities to learn in rural Rwandan schools, but some children are simply left behind – the "internally excluded." Strategies to identify and support these children should be developed and implemented as a priority. Effective remediation of zero scores can help reduce repetition and drop-out rates in addition to ensuring all children have the opportunity to learn. Significantly reducing the substantial proportion of zero scores will also be the most efficient way for Rwanda to increase the overall proportion of students meeting reading benchmarks.
- Investigate reasons for teacher/student absence/tardiness and establish remediation of absence/tardiness as a priority for basic education.
- While the broader issues are being addressed, based on "by day" and "by hour" absence and tardiness data, revise scheduling of Kinyarwanda and other foundational classes so a higher percentage of classes are conducted on days and at times with lower absence and tardiness rates. This analysis should recognize differences between urban and rural schools and ensure that all foundational subjects receive sufficient instructional time.
- Rwanda benefits from a strong de-centralized political structure with planning and accountability mechanisms in place. We encourage discussion about the findings of this report with local government to develop strategies for remediation and integration of improvement goals into perofrmance contracts with realistic progressive targets for reducing absenteeism and tardiness.[75]

[75]For additional information on imihigo/performance contracts, see http://documents.worldbank.org/curated/en/833041539871513644/122290272_20181348045807/additional/131020-WP-P163620-WorldBankGlobalReport-PUBLIC.pdf

- School inspection and quality education campaigns should focus on the issues of teacher and student absenteeism and tardiness, communicating how these factors negatively affect learning outcomes and squander valuable resources.
- Develop information systems that provide timely and reliable data about teacher and student absence/tardiness, complemented by joint reviews at the national and local levels.
- Consider more effective response systems for when teachers are unable to attend school so that lessons can be delivered despite teachers absence, some of which may be out of the control of the school administation and teachers.
- Teacher preparation institutions and in-service teacher professional development programs should train and equip (e.g., with materials) teachers to deliver more "active" instruction.
- Train and incentize those supporting teachers (e.g., head teachers, directors of studies, etc.) in reading instruction, active pedagogies and student engagement strategies.
- Implement strategies and investments to ensure children have sufficient access to materials aligned with the lesson. Only 30.6% of students in urban schools used textbooks during the observed lessons and only 42.5% did so in rural schools and less than 2% of students across the random sample used supplemental materials. This could be a function of availability or teacher's orientations/strategies for using texts.
- Develop practical and timely strategies for classroom-based assessments to close the gap between teachers' estimation of student reading skills and actual student reading skills. Teachers should also be made familiar with the evidence about the age at which children are able to learn reading and writing skills.
- Review of the school leadership policy to incorporate time on task focus in basic education.
- Develop clear standards for classroom observation frequency and goals by different roles (e.g., head teacher, director of studies) and include guidance about observation and feedback strategies to support more effective instructional time and student engagement.
- The analysis indicates that students have lower reading scores if they are taught by teachers who exclusively teach lower primary Kinyarwanda classes. This should be further explored and if less effective teachers are being slotted into teaching lower primary Kinyarwanda classes, local governments need to be engaged to reconsider this teacher placement approach.
- Given the investments in student textbooks, investigate why 69.4% of urban and 57.5% of rural lessons observed did not use any textbooks.
- Utilize simple technology such as cell phones to send SMS reminders to teachers about the importance of instructional time.

LOCAL GOVERNMENT

- Work with school heads to implement immediate changes to the times at which early grade Kinyarawanda lessons (and other foundational classes) take place to avoid those periods/hours with high absenteeism/tardiness (e.g., first period of AM and PM shifts) by teachers and students. This is a change that can be put into place immediately while schools and communities work with students and parents to reduce teacher and student absence and tardiness.
- Given local governments' central role in recruiting, placing and retaining teachers, convey and enforce expectations about school staff attendance during school hours.
- Implement information campaigns and other mechanisms to increase the number of children who have access to reading materials in the home.

SCHOOL LEADERSHIP

- Incentivize and motivate teachers to reduce absences and tardiness. This will require gaining a better understanding, for each school, the drivers of absenteeism and tardiness for teachers.
- Consistently remind teachers how important instructional time is to improve learning outcomes.
- In collaboration with REB and local authorities, implement changes to the time table to minimize the scheduling of lower primary Kinyarwanda lessons during lesson periods with high absenteeism/tardiness (e.g., first period of AM and PM shifts).
- Engage with school and community management bodies to idenify strategies to reduce student absenteeism and tardiness. Many of the reasons for absence or tardiness given by children need to be addressed by parents.
- Regularly coach teachers on the use of active instructional strategies. This is critical given the positive and significant relationship found between student reading scores and the proportion of time spent on active instruction in the classroom. Recalling that 24.7% of teachers report their head teacher had not observed one of their lessons in the past year, this recommendation will require incentives and accountability to ensure implementation.
- Regularly coach teachers on student engagement strategies (observations note that 27.6% of teachers try engaging inattentive students none of the time, a little of the time or some of the time).

TEACHERS

- Work with school and community leaders to address the challenges that are creating high teacher absenteeism rates.
- Consistently begin lessons on time to model expectations to students about reducing tardiness.
- Clearly communicate expectations about student on-time arrival to students and parents, with support from school leaders.
- Implement strategies to identify students with low learning levels and coordinate with school leadership and parents about how to support learners who have fallen behind.
- Reflect on instructional strategies and increase the proportion of time spent on active instructional activities across all reading skills. The data show clearly that more active instructional activities is associated with higher student engagement levels.
- Consistently utilize and follow the teachers guide and scripted lesson (note: the Soma Umenye program with scripted lessons is being rolled out for P2 Kinyarwanda classrooms in the 2019-2020 school year and was not available in P2 classrooms during this study.)
- Increase attempts to bring unengaged or distracted students into the learning process.

COMMUNITIES & PARENTS

- Prioritize school and getting to school on-time over household chores (52% of children who were late in the last week cite household chores as the primary reason they were late).
- Develop extra-curriculur community-based programs to support students not meeting reading benchmarks.
- Develop strategies to increase the number of children who have access to leveled, decodable reading materials in the home. Less than 20% of students stated there were children's storybooks in their house and 94% of those households only have 1-5 children's storybooks.

DEVELOPMENT PARTNERS

- Work with REB and school leadership to support supplemental instruction programs to support struggling readers to complement the efforts of teachers during regular instruction time.
- Link investments in school leadership and teacher training programs to the findings about the most effective types of time use and the need to improve student engagement levels in the classroom.
- Leverage opportunities during program design and implemenation to share evidence with parents and communities about high absenteeism and tardiness rates and how it is adversely affecting learning outcomes.
- Fund additional research to better understand the drivers of high absenteeism and tardiness and test strategies to increase in-classroom time.
- Invest in strategies to increase the availabilty of high-quality, appropriately leveled reading materials available to children outside of school.