The LEGO Foundation

LEGO Play Accelerator

Baseline Report

February 18, 2022

Prepared for
The LEGO Foundation Play Accelerator Research Initiative

Prepared by
Jonathan Stern, Rachel Jordan, and Matthew Jukes
RTI International[1]
3040 East Cornwallis Road
Research Triangle Park, NC 27709-2194

RTI International is an independent, nonprofit research institute dedicated to improving the human condition. Clients rely on us to answer questions that demand an objective and multidisciplinary approach—one that integrates expertise across the social and laboratory sciences, engineering, and international development. We believe in the promise of science, and we are inspired every day to deliver on that promise for the good of people, communities, and businesses around the world. For more information, visit www.rti.org.

With the support of the LEGO Foundation.

[1] RTI International is a registered trademark and a trade name of Research Triangle Institute.

RTI INTERNATIONAL

TABLE OF CONTENTS

ABBREVIATIONS I

EXECUTIVE SUMMARY II

1.1 Overview ii

1.2 Methodology, Instruments and Sample ii

1.3 Findings and Recommendations iv
Learning through Play in the Classroom iv
LtP Enabling Environment v
Perceptions of LtP vi
Student Holistic Outcomes vi

1. INTRODUCTION 1

1.1 Overview of Activity 1

1.2 Purpose and Objectives of Baseline 1

1.3 Evaluation Design, Instruments, and Sample 2
Evaluation Design 2
Instruments 2
Program Evaluation Designs and Samples 6
Training and Data Collection 9

2. FINDINGS 11

2.1 Partners in Play Program (P3) - Ghana 12
What is the baseline state of teachers' pedagogical practices in the classroom (including LtP) among Grade 1 teachers in treatment and control schools? 12
What are the baseline perceptions of LtP among school, district, and community-level stakeholders? 21
What are the enabling factors associated with implementing LtP pedagogical methods in classrooms?
What are the barriers associated with implementing LtP pedagogical methods in classrooms? 23
What are the baseline levels of student literacy, numeracy and SEL among Grade 1 students in treatment and control schools? 28

2.2 Tucheza Kujifunza (TuKu) - Kenya 36
What is the baseline state of teachers' pedagogical practices in the classroom (including LtP) among Grade 1 teachers in treatment and control schools? 36
What are the baseline perceptions of LtP among school, district, and community-level stakeholders? 45
What are the enabling factors associated with implementing LtP pedagogical methods in classrooms?
What are the barriers associated with implementing LtP pedagogical methods in classrooms? 46
What are the baseline levels of student literacy, numeracy, and SEL among Grade 1 students in treatment and control schools? 51

2.3 Scaling Learning through Play - Rwanda 58
What is the baseline state of teachers' pedagogical practices in the classroom (including LtP) among Grade 1 teachers in treatment and control schools? 58
What are the baseline perceptions of LtP among school, district, and community-level stakeholders? 65
What are the enabling factors associated with implementing LtP pedagogical methods in classrooms?
What are the barriers associated with implementing LtP pedagogical methods in classrooms? 67
What are the baseline levels of student literacy, numeracy, and SEL among Grade 1 students in treatment and control schools? 72

3. CONCLUSION 80

3.1 Measurement Considerations 80

3.2 Implications and Recommendations 80
Learning through Play in the Classroom 80
LtP Enabling Environment 82
Perceptions of LtP 82
Student Holistic Outcomes 82

ANNEX A – ADDITIONAL TABLES 84

A.1 Ghana Stakeholder Interviews 84

A.1.1 Teacher Interview - Ghana 84
A.1.2 Coach Interview – Ghana 86
A.1.3 Trainer Interview – Ghana 87
A.1.4 DEO Interview – Ghana 88
A.1.5 Caregiver Interview – Ghana 88
A.2 Kenya Stakeholder Interviews 90
A.2.1 Teacher Interview – Kenya 90
A.2.2 Trainer Interview – Kenya 92
A.2.3 DEO Interview – Kenya 93
A.2.4 Caregiver Interview – Kenya 94
A.3 Rwanda Stakeholder Interviews 95
A.3.1 Teacher Interview – Rwanda 95
A.3.2 Coach Interview – Rwanda 97
A.3.3 Trainer Interview – Rwanda 97
A.3.4 DEO Interview – Rwanda 99
A.3.5 Caregiver Interview – Rwanda 99
A.4 Ghana SEL Score Distributions 101
A.5 Kenya SEL Score Distributions 102
A.6 Rwanda SEL Score Distributions 104
ANNEX B – RELIABILITY ANALYSIS 106
B1. EGRA – Cronbach's Alpha 106
B2. EGMA – Cronbach's Alpha 106
B3. SEL 106
B3.1 Behavioral Observation- Alpha 106
B3.2 Relationship Scale- Alpha 106
B3.3 Conflict Perception: Item-rest and Alpha 106
B3.4 Confidence- Item-Rest Correlation and Alpha 107
B3.5 Learning Environment- Item-Rest Correlation and Alpha 108
B4. Classroom Observation – Factor Analysis 109
B4.1 Ghana – Factor Analysis 109
B4.2 Ghana – Factor Loading 109
B4.3 Kenya – Factor Analysis 110
B4.4 Kenya – Factor Loading 110
B4.5 Rwanda – Factor Analysis 111
B4.6 Rwanda – Factor Loading 112
ANNEX C – INSTRUMENTS 113
C.1 Attitudes Scale – Teachers 113
C.2 Attitudes Scale - Caregivers 116
C.3 Timed Classroom Observation 117
C.4 Post-Observation Checklist 119
ANNEX D – OVERVIEW OF PROGRAMS 121

LIST OF FIGURES

Figure 1. Generalized Theory of Change for LtP Interventions 2
Figure 2. Student Engagement During Lessons (Percentage of Observed Class Time), by Treatment Group in Ghana 13
Figure 3. Teacher Actions During Lessons (Percentage of Observed Class Time), by Treatment Group in Ghana 14
Figure 4. Student Activity During Lessons (Percentage of Observed Class Time), by Treatment Group in Ghana 15

Figure 5. Use of Materials During Lessons (Percentage of Observed Class Time), by Treatment Group in Ghana 16
Figure 6. Student Grouping During Lessons (Percentage of Observed Class Time), by Treatment Group in Ghana 17
Figure 7. Teachers' Self-Efficacy Using LtP, by Treatment Group 24
Figure 8. Caregiver: Do You Have Any Play Materials at Home? 25
Figure 9. Student Engagement During Lessons (Percentage of Observed Class Time), by Treatment Group in Kenya 37
Figure 10. Teacher Actions During Lessons (Percentage of Observed Class Time), by Treatment Group in Kenya 38
Figure 11. Student Activity During Lessons (Percentage of Observed Class Time), by Treatment Group in Kenya 39
Figure 12. Use of Materials During Lessons (Percentage of Observed Class Time), by Treatment Group in Kenya 40
Figure 13. Student Grouping During Lessons (Percentage of Observed Class Time), by Treatment Group in Kenya 41
Figure 14. Teachers' Self-Efficacy Using LtP, by Treatment Group 47
Figure 15. Caregiver: Do You Have Any Play Materials at Home? 48
Figure 16. Student Engagement During Lessons (Percentage of Observed Class Time), by Treatment Group in Rwanda 58
Figure 17. Teacher Actions During Lessons (Percentage of Observed Class Time), by Treatment Group in Rwanda 59
Figure 18. Student Activity During Lessons (Percentage of Observed Class Time), by Treatment Group in Rwanda 60
Figure 19. Use of Materials During Lessons (Percentage of Observed Class Time), by Treatment Group in Rwanda 60
Figure 20. Student Grouping During Lessons (Percentage of Observed Class Time), by Treatment Group in Rwanda 61
Figure 21. Teachers' Self-Efficacy Using LtP, by Treatment Group 68
Figure 22. Caregiver: Do You Have Any Play Materials at Home? 69

LIST OF TABLES

Table ES1. Target Interview Respondents, by Program (Country) iii
Table ES2. School Samples by Program iv
Table 1. Target Interview Respondents, by Program (Country) 5
Table 2. Ghana School Sample 6
Table 3. Respondent Counts by Treatment Group —P3, Ghana 7
Table 4. Kenya School Sample 8
Table 5. Respondent Counts by Treatment Group—TuKu, Kenya 8
Table 6. Rwanda School Sample 9
Table 7. Respondent Counts by Treatment Group —Scaling Learning Through Play, Rwanda 9
Table 8. Comparison of Learning Environment and SEL Scores Across Countries. 12
Table 10. Observed Classroom Climate (Percentage of Observed Class Time), by Treatment Group in Ghana 13
Table 11. Student Agency (Percentage of Observed Class Time), by Treatment Group in Ghana 17

Table 12. Post-Observation Checklist (Percentage of Observed Lessons), by Treatment Group in Ghana .. 18
Table 13. Post-Observation Checklist Scores, by Treatment Group in Ghana 19
Table 14. Post-Observation Reading Classroom Inventory (Availability in Percentage of Classrooms), by Treatment Group in Ghana 20
Table 15. Responses to Learner Environment Questions by Students in Ghana 20
Table 16. Learning Environment Score by Treatment Status 20
Table 17. Stakeholder Attitudes toward LtP—Ghana 21
Table 18. Mean Attitudes Scale Scores toward LtP—Teachers and Head Teachers 22
Table 19. Caregiver Attitudes toward LtP - Ghana 22
Table 20. Which of the Following Supports Do You Receive as a Teacher? 23
Table 21. Caregiver: In the Past Three Days, Have You or Another Caregiver Engaged in Any of These Activities with Your Child? 24
Table 22. Caregiver: What Types of Play Materials Do You Have at Home? 25
Table 23. Head Teacher: How Do You Know if a Teacher in Your School Is Not Performing as Expected? . 26
Table 24. Head Teacher: What Action Do You Take if a Teacher Is Not Performing as Expected? 26
Table 25. Coaches: What Do You Think Are the Biggest Challenges to Increasing the Amount of Playful or Interactive Activities in the Classroom? 27
Table 26. Trainers: What playful, Interactive, or Student-Directed Learning Content Areas Are Covered in Program Trainings? 27
Table 27. DEOs: What Is Your (or Your Office's) Role in Supporting P3? 28
Table 28. Means Scores for Six Observed Student Behaviors 28
Table 29. Observed Behavior by Treatment Group 29
Table 30. Relationship Item Scores 29
Table 31. Average Percent of Relationship Items 29
Table 32. Percentage of Responses to Conflict Resolution Scenarios in Ghana 31
Table 33. Average Percent of Responses to Conflict Resolution Scenarios by Treatment Group 31
Table 34. Percentage Negative Attributions by Students 32
Table 35. Percentage Negative Attributions by Treatment Group 32
Table 36. Percentage Students Identifying with Confident Person 32
Table 37. Percentage Students Identifying with Confident Person by Treatment Group 33
Table 38. Average Creativity Score by Treatment Group 33
Table 39. EGRA and EGMA Results for Grade 1 Students in Ghana, by Subtask and Treatment Group 34
Table 40. Observed Classroom Climate (Percentage of Observed Class Time), by Treatment Group in Kenya 37
Table 41. Student Agency (Percentage of Observed Class Time), by Treatment Group in Kenya 41
Table 42. Post-observation checklist (percentage of observed lessons), by treatment group in Kenya 42
Table 43. Post-Observation Checklist Scores, by Treatment Group in Kenya 43
Table 44. Post-Observation Reading Classroom Inventory (Availability in Percentage of Classrooms), by Treatment Group in Kenya 44
Table 45. Responses to Learner Environment Questions by Students in Kenya 44
Table 46. Learning Environment Score by Treatment Status 44
Table 47. Stakeholder Attitudes toward LtP - Kenya 45
Table 48. Mean Attitudes Scale Scores toward LtP—Teachers and Head Teachers 45
Table 49. Caregiver Attitudes toward LtP—Kenya 46

Table 50. Which of the Following Supports Do You Receive as a Teacher? 47
Table 51. Caregiver: In the Past 3 Days, Have You or Another Caregiver Engaged in Any of These Activities with Your Child? 47
Table 52. Caregiver: What Types of Play Materials Do You Have at Home? 48
Table 53. Head Teacher: How Do You Know if a Teacher in Your School Is Not Performing as Expected? 49
Table 54. Head Teacher: What Action Do You Take if a Teacher Is Not Performing as Expected? 49
Table 55. Trainers: What Playful, Interactive, or Student-Directed Learning Content Areas Are Covered in Program Trainings? 50
Table 56. DEOs: What Is Your (or Your Office's) Role in Supporting TuKu? 50
Table 57. Student Behaviors Observed 51
Table 58. Observed Behavior Score by Treatment Group 51
Table 59. Relationship Item Scores in Kenya 51
Table 60. Average Percent of Relationship Items 52
Table 61. Percentage of Responses to Conflict Resolution Scenarios in Kenya 53
Table 62. Average Percent of Responses Conflict Resolution Scenarios by Treatment Group 53
Table 63. Percentage Negative Attributions by Students in Kenya 54
Table 64. Average Percentage of Negative Perception by Treatment Group 54
Table 65. Percentage of Students Identifying with Confident Person 54
Table 66. Percentage Students Identifying with Confident Person by Treatment Group 55
Table 67. Average Creativity Score by Treatment Group 55
Table 68. EGRA and EGMA Results for Grade 1 Students in Kenya, by Subtask and Treatment Group 56
Table 69. Observed Classroom Climate (Percentage of Observed Class Time), by Treatment Group in Ghana 58
Table 70. Student Agency (Percentage of Observed Class Time), by Treatment Group in Rwanda 61
Table 71. Post-Observation Checklist (Percentage of Observed Lessons), by Treatment Group in Rwanda 62
Table 72. Post-Observation Checklist Scores, by Treatment Group in Rwanda 63
Table 73. Post-Observation Reading Classroom Inventory (Availability in Percentage of Classrooms), by Treatment Group in Rwanda 64
Table 74. Responses to Learner Environment Questions by Students in Rwanda 64
Table 75. Learning Environment Score 65
Table 76. Stakeholder Attitudes toward LtP - Rwanda 65
Table 77. Mean Attitudes Scale Scores toward LtP—Teachers and Head Teachers 66
Table 78. Caregiver Attitudes toward LtP—Rwanda 66
Table 79. Which of the Following Supports Do You Receive as a Teacher? 67
Table 80. Caregiver: In the Past 3 Days, Have You or Another Caregiver Engaged in Any of These Activities with Your Child? 68
Table 81. Caregiver: What Types of Play Materials Do You Have at Home? 69
Table 82. Head Teacher: How Do You Know if a Teacher in Your School Is Not Performing as Expected? 70
Table 83. Head Teacher: What Action Do You Take if a Teacher Is Not Performing as Expected? 70
Table 84. Coaches: What Do You Think Are the Biggest Challenges to Increasing the Amount of Playful or Interactive Activities in the Classroom? 71
Table 85. DEOs: What Is Your (or Your Office's) Role in Supporting Scaling Learning Through Play? 71
Table 86. Student Behaviors 72
Table 87. Observed Behavior Score by Treatment Group 72

Table 88. Relationship Item Scores in Rwanda 72
Table 89. Average Percent of Relationship Items 73
Table 90. Percentage of responses to Conflict Resolution Scenarios in Rwanda 74
Table 91. Average Percent of Responses Conflict Resolution Scenarios by Treatment Group 74
Table 92. Percentage Negative Attributions by Students 75
Table 93. Percentage Negative Attributions by Treatment group 75
Table 94. Percentage Students Identifying with Confident Person 75
Table 95. Percentage Students Identifying with Confident Person by Treatment Group 76
Table 96. Average Creativity Score by Treatment Group 76
Table 97. EGRA and EGMA Results for Grade 1 Students in Rwanda, by Subtask and Treatment Group ... 77

ABBREVIATIONS

AAM	assessor accuracy measure
AKF	Aga Khan Foundation
COL	Community of Learning
COP	Community of Practice
COVID-19	coronavirus disease 2019 (SARS-CoV-2)
CPD	Continuous Professional Development
DED	District Education Director
DEO	District Education Official
DTO	District Training Officer
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
G	grade
GALOP	Ghana Accountability for Learning Outcomes Project
LtP	Learning through Play
PLAY	Playful Learning Across the Years
RISE	Research on Improving Systems of Education
RTI	RTI International (registered trademark and trade name of Research Triangle Institute)
RtP	Right to Play
SEL	social and emotional learning
SBM	School-Based Mentor
SBMT	School-Based Mentor Trainer
SISO	School Improvement Support Officer
TPD	teacher professional development
TuKu	Tucheza Kujifunza
UNICEF	United Nations Children's Fund
VVOB	Flemish Association for Development Cooperation and Technical Assistance (VVOB – education for development)
VNIES	Vietnam Institute of Educational Sciences
WERK	Women Educational Researchers of Kenya
Y	year

EXECUTIVE SUMMARY

1.1 Overview

The Play Accelerator research partner grant presents a unique opportunity to examine five large-scale educational interventions focused on improving learning through play (in Bangladesh, Ghana, Kenya, Rwanda and Vietnam, respectively). One aspect of this activity is determining if and how these five programs change the practice and attitudes of key education stakeholder and whether or not this leads to improved holistic learning outcomes through increasing playful pedagogies. This, in turn, will generate much-needed rigorous evidence on implementing playful pedagogies successfully through government teacher professional development systems at scale.

Due to prolonged school closures and other obstacles resulting from COVID-19, this report presents baseline findings collected between June and November 2021 from three of the five Play Accelerator partner programs: Partners in Play Program (P3) implemented by Right to Play in Ghana; Tucheza Kujifunza (TuKu) implemented by Aga Khan Foundation (AKF) in Kenya; and the Scaling Learning Through Play program implemented by the United Nations Children's Fund (UNICEF) in Rwanda.

The goal of this report is to provide comprehensive baseline estimates for these three Learning through Play (LtP) programs. These results will serve two purposes: 1) to provide a starting point against which program impact will be measured at midline and endline, and 2) to provide programs with data-driven recommendations of areas to focus on for improved implementation throughout the remainder of their programming.

While the overall evaluation of these programs includes a wide range of research questions that will be addressed at midline and endline, the baseline study focuses specifically on four main research questions:

1. What are the baseline levels of student literacy, numeracy, and social-emotional learning among Grade 1 students in treatment and control schools?
2. What is the baseline state of teachers' pedagogical practices in the classroom (including LtP) among Grade 1 teachers in treatment and control schools?
3. What are the baseline perceptions of LtP among school-, district-, and community-level stakeholders?
4. What are the enabling factors associated with implementing LtP pedagogical methods in classrooms? What are the barriers associated with implementing LtP pedagogical methods in classrooms?

1.2 Methodology, Instruments and Sample

This evaluation employs a quasi-experimental design that is longitudinal at the student level. The study includes three data collection time points for each program (baseline, midline, and endline). While the initial design was intended to cover a full 3 years of implementation, COVID-19 delays in implementation (along with school closures) led to a shortened time frame. For each of the three studies included in this report, the evaluation timeline now consists of a baseline in 2021, an anticipated midline in 2022, and an anticipated endline in 2023.

At baseline, students in this study were all in Grade 1. Program effects will be estimated as the difference between treatment and control students at the end of Grade 3, accounting for

basic student and school demographics, as well as baseline performance. Due to the student-level longitudinal nature of the design, the same schools will be visited at each time point but teachers will be re-sampled at each wave. At each time point, these student-level findings will be complemented with classroom observations and education stakeholder and caregiver interviews. The same instruments are administered in treatment and control schools for each program.

Accordingly, this report details baseline findings across five main instruments: 1) classroom observations administered in Grade 1 reading and mathematics lessons; 2) teacher and caregiver attitude scales; 3) fixed-response stakeholder interviews administered to teachers, head teachers, coaches, teacher meeting facilitators, trainers and district education officials (varying by program); 4) a shortened version of the Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA) administered to Grade 1 students; and 5) a social and emotional learning (SEL) and learning environment assessment administered to the same Grade 1 students as the EGRA and EGMA. Target baseline respondents for each instrument are listed, by program, in Table ES1 (final sample sizes are detailed in the report).

Table ES1. Target Interview Respondents, by Program (Country)

Interview	P3 (Ghana)	TuKu (Kenya)	Scaling LTP (Rwanda)
Teacher Interview & Attitudes Scale	Grade (G)1 teacher (*1 per school)*		
Caregiver Interview & Attitudes Scale	G1 caregiver (*2 per school)*		
Head Teacher Interview & Attitudes Scale	Head teacher (*1 per school)*		
Coach Interview	School Improvement Support Officer 1 per circuit	No coach interview	Director of Studies 1 per school
Trainer Interview	District Training Officer 1 per District	TuKu Trainers (from civil society organizations) ~10	School-Based Mentor 1 per school
Teacher Meeting Facilitator Interview	Curriculum Lead 1 per school	Assistant Head Teacher 1 per school	School-Based Mentor 1 per school
District Education Official Interview	District Education Director 1 per District	Sub-County Director ~10	District Education Officer 1 per District
Classroom Observations	G1 Literacy & Numeracy (1 each per school)		
EGRA, EGMA and SEL Assessment	G1 Students (16 per school; 8 boys, 8 girls)		

For each of the three programs, RTI worked in consultation with program staff to identify treatment and control groups for this evaluation. The sample sizes by program and treatment group are displayed in Table ES2. Additional details on treatment and control groups are provided in the report.

Table ES2. School Samples by Program					
Country	**Treatment / Control**				**Total**
Ghana	Treatment 1 (Direct)	Treatment 2 (Indirect)	Treatment 3 (GALOP)	Control	
	57	48	46	46	197
Kenya	Treatment 1 (Phase I)	Treatment 2 (Phase II)	Control		
	40	50	50		140
Rwanda	Treatment 1 (SBMT)	Treatment 2 (Non-SBMT)	Control 1 (SBMT)	Control 2 (Non-SBMT)	
	52	51	33	50	186

1.3 Findings and Recommendations

Findings in the report are disaggregated by program and treatment/control. For the purpose of this executive summary, we have provided high-level findings and recommendations that are applicable across programs.

Learning through Play in the Classroom

Teachers in all programs incorporated many basic best teaching practices in their lessons (e.g., student monitoring and support, material preparation, good instructional pace, making connections and sharing ideas, inclusionary practices). However, we recommend providing additional space and opportunities to incorporate more playful pedagogies in the classroom, as laid out below.

Student Engagement. Classrooms in all three programs showed evidence of strong levels of student engagement. This is a promising starting point for these programs. ***We recommend that programs now focus on improving the quality of engagement through more playful practices and socially interactive methods.***

Classroom Climate. For the large majority of observed lessons, the classroom climate was found to be 'ordinary' (meaning that there was neither a noticeably positive nor negative tone). While it's great to see extremely limited use of negative language or physical punishment, ***we recommend that programs seek to increase levels of positive climate over the course the implementation, as well as increased perseverance (or recognition of effort and encouragement of goal setting among students).***

Teaching Approach. Across all three programs, teachers spent the largest proportion of time explaining concepts, modeling activities and monitoring students' understanding, while students spent the largest proportion of time listening. These are all valuable and important components of teaching and learning but there was a noticeably limited use of more interactive or student-focused actions. ***We recommend that programs provide teachers with additional opportunities to incorporate more interactive approaches into their teaching, such as the use of open-ended questioning and turn-taking.***

Material Usage. Materials provide a great opportunity for programs to increase the use of playful practices in the classroom. Chalkboards and oral language were the most commonly used materials in two of the three programs, while one program had a larger proportion of book usage. Across the board, this left little time for the use of other written materials (such as letter/number cards or pocket charts), manipulatives and alternative materials (e.g.,

crafts, toys, illustrations, space, etc.). ***We recommend an increased use of these types of alternative materials (as well as book usage), in order to support improved access and use of print and engagement among students.***

Material Availability. In order for students to engage with materials, it is essential for those materials to be available in the classroom. Although some materials were available, **we recommend all programs to carefully examine the availability of materials and consider providing additional resources where possible (e.g., teaching aids, supplementary readers, workbooks, etc.). Increasing the availability of these materials in classrooms will improve students access to print and could be important for improving their performance on academic outcomes.**

Group Work. Although the majority of teachers in all three programs noted the importance of group work for student learning, less than 10% of the total class time was spent with students working in large or small groups. ***We recommend increased usage of group work and/or learning corners in order to improve social interaction and iteration among students.***

Student Agency. Arguably the area with the greatest room for improvement in these programs is student agency. Nearly all of class time across programs was dedicated to activities that were initiated and directed by teachers. Whether student-initiated or teacher-initiated and student-directed, ***we recommend providing students with increased levels of control over how they engage in an activity or, at a minimum, with improved opportunities for asking questions, expressing opinions and practicing skills. Increased student agency is at the foundation of improving learning through play in the classroom.***

LtP Enabling Environment

Learning Environment. Students in all programs reported incidences of physical and verbal aggression by their teachers in the past week. Roughly half (45%-55% across countries) of students said that teachers yell at them when they get the answer wrong. Interestingly, while students reported physical and verbal punishment, they also largely reported that they feel supported and listened to by teachers. Furthermore, verbal and physical aggression was rarely seen in classroom observations (which was likely impacted by the presence of an outside observer in the classroom). ***We recommend that programs further investigate the degree to which children feel safe and supported in classrooms and focus on ensuring that there are more positive and constructive interactions between teachers and students.***

Teacher Support. Head teachers in all programs often claimed that they provide additional support to struggling teachers themselves, although there was variation in who assists them with identifying a struggling teacher in the first place. ***We recommend that programs consider further investigating head teachers' capacity and availability to support struggling teachers in their schools, while identifying alternative avenues for support through other system actors (e.g., coaches, meeting facilitators, etc.).***

Training and Materials. Teacher interviews underscored the ongoing importance of training as a major support teachers receive – as well as teacher meetings. While teachers' understanding of how to use playful approaches in the classroom was a common barrier noted by trainers and coaches, the largest concern was with the availability of play-based materials (or lack thereof). ***We recommend that programs continue to focus both on training teachers on playful pedagogies, as well as the development/provision of***

materials (or through advocacy for approaches that require few or no additional materials).

Ministry Stakeholders. As anticipated, DEOs largely see their role as communicating program activities and monitoring implementation in schools. **We recommend leveraging district and sub county officials to clearly communicate expectations around learning through play (including what learning looks like, in terms of student agency and collaboration). Support from district officials can often play a large role in ensuring that activities are implemented as intended in schools.**

Perceptions of LtP

Attitudes Toward Play. Generally speaking, caregivers, teachers and education stakeholders are in support of LtP in theory but are less in support of a more student-centered, "free" classroom environment. This may be due in part to a misconception that student agency and more student-centered approaches will lead to a less disciplined and more chaotic classroom. ***We recommend that programs continue their work of orienting key stakeholders on the benefits of learning through play in the classroom and understanding their concerns/objections in order to provide solutions for improving attitudes toward specific playful practices.***

Student Holistic Outcomes

Conflict Resolution. Across programs, there was little evidence of advanced conflict resolution skills among students. Only 5-10% of students responded to another child in a conflict situation by asking for reasons for their behavior or by trying to influence them. Perhaps this is unsurprising among Grade 1 students. ***We recommend increasing activities that involve collaboration and interaction with other students to enhance children's conflict resolution skills.***

Conflict Resolution. Interestingly, Rwandan students were much more likely (nearly twice as likely as those in Ghana) to attribute a hostile intention to another student in a conflict situation. This may reflect difference in cultural and experience of students across countries. It may also be the result of different interpretations by assessors across countries. ***We recommend that that the Scaling LTP team explores this issue further in Rwanda, while we plan to recalibrate and standardize coding procedures across teams at midline.***

Creativity. There were low creativity scores across all programs, with only an average of about 4 ideas generated in the alternate uses test. ***We recommend increasing play-based approaches (such as the use of games and symbolic play) in order to improve creativity rates.***

Literacy and Numeracy. EGRA and EGMA results represent baseline estimates for students in Grade 1 for all programs. Therefore, results were relatively low across all subtasks (as expected). This led to a large proportion of zero scores (i.e., students who were unable to identify a single item correctly), particularly for higher order skills. These same skills will be measured at each time point and low baseline estimates should not be seen as a point of concern for these programs.

1. INTRODUCTION

1.1 Overview of Activity

The Play Accelerator research partner grant presents a unique opportunity to examine five large-scale educational interventions focused on improving learning through play. One aspect of this activity is determining if and how these five programs (implemented in Bangladesh, Ghana, Kenya, Rwanda and Vietnam, respectively) change the practice and attitudes of key education stakeholder and whether or not this leads to improved holistic learning outcomes through increasing playful pedagogies. This, in turn, will generate much-needed rigorous evidence on implementing playful pedagogies successfully through government teacher professional development systems at scale. A key component of this research is a longitudinal study aimed at measuring the learning experiences and holistic outcomes of children attending Grade 1 through 3 in schools participating in one of the five LEGO Play Accelerator Partner Programs, as compared with selected comparison (control) schools. These student-level findings are complemented with classroom observations and education stakeholder and caregiver interviews.

Due to prolonged school closures and other obstacles resulting from COVID-19, this report presents baseline findings collected between June and November 2021 from three of the Play Accelerator partner programs: Partners in Play Program (P3) implemented by Right to Play in Ghana; Tucheza Kujifunza (TuKu) implemented by Aga Khan Foundation (AKF) in Kenya; and the Scaling Learning Through Play program implemented by the United Nations Children's Fund (UNICEF) in Rwanda. An overview of these programs can be found in ***Annex D.*** Baseline findings for the other two programs will be reported in an addendum following data collection.

1.2 Purpose and Objectives of Baseline

The goal of this report is to provide comprehensive baseline estimates for the Learning through Play (LtP) programs in Ghana, Kenya, and Rwanda. These results will serve two purposes: 1) to provide a starting point against which program impact will be measured at midline and endline, and 2) to provide programs with data-driven recommendations of areas to focus on for improved implementation throughout the remainder of their programming.

While the overall evaluation of these programs includes a wide range of research questions that will be addressed at midline and endline, the baseline study focuses specifically on four main research questions:

1. What are the baseline levels of student literacy, numeracy, and social-emotional learning among Grade 1 students in treatment and control schools?
2. What is the baseline state of teachers' pedagogical practices in the classroom (including LtP) among Grade 1 teachers in treatment and control schools?
3. What are the baseline perceptions of LtP among school-, district-, and community-level stakeholders?
4. What are the enabling factors associated with implementing LtP pedagogical methods in classrooms? What are the barriers associated with implementing LtP pedagogical methods in classrooms?

1.3 Evaluation Design, Instruments, and Sample

Evaluation Design

This evaluation uses a quasi-experimental design that is longitudinal at the student level. The study includes three data collection time points for each program (baseline, midline, and endline). While the initial design was intended to cover a full 3 years of implementation, COVID-19 delays in implementation (along with school closures) led to a shortened time frame. For each of the three studies included in this report, the evaluation timeline now consists of a baseline in 2021, an anticipated midline in 2022, and an anticipated endline in 2023.

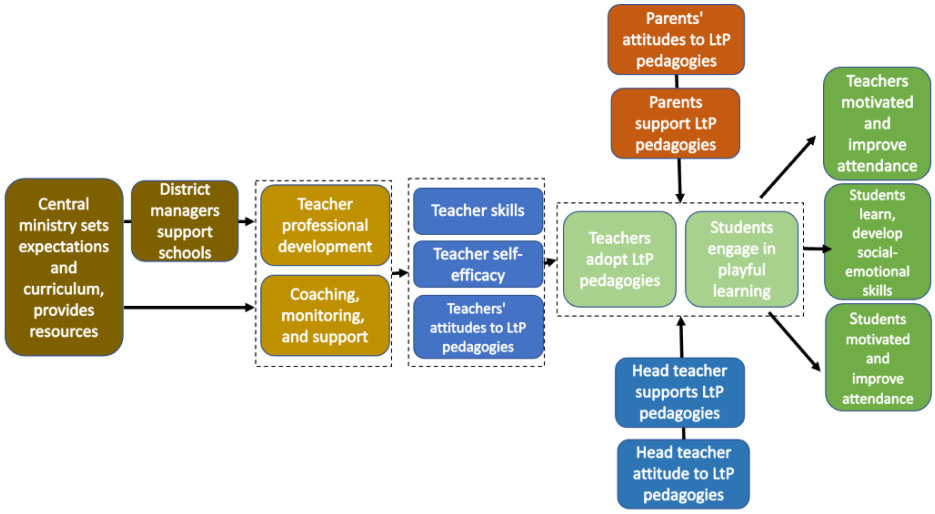
Since the main focus of this evaluation is on the impact on student holistic outcomes, the design is longitudinal at the student level. Baseline data collection occurred with students in Grade 1. These same students will be tracked over time and assessed in Grade 2 at midline and at the end of Grade 3 at endline. Program effects will be estimated as the difference between treatment and control students at the end of Grade 3, accounting for basic student and school demographics, as well as baseline performance. Due to the student-level longitudinal nature of the design, the same schools will be visited at each time point, but teachers will be re-sampled at each wave.

While the intent was for all baseline data collection to occur prior to the start of implementation for any of these programs, this was not ultimately possible due to COVID-19-related delays. Therefore, some intervention activities were conducted in each of the three programs prior to baseline—though no comprehensive school-level intervention occurred, so we expect these activities to have only minor impacts on the baseline results. All impact estimates for these programs (at midline and endline) will account for baseline differences (either via difference-in-differences models or by adjusting for baseline covariates in regression models).

Instruments

The instruments used in this study were developed to measure the components in the LtP Theory of Change *(***Figure 1**) as well as the relationship between different components.

Figure 1. Generalized Theory of Change for LtP Interventions



Classroom Observation

To understand the learning environment, classroom observation tools were used at two timepoints, during the lesson and following its conclusion. The tool used during the observation was a timed instrument that captured what happened every 2 minutes at seven levels of classroom practice (Facilitator, Student, Source [State of Play], Materials, Focus [the organization and the nature of the social interaction], Engaged, Climate [tone and initiator]). The tool used after the observation was a binary checklist with 29 items organized under six categories (Integrated Pedagogies, Application, Responsiveness, Resources, Efficiency, Student Autonomy). The post-observation tool also included a materials inventory with four categories (teacher materials, student materials, textbook ratio, library) and a means to capture a statement from the observer.

The constructs and the items that were included in the classroom observation tool were informed by the descriptions of play with an emphasis to include items that could be directly observed. The timed tool is considered low inference as the observers did not make evaluative statements. The post-observation checklist required more inferencing as it captured the presence of a desired behavior at least once during the observation. The materials checklist noted the presence of materials to support mathematics and literacy development. The observer statement was a means to capture an opinion of the observer and potentially to be used to help further validate our understanding of the environment.

Classroom observations were conducted following school-level introductions and getting the teacher's consent. Observers sat in a part of the classroom that was unobtrusive so as not to distract the students or the instruction. They coded on an A4-sized paper version of the timed instrument so that all seven levels and their codes would be visible simultaneously. Following the observation, the codes written on paper for each 2-minute interval were entered into Tangerine. The responses for the post-observation tools were entered directly into Tangerine. Item definitions and examples that were used in training were readily available to the observers during the observations.

Child Assessment: Social and Emotional Learning (SEL) and Learning Environment Tool

The SEL tool consists of six subtests. It was developed after a review of the SEL literature and is based on a number of existing tools (see below for details). We conducted two rounds of piloting with the SEL tool, assessing both learning environment and attitudes toward play, in Nairobi and Kitui counties in Kenya. The first pilot was with 150 children, 50 teachers, and 100 parents in 6 schools. Results pointed to the need for more tool adaptation. We made the adjustments and then conducted a second round of piloting in Kitui with 75 children, 50 teachers, and 50 parents. After these pilots, we used the results to shorten the subtests by dropping items that performed less well. Below is a summary of what we learned about each assessment.

Child behavior during assessment. This subtest involves the assessor rating students on six behaviors during the SEL assessment. It has been used in a number of previous studies; many items were adapted from the Preschool Self-Regulation Assessment. It had high reliability during the pilots in Kenya and Vietnam and is a helpful measure to triangulate with children's self-ratings of confidence, for example.

Relationships: This subtest assesses whether children have friends they can ask for help in difficult situations or when studying. It was adapted from the International Social and Emotional Learning Assessment (ISELA). It was quick and easy for children to understand. It had reasonable reliability.

Solving conflict: This subtest assesses students' response in a conflict situation and their hostile attribution bias (the tendency to attribute a negative motive to another child in the scenario). This subtest was adapted from the NYU/International Rescue Committee Social-Emotional Response and Information Scenarios (SERAIS) test (with items added from the ISELA). A key adaptation from the original version was that we used an open-ended question about children's response to conflict instead of a series of closed-ended questions. The tool had excellent psychometrics and was enjoyable for children.

Confidence paired statements: This subtest involves reading a series of paired statements to students, one of which involves a confident child and the other a non-confident child. The student is asked which child is most similar to them. The subtest was developed by RTI for use in Tanzania. The statements had good reliability in Kenya, and a reduced version had reasonable reliability in Vietnam. We piloted 20 items and selected the best 7.

Creativity: This subtest requires students to suggest alternative uses for ordinary objects (e.g., a spoon). It was based on "The Alternative Uses Test," a well-established test created in 1967. The data on this assessment had a good distribution.

Child assessment of learning environment: This scale involved students' rating their learning environment on 10 items. This scale took items from a number of other assessments, including RTI's school-related gender-based violence (SRGBV) classroom climate assessment and a 2012 tool developed for UNICEF child-friendly schools. Children's assessment of their learning environment in school had good reliability around two sets of items: positive aspects of the environment (warm, encouraging teachers) and negative aspects (verbally and physically aggressive teachers).

EGRA and EGMA

For measuring literacy and numeracy, versions of the Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA) were adapted. For the EGRA, four subtasks were selected: letter sound identification, nonword reading, oral passage reading, and reading comprehension. The letter sound and nonword subtasks for all evaluations were drawn from existing adaptations of the EGRA in each context. A new oral reading passage and accompanying reading comprehension questions was developed, in lieu of using existing passages. The decision to use a universal passage was spurred by the need for comparability across evaluations (as EGRA passages can vary greatly in their level of difficulty). The newly developed passages and comprehension questions were purposefully designed to be context-agnostic, by focusing on issues common to all students. However, these passages and questions underwent translation, back-translation, and review, in an effort to ensure that all language was contextually appropriate. All subtasks were administered in local languages of instruction.

For numeracy, we opted to use four of the core EGMA subtasks: level 1 addition, level 1 subtraction, quantity discrimination, and missing number. Instead of creating new items, we used the items from the EGMA Toolkit, which have been rigorously vetted and piloted.

Teacher and Caregiver Attitudes to Play Scale

This tool was based on a method developed by RTI to assess attitudes by asking respondents to endorse one or the other of a pair of contrasting statements. Statements are based on activities and attitudes that facilitate play in classrooms. Initially the same set of statements was piloted with parents and teachers, but responses were substantially different, which led to the development of two separate scales. The parent scale focuses on attitudes related to child agency and teacher behaviors. The teacher scale focuses on child

agency and aspects of the environment and teacher behavior that may facilitate play. Both scales had good reliability after poorly performing items were removed. Although initially developed for teachers, the teacher attitudes to play scale was also conducted with head teachers and a non-representative sample of trainers, coaches, district officials, and the designated facilitators of teacher meetings (teacher meeting facilitators). English versions of both scales can be found in ***Annex C*** while reliability test outcomes can be found in ***Annex B***.

Stakeholder Interviews

Stakeholder interviews focus on the support provided by the program, as well as the supports that these stakeholders offer to teachers and students.

This study developed and employed fixed-response survey instruments for teachers, head teachers, coaches, teacher meeting facilitators, trainers, and district education officials. These instruments were designed to collect background information while also measuring attitudes toward play and information about key program supports to a play-enabling environment. These survey instruments were informed by prior RTI work but were significantly adapted to ensure their alignment with and appropriateness for the Play Accelerator evaluation. All program partners reviewed these instruments and gave feedback, to ensure that the tools would adequately measure relevant constructs and components for each program. All instruments were revised to accommodate feedback from program partners and the LEGO Foundation. **Table 1** outlines the target respondents for each interview at baseline, by program.

Table 1. Target Interview Respondents, by Program (Country)

Interview	P3 (Ghana)	TuKu (Kenya)	Scaling LTP (Rwanda)
Teacher Interview & Attitudes Scale	Grade (G)1 teacher *1 per school*	G1 teacher *1 per school*	G1 teacher *1 per school*
Caregiver Interview & Attitudes Scale	G1 caregiver *2 per school*	G1 caregiver *2 per school*	G1 caregiver *2 per school*
Head Teacher Interview & Attitudes Scale	Head teacher *1 per school*	Head teacher 1 per school	Head teacher 1 per school
Coach Interview	School Improvement Support Officer 1 per circuit	No coach interview	Director of Studies 1 per school
Trainer Interview	District Training Officer 1 per District	TuKu Trainers (from civil society organizations) ~10	School-Based Mentor 1 per school
Teacher Meeting Facilitator Interview	Curriculum Lead 1 per school	Assistant Head Teacher 1 per school	School-Based Mentor 1 per school
District Education Official Interview	District Education Director 1 per District	Sub-County Director ~10	District Education Officer 1 per District
Classroom Observations	G1 Literacy & Numeracy 1 per school	G1 Literacy & Numeracy	G1 Literacy & Numeracy
SEL Assessment	G1 Students 16 per school (8 boys, 8 girls)	G1 Students 16 per school (8 boys, 8 girls)	G1 Students 16 per school (8 boys, 8 girls)

Table 1. Target Interview Respondents, by Program (Country)			
Interview	**P3 (Ghana)**	**TuKu (Kenya)**	**Scaling LTP (Rwanda)**
EGRA & EGMA	G1 Students 16 per school (8 boys, 8 girls)	G1 Students 16 per school (8 boys, 8 girls)	G1 Students 16 per school (8 boys, 8 girls)

Program Evaluation Designs and Samples

While it was not possible to randomly assign treatment and control groups for the LEGO Play Accelerator programs, the approach for each program included a random selection of treatment schools along with a random (or matched) sample of control schools. Details on the selection of each program's evaluation sample, and final instrument counts, are provided below.

Ghana – P3

P3's direct and indirect intervention activities are being implemented in 7 districts across 3 regions (Greater Accra, Volta and Northern). Overall, the program includes three distinct approaches to implementation that are to be included in this evaluation: 1) Direct intervention (100 schools within 7 districts); 2) Indirect intervention (256 schools within 7 districts); and 3) Ghana Accountability for Learning Outcomes Project (GALOP) schools (4,000 schools, including those in 5 P3 districts, are receiving master training and direct district-level support from P3; 6,000 additional GALOP schools are receiving indirect support through the master training only). The proposed design for this evaluation included 50 schools from each of these three treatment arms, along with 50 control schools.

In order to streamline our work and maximize the value of the data collected, RTI worked with the United Kingdom Open University's Centre for the Study of Global Development (CSGD)—which led the P3 internal evaluation baseline—to select the sample for this evaluation. Open University's multistage cluster sampling approach led to a final sample of 75 schools for each of the three intervention groups, as well as 75 control schools (which were selected from nearby districts in each of the three regions). Our evaluation sample consisted of 50 randomly selected schools from each of these 4 groups (3 treatment arms and 1 control arm). This random selection was stratified across the three regions (with 30 Northern schools, 10 Volta schools and 10 Greater Accra schools in each group). The final sample included schools from 11 districts across the three intervention regions.

The intended sample consisted of 200 schools, evenly distributed across the four groups. However, due to logistical complications during data collection, the final sample numbers varied slightly, as shown in ***Table 2***.

Table 2. Ghana School Sample					
Region	**Direct**	**Indirect**	**GALOP**	**Control**	**Total**
Greater Accra	10	10	10	10	40
Northern	36	29	27	25	117
Volta	11	9	9	11	40
Total	57	48	46	46	197

A final count of respondents, per instrument and treatment group in Ghana, is presented in ***Table 3.***

Table 3. Respondent Counts by Treatment Group —P3, Ghana				
Instrument	**CONTROL**	**DIRECT**	**INDIRECT**	**GALOP**
G1 Teacher Interviews	46	57	47	46
Caregiver Interviews	91	115	97	92
Head Teacher Interviews	46	55	46	45
Coach Interviews	13	9	12	13
Trainer Interviews		9		
Teacher Meeting Facilitator Interviews	37	55	45	45
District Education Official Interviews		10		
Classroom Observations – G1 Literacy	43	53	43	41
Classroom Observations – G1 Numeracy	42	55	45	39
SEL Assessments	665	854	737	692
EGRA & EGMA	677	859	741	699

Kenya - TuKu

The TuKu program's original scope consisted of implementation in 201 treatment schools across Lamu and Mombasa counties. Although the program has since expanded into two additional counties, the original scope was used to define the treatment sample for this evaluation. Implementation in Lamu and Mombasa is broken into three phases. Phase I consists of 20 schools in each county. These 'innovation schools' were purposefully selected by the program and will receive direct intervention, beyond that which will be received by schools in Phase II and Phase III. Phase II schools consist of an additional 40 schools per county to be added in Year 3, while Phase II schools consist of another 40 per county to be added in Year 4. These 160 Phase II and III 'network' schools will all be supported by the 40 innovation schools (for a total of approximately 100 schools per county).

At the time of sample selection for this evaluation, TuKu had not yet selected its Phase II schools. Therefore, the RTI evaluation team proposed 80 Phase II schools (randomly assigned across Lamu and Mombasa counties) for the program to use. This was necessary in order to collect baseline data in these schools, prior to the start of implementation.

Since the innovation schools were purposefully selected, it was not possible to identify a matched set of control schools for this phase. However, it was possible to identify a matched set of control schools for the selected Phase II network schools. These control schools were selected from Kilifi County.

Despite TuKu's expansion into Kilifi County, AKF agreed to withhold implementation in the zones that RTI selected for comparison schools, until the end of the 2023 school year. This includes all schools in Malindi Central, Kilifi Central, and Watamu zones. For the control sample selection, schools in Lamu were matched with schools in Kilifi Central, while schools in Mombasa were matched with schools in Malindi Central and Watamu.

Ultimately, the sampling was as follows (and shown in ***Table 4***):

1. All 40 innovation (Phase I) schools selected (20 in Lamu and 20 in Mombasa) at baseline.
2. 50 network (Phase II) schools randomly sampled (20 in Lamu and 20 in Mombasa).
3. 50 control schools matched with network schools (Kilifi County: 25 in Kilifi Central, 16 in Malindi Central, and 9 in Watamu).

Table 4. Kenya School Sample				
County	**Phase I**	**Phase II**	**Control**	**Total**
Mombasa	20	25	-	45
Lamu	20	25	-	45
Kilifi	-	-	50	50
Total	40	50	50	140

A final count of respondents, per instrument and treatment group in Kenya, is presented in ***Table 5***.

Table 5. Respondent Counts by Treatment Group—TuKu, Kenya			
Instrument	**Phase I**	**Phase II**	**Control**
G1 Teacher Interviews	40	50	49
Caregiver Interviews	80	99	96
Head Teacher Interviews	39	47	47
Trainer Interviews	12	2	0
Teacher Meeting Facilitator Interviews	39	50	47
District Education Official Interviews	4	5	2
Classroom Observations – G1 Literacy	40	50	50
Classroom Observations – G1 Numeracy	39	50	50
SEL Assessments	623	762	765
EGRA & EGMA	626	763	766

Rwanda – Scaling Learning through Play

The UNICEF program in Rwanda is a national scale intervention, working in all 416 sectors (sub-division of districts) across the country. In each sector, LEGO/UNICEF-supported district-based Continuous Professional Development advisors will train School-Based Mentor Trainers (SBMTs). These SBMTs will then train the remaining School-Based Mentors (SBM) in their sector (and SBMs will in turn train the remaining teachers in their respective schools). There is one SBM in each school and an average of two SBMTs per sector.

Since SBMTs were purposefully selected and will be supported directly by program staff at the district level—while the remaining schools in each sector will receive support using a cascade approach led mostly by the initially trained SBMs—the two groups are treated as separate intervention arms for this evaluation. Each intervention group also has its own control group. With a proposed sample of 50 schools in each group, the originally intended sample for this evaluation consisted of 200 schools.

Although the program was intended to be conducted at national scale from the start, UNICEF agreed to withhold the intervention from Grades 1–3 until after the midline data collection in 25 randomly selected sectors.[2] The number of sectors per province was determined as a proportion of population by province, while sectors within each province were then randomly selected.

[2] However, the approach itself is school level, so there are still some concerns about spillover (even though the program will not explicitly target grades 1–3).

Educational experts in Rwanda then selected matched treatment sectors (based on geography, size, urbanicity, etc.) for each of the 25 control sectors. Since the number of SBMTs in these sectors was insufficient for our proposed sample size (with an average 1.5 SBMTs per sector), an additional 8 treatment sectors were added. Unfortunately, it was not possible to add additional control sectors. Therefore, the final sample consisted of schools from 33 treatment sectors and 25 control sectors. Within each sector, all SBMT schools were selected, while an additional 2 non-SBMT schools were selected from the original 50 sectors. Ultimately, this led to a final sample of 186 schools, as shown in ***Table 6***.

Table 6. Rwanda School Sample

	SBMT	Non-SBMT	Total
Treatment	52	51	103
Control	33	50	83
Total	85	101	186

A final count of respondents, per instrument and treatment group in Rwanda, is presented in ***Table 7.***

Table 7. Respondent Counts by Treatment Group —Scaling Learning Through Play, Rwanda

Instrument	SBMT Treatment	Non-SBMT Treatment	SBMT Control	Non-SBMT Control
G1 Teacher Interviews	51	51	31	49
Caregiver Interviews	103	98	62	99
Head Teacher Interviews	52	50	33	49
Coach Interviews	50	51	32	50
Trainer Interviews	51	49	32	50
Teacher Meeting Facilitator Interviews	48	49	31	50
District Education Official Interviews	11			
Classroom Observations – G1 Literacy	52	51	32	50
Classroom Observations – G1 Numeracy	45	32	26	38
SEL Assessments	822	800	511	787
EGRA & EGMA	829	807	531	789

Training and Data Collection

During the three baseline assessor trainings, participants were divided into three groups—classroom observers, EGRA/EGMA assessors, and SEL interviewer—based on their previous experience and skillsets. Within groups, RTI technical staff trained participants to specialize in specific content and tools. Trainees participated in an assessor accuracy measure (AAM) that evaluated their ability to accurately score items. Participants watched a video and used a tablet to score as if they were conducting the assessment or observation. The scores were then compared to a Gold Standard scoring, which had previously been completed by a trainer scoring the same video. Problematic items identified through the AAM exercise were addressed and improved upon during the training. AAM scores, as well as observations conducted during two practice school visits (on Days 3 and 5 of training), were used to select the top performing assessors to participate in fieldwork. All trainings ran for 5 days, with a sixth day of training for a subset of participants selected to serve as team leaders.

All data collections spanned a 10-to-15-day period, during which each team of three assessors visited one school per day. In advance of school visits, a team of trained "runners" visited each school to make introductions to the head teacher, randomly sample students, and to work with the head teacher to request informed consent from parents. Recognizing the risk to our sample posed by selected students being absent on the day of data collection and the likelihood of not all parents being reachable or giving consent, teams oversampled six additional students at each school. During the data collection, student assessments, classroom observations, and caregiver interviews were prioritized during the schools' morning hours to adhere to the schools' timetable and caregivers' schedules. School-level stakeholder interviews were conducted during the school day when respondents were available. District Education Office (DEO) and trainer interviews were conducted at the respondent's convenience at district offices or other agreed-upon meeting places.

Details on each training and data collection activity, including AAM scores, are provided below.

Ghana—P3

In Ghana, DevtPlan Consult, a local research firm subcontracted by RTI, recruited 48 experienced data collectors to participate in a 5-day training held May 17–21, 2021, outside of Accra. A subset of participants who were selected to serve as team leaders received an additional day of training on May 22. Advance school visits for sampling and consent were conducted from May 24 to 28, 2021, followed by the full data collection from May 31 to June 18, 2021.

Average scores on the final AAM are as follows:

- Classroom Observation—85% agreement
- SEL—88% agreement
- EGMA subtasks (all assessors)—90% agreement
- EGRA subtasks (Ga assessors)—91% agreement
- EGRA subtasks (Asante Twi assessors)—85% agreement
- EGRA subtasks (Dagbani assessors)—86% agreement
- EGRA subtasks (Ewe assessors)—88% agreement

Kenya—TuKu

In preparation for data collection, 45 experienced assessors from the RTI and Women Educational Researchers of Kenya (WERK) (subcontractor) database were contacted and invited to attend a one-week training held between October 18 and 23. Advance school visits for sampling and consent were conducted alongside training from October 18 to 23, 2021, followed by the full data collection from October 25 to November 5, 2021.

Average scores on the final AAM are as follows:

Classroom Observation—87% agreement

- SEL—87% agreement
- EGMA subtasks—97% agreement
- EGRA subtasks (Kiswahili)—97% agreement

Rwanda—Scaling LtP

Training of 60 assessors was conducted from May 24 to 29, 2021, with support from Incisive Africa (subcontractor). Advance school visits for sampling and consent were conducted alongside training from May 31 to June 4, 2021, and the full data collection was conducted from June 1 to 15, 2021.

Average scores on the final AAM are as follows:

- Classroom Observation—83% agreement
- SEL—84% agreement
- EGMA subtasks—95% agreement
- EGRA subtasks (Kinyarwanda) —95% agreement

2. FINDINGS

Overall, at baseline, school- and district-level staff had positive attitudes about the importance of learning through play in the classroom, including the idea of engaging children in small groups. However, these same staff tended to have a more difficult time with supporting more student autonomy and self-direction which, although necessary for group work and collaborative learning, was viewed as risking a loss of control and order in the classroom. Similar to the responses from teachers, parents and caregivers were in favor of LtP (though at a lower rate), but not at the expense of a disciplined and serious learning environment. Teacher meeting facilitators, coaches, and DEOs noted that a lack of materials to support play-based learning is a major challenge.

During classroom observations, there was evidence of teachers incorporating best teaching practices in their lessons (such as modelling activities, monitoring student progress, providing support to students, material preparation, etc.). However, the classrooms in all programs tended to have limited interactive components, and student choice and autonomy were noticeably absent. Additionally, although teachers praised the importance of group work, classroom observations showed that teachers implemented group work or learning corners quite infrequently during their lessons.

Students were consistently engaged in the majority of observed lessons across programs, but the quality of this engagement was hindered by a lack of available materials in classrooms.

Students reported generally helpful, encouraging learning environments, with the exception of significant incidences of yelling and harsh punishments by teachers for children in school. Although social and emotional competencies were promising, few students responded to conflict-based scenarios in more socially and emotionally productive ways, preferring to disengage entirely, commanding a peer to stop, or telling the teacher.

Table 8 shows the summary scores across countries of the Learning Environment and SEL assessments. There was a trend in the data for Rwanda to score more highly than Kenya, which scored more highly than Ghana. This was true in all assessments except for "observed behavior" and "median number of friends." It was also the case for percentage of children attributing hostile intentions to another child. This subtest showed the biggest cross-country differences with students attributing a hostile intention in more than 80% of scenarios in Rwanda but only 42% of scenarios in Ghana. In all other cases the differences between countries were small.

Table 8. Comparison of Learning Environment and SEL Scores Across Countries.			
Assessment	**Ghana**	**Kenya**	**Rwanda**
Learning Environment (/20)	11.7	11.8	14.9
Observed Behavior (/15)	7.8	9.3	8.2
Relationship (/5)	3.6	4.0	4.2
Median Number of Friends	4	4	3
Conflict Resolution (/12)	5.5	6.05	6.27
Hostile Attribution (%)	42.4	58.5	81.6
Confidence (/7)	3.9	4.4	4.8
Mean Creativity Score	4.9	4.0	3.4

EGRA and EGMA results represent baseline estimates for students in Grade 1. Therefore, results were relatively low across all subtasks (as expected). This led to a large proportion of zero scores (i.e., students who were unable to identify a single item correctly), particularly for higher order skills.

In the remainder of this section, we describe the baseline findings for each program in detail.

2.1 Partners in Play Program (P3) - Ghana

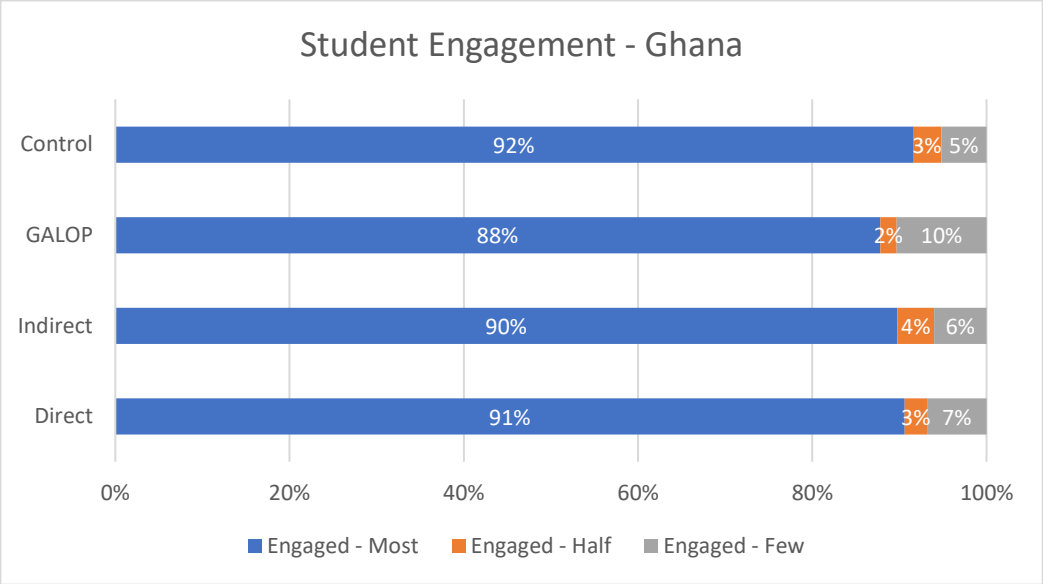
What is the baseline state of teachers' pedagogical practices in the classroom (including LtP) among Grade 1 teachers in treatment and control schools?

Classroom observations in Ghana were conducted in Grade 1 classrooms for both reading and mathematics lessons. This decision was made in order to ensure variation in pedagogical practices (without being limited by approaches used for a particular subject). Accordingly, results for reading and mathematics lessons are combined throughout this section. Results are averaged across classroom observations at the school level and are disaggregated by treatment group throughout this section, with observations occurring in 56 direct implementation schools, 46 indirect implementation schools, 44 GALOP schools, and 44 control schools.

Timed Classroom Observations

Across all observed lessons, assessors in Ghana reported that most students were engaged approximately 90% of the time, with little variation across treatment groups (**Figure 2**). This strong level of engagement is a promising starting point for the program to continue to improve the depth and breadth of student engagement through implementation.

Figure 2. Student Engagement During Lessons (Percentage of Observed Class Time), by Treatment Group in Ghana



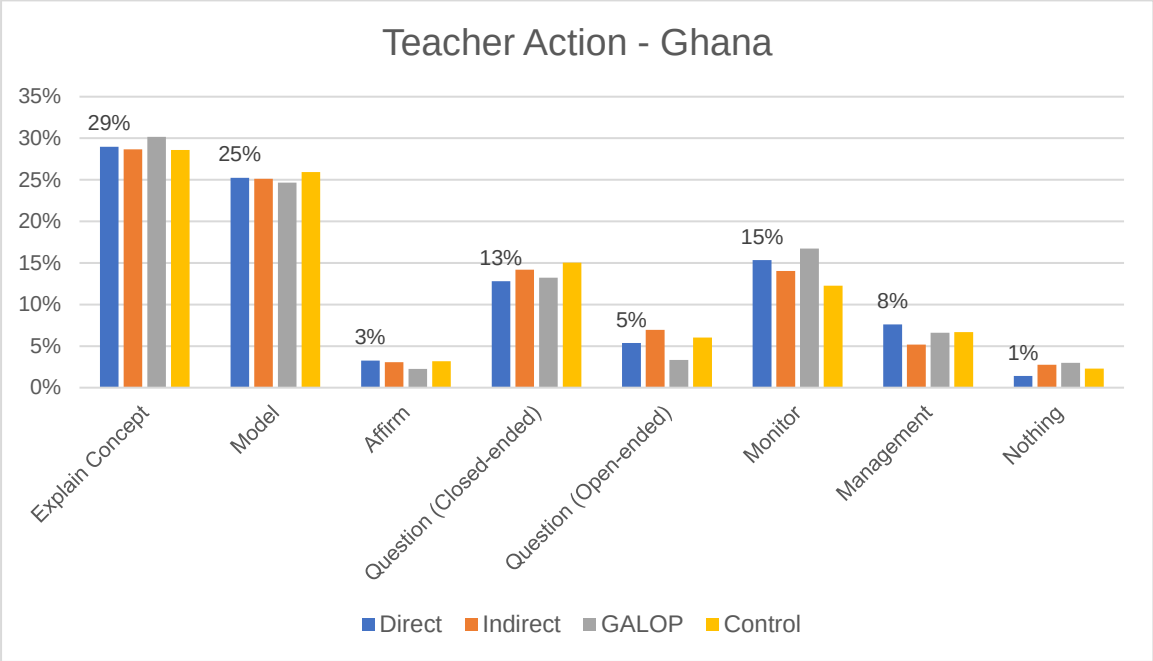
Assessors were also tasked with identifying the classroom climate during observed lessons, focused on four main criteria: perseverance (e.g., recognition of effort or encouragement of goal setting), positive verbal/physical (e.g., use of positive language of physical evidence of joy), negative verbal/physical (e.g., use of negative language or physical punishment), and ordinary (e.g., tone is neither noticeably positive nor negative). Overall, assessors observed an ordinary climate between 76% and 84% of the time, with the remaining time generally made up by positive climate (***Table 10***). There was very little perseverance or negative climate.

Table 10. Observed Classroom Climate (Percentage of Observed Class Time), by Treatment Group in Ghana

	Direct	Indirect	GALOP	Control
Ordinary	80%	76%	84%	82%
Positive verbal/physical	19%	22%	14%	17%
Negative verbal/physical	1%	1%	0%	1%
Perseverance	0%	1%	1%	0%
Number of schools	56	46	44	44

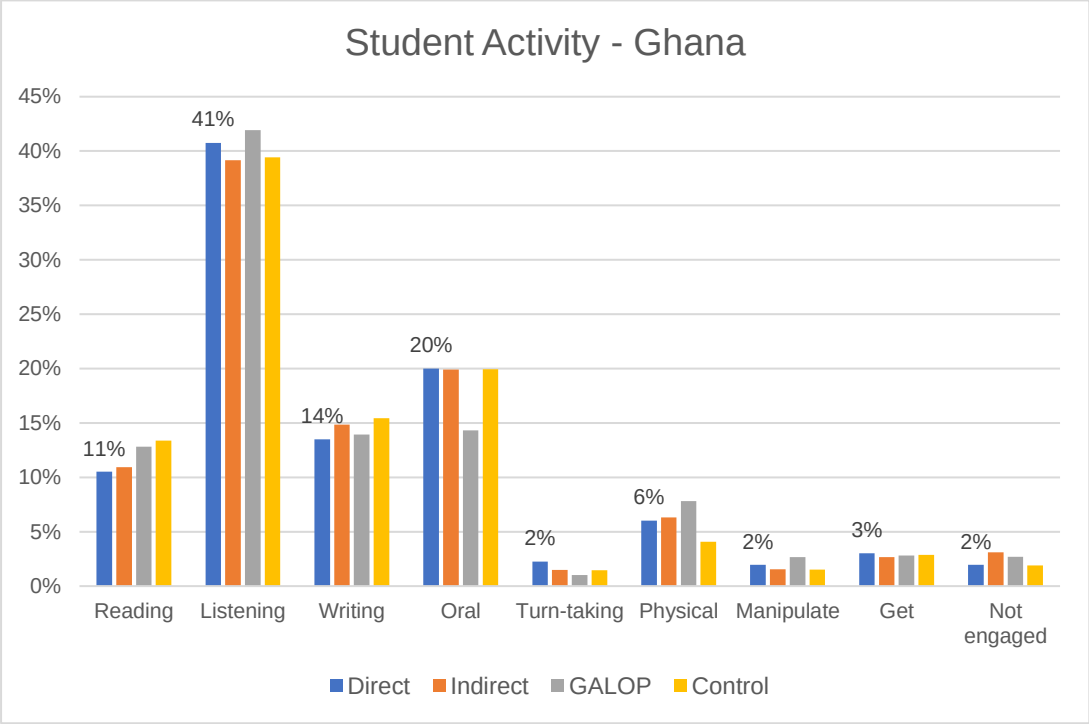
During the observed lessons, teachers spent more than half of their time either explaining concepts or modeling activities (**3**). They spent the next largest amount of time monitoring students' understanding (or tracking participation in an activity). More interactive approaches were used less frequently in the classroom, which is a viable area of improvement for the program. For example, open-ended questioning was used by teachers approximately half as often as closed-ended questioning. Teacher affirmations or recognition of student input made up a small proportion of time, while management or an unengaged teacher (i.e., nothing) accounted for just under 10% of class time.

Figure 3. Teacher Actions During Lessons (Percentage of Observed Class Time), by Treatment Group in Ghana



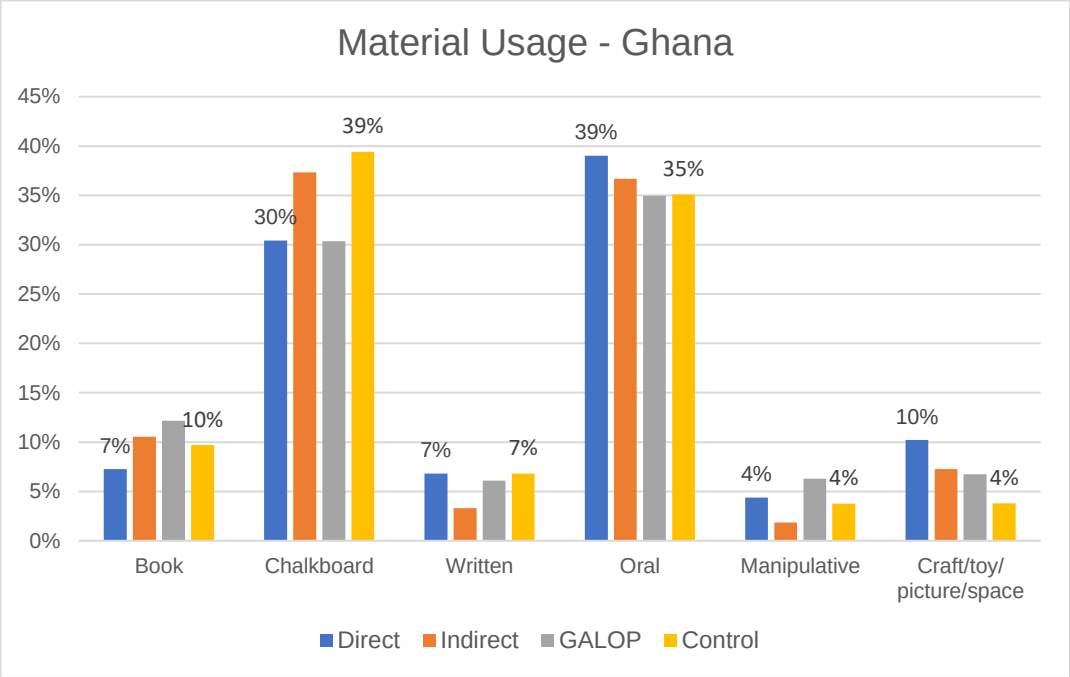
Supporting these findings, student activities are displayed in **Figure 4**. Approximately 40% of class time (across all treatment groups) was spent with children listening, while about one-fifth of the time consisted of students orally responding. Interestingly, even though these are Grade 1 classrooms, slightly more time was spent writing than reading. About 10% of class time overall was dedicated to students either discussing/responding to one another (turn-taking), having a physical response, or manipulating something with their hands. Students spent approximately 5% of class getting something to work on, waiting for the teacher, or disengaged (i.e., 'get' and 'not engaged' in **Figure 4**).

Figure 4. Student Activity During Lessons (Percentage of Observed Class Time), by Treatment Group in Ghana



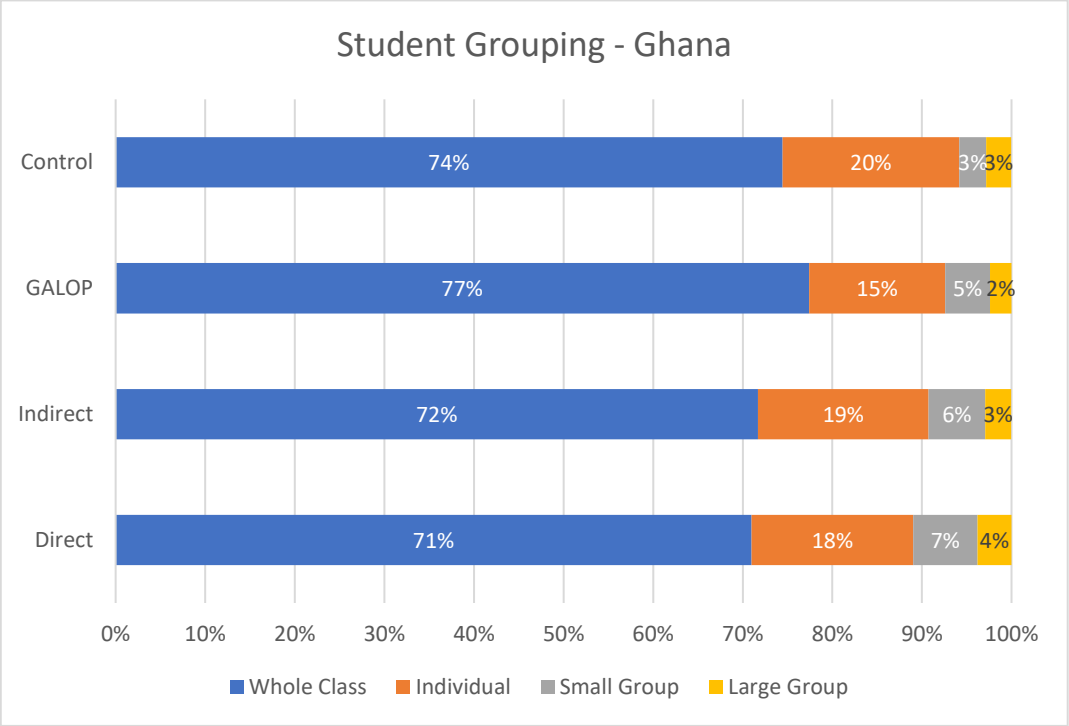
Materials provide a great opportunity for programs to increase the use of playful practices in the classroom. At baseline, between 65% and 75% of class time across treatment groups was spent either using a chalkboard or with material being presented orally (**Figure 5**). The remainder of the time was split somewhat evenly among the use of books, written materials (such as letter/number cards or pocket charts), manipulatives and alternative materials (e.g., crafts, toys, illustrations, space, etc.). While the direct implementation schools tended to have slightly less book usage, they had slightly higher levels of other written and alternative material usage than the other treatment groups (and control).

Figure 5. Use of Materials During Lessons (Percentage of Observed Class Time), by Treatment Group in Ghana



Although the majority of teachers in Ghana said that children learn language/literacy best through group work and that student exploration is important (on the attitude survey), these are still areas for improvement in actual instruction. The results of the classroom observations showed that approximately three-quarters of the class time was spent on whole class instruction, with 15% to 20% of the time used for individual work (**Figure**). Limited proportions of the total class time were spent with students working in large or small groups.

Figure 6. Student Grouping During Lessons (Percentage of Observed Class Time), by Treatment Group in Ghana



Note: "Corners" omitted from the table because <1% of time was spent using learning corners in observed classrooms.

There was also evidence of limited student agency found in the classroom observations, with nearly all of class time dedicated to activities that were initiated and directed by teachers (***Table 11***). Only about 5% of time across treatment groups was available for either student-directed (e.g., students doing an activity by their choice) or student-initiated work (students doing an activity that was initiated by the teacher but with ability to make decisions about how they do it).

Table 11. Student Agency (Percentage of Observed Class Time), by Treatment Group in Ghana

	Direct	Indirect	GALOP	Control
Teacher directed	95%	94%	95%	96%
Teacher initiated; Student directed	2%	1%	1%	1%
Student initiated	3%	4%	4%	3%

Post-Classroom Observation Indicators

At the end of each timed classroom observation, assessors completed a checklist to determine whether a range of activities occurred (at least once) during the lesson. These indicators are divided into six constructs: 1) integrated pedagogies, 2) application (use of games), 3) responsiveness (to student behaviors), 4) resources (availability and use), 5) efficiency (maximized instructional time), and 6) student-directed (ownership or autonomy).

As shown in ***Table 12***, there was significant variation in the prevalence of activities occurring throughout the observed lessons (with high prevalence in green, moderate prevalence in yellow, low prevalence in red). In terms of constructs, teachers tended to perform best on

responsiveness and efficiency indicators, with the most limited use of student-directed opportunities. Integrated pedagogies, application, and resources all include some high prevalence and some low prevalence items. Overall, basic best practices (e.g., student monitoring and support, material preparation, good instructional pace, making connections and sharing ideas, inclusionary practices) were used throughout lessons, but more playful pedagogies have significant room for improvement.

Table 12. Post-Observation Checklist (Percentage of Observed Lessons), by Treatment Group in Ghana

	Direct	Indirect	GALOP	Control
Integrated Pedagogies				
The teacher asked questions or raised topics that got students to respond to each other, not to her.	46%	33%	30%	38%
Student(s) persisted with the activity after the rest of the class has started to do something else.	21%	17%	11%	20%
Student(s) started the task again or adjusted when they thought of a better idea/approach.	21%	23%	15%	15%
Student(s) had a chance to make connections between the new experiences/content to their prior knowledge or experiences.	56%	41%	50%	44%
Students problem solved or had a chance to think critically or reflect about a topic.	74%	59%	52%	71%
Application: Students have generalized games				
Songs or energizers to start or divide activities.	90%	77%	80%	79%
Students played games with obvious rules.	40%	31%	24%	17%
Students had physical play.	44%	27%	33%	20%
Students used objects to represent something else.	15%	13%	12%	9%
Students had symbolic play.	16%	13%	3%	7%
Responsiveness: The teacher adapts to student behaviors				
The teacher gave students hints, suggestions, or redirection when they needed it.	91%	88%	88%	86%
The teacher circulated, leaned in, redirected, to support students as needed.	72%	68%	66%	66%
The teacher asked the class or a group for a solution to a student question or problem.	60%	49%	37%	47%
The teacher included students who did not volunteer to answer.	86%	76%	74%	65%
Resources: Materials that are needed for implementation are used				
Students used a resource that was dedicated for their use (e.g., book, workbook, slate, tablet).	47%	38%	52%	36%
Students used something to create something else (e.g., art supplies, items in nature, objects repurposed).	8%	2%	4%	3%
Students used a resource to solve a problem (e.g., manipulatives, math charts, dictionary).	32%	19%	28%	28%
Instruction used an item designed for interaction between teacher and students or between students (e.g., read aloud, pocket chart, counters, puzzle).	58%	47%	47%	45%
Students used physical space for learning activities (e.g., counting large steps for measurement, sit in a circle for sharing).	33%	20%	27%	13%
Students used stations/corner work/learning centers.	19%	6%	19%	5%
Efficiency: Instructional time is maximized				

Table 12. Post-Observation Checklist (Percentage of Observed Lessons), by Treatment Group in Ghana

	Direct	Indirect	GALOP	Control
The teacher had materials ready at the start of an activity or class.	82%	90%	88%	79%
The pace of the instruction was lively.	92%	80%	83%	88%
Student-directed (Ownership/autonomy)				
Students developed and shared ideas with classmates or teacher.	60%	47%	37%	46%
Students had a chance to select from choices.	20%	16%	8%	19%
Students chose collaborators or roles.	16%	7%	10%	2%
Students made or changed rules for an activity.	5%	2%	0%	0%
Students chose how long to do an activity.	8%	6%	12%	0%
Students set goals or helped to plan how to use time.	4%	0%	1%	1%
Students chose when to move around.	10%	9%	9%	5%
Number of schools	57	48	46	46

Note: The color scale is a gradient based on the proportion of lessons with observed indicators. Green = high prevalence; yellow = moderate prevalence; orange = low prevalence; red = extremely low prevalence.

In addition to the item level estimates in Table 22, construct level scores were calculated (***Table 13***). These scores will be used as the basis for measuring change/impact across treatment groups, over time. The scores are calculated as simple averages of the number of items occurring per construct for each treatment group. At baseline, direct implementation schools tended to have slightly higher scores than the comparison groups with several indicators showing statistically significant differences between direct and control (which will be accounted for in midline/endline impact estimates).

Table 13. Post-Observation Checklist Scores, by Treatment Group in Ghana

	Direct	Indirect	GALOP	Control
Integrated Pedagogies (0–5)	2.2	1.7	1.6	1.9
Application: Students have generalized games (0–5)	2.0*	1.6	1.5	1.3
Responsiveness: The teacher adapts to student behaviors (0–4)	3.1*	2.8	2.7	2.6
Resources: Materials that are needed for implementation are used (0–6)	2.0*	1.3	1.8	1.3
Efficiency: Instructional time is maximized (0–2)	1.7	1.7	1.7	1.7
Student-Directed (Ownership / Autonomy) (0–7)	1.2*	0.9	0.8	0.7
Overall (0–29)	12.2*	10.0	10.0	9.5

*p<0.05 (comparison between Direct and Control)

The post-observation component of the data collection also included a classroom inventory. The inventory was focused on determining the availability of basic materials in each classroom. For simplicity, the results displayed in ***Table 14*** are for reading lessons only. Overall, the availability of materials in the observed classrooms (across all treatment groups) was low. Approximately two out of every five classrooms had a teacher's guide and teaching aids (though direct implementation schools did notably have the largest proportion of teaching aids available). While textbooks were present in 29% of classrooms, on average only 14% of individual students had a textbook available. Lastly, only one-quarter of classrooms had workbooks and only 6% had supplementary readers.

Table 14. Post-Observation Reading Classroom Inventory (Availability in Percentage of Classrooms), by Treatment Group in Ghana

	Direct	Indirect	GALOP	Control	Overall
Teacher's guide	39%	29%	28%	46%	36%
Teaching aids	55%	27%	39%	34%	40%
Textbooks (Present in classroom)	21%	31%	35%	32%	29%
Textbook ratio	10%	18%	16%	12%	14%
Workbooks	29%	23%	24%	30%	26%
Supplementary readers	5%	4%	13%	2%	6%
N	56	48	46	44	194

Learning Environment (Student Survey)

Students were asked 10 questions about the learning environment (see ***Table 15***) with each response giving a score of 0, 1, or 2. ***Table 16*** shows that the learning environment score was similar across treatment groups, with a mean of 11.7.

Table 15. Responses to Learner Environment Questions by Students in Ghana

Item	No/Never	Sometimes	Yes/Often	Missing/No Reply
Does your teacher praise you when you give the right answer?	3	48	48	1
Can you talk to at least one teacher in the school about things that are bothering you?	29	37	33	1
Last week, did children in your class ask the teacher questions in lessons?	30	50	19	1
Last week, did some children in your class ask the teachers for help with their classwork?	31	51	16	
Do your teachers encourage you to try out new things in class?	14	49	36	1
Do your teachers get angry or yell when you get the answer wrong?	49	40	10	
Do your teachers help those who do not understand?	7	54	38	1
In the last week, did an adult in your school scream or yell at a child?	40	48	10	
Can students share their thoughts in front of the class?	30	41	28	1
In the last week, did an adult in your school hit a child or use harsh punishment?	32	54	12	1

Table 16. Learning Environment Score by Treatment Status

Treatment Status	Mean Score
Control	12.0
Indirect	11.7
Direct	11.6
GALOP	11.4
Overall	11.7

What are the baseline perceptions of LtP among school, district, and community-level stakeholders?

Attitudes toward LtP

Overall, attitudes towards LtP were similar across stakeholders, with trainers and DEOs having slightly lower support for LtP practices than head teachers, teachers, coaches, and teacher meeting facilitators (see ***Table 17***).

Nearly all of the interview participants strongly supported play as an important component of learning for students in primary school (Item 1), believed that parents and caregivers also welcome play-based approaches (Item 2), and prioritized student engagement and exploration (Items 3 and 4). The areas that need the most improvement are the views on student self-directed (or peer-directed) work, as shown by the low agreement in Items 8 through 10.

Table 17. Stakeholder Attitudes toward LtP—Ghana

Table Heading	Head Teacher	Teacher	Coach	Meeting Facilitator	Trainer	DOE
1-Play is an important part of learning for children in primary school	78%	85%	72%	81%	78%	90%
2-Parents welcome more play-based approaches in the classroom	73%	70%	68%	56%	56%	40%
3-Children learn best by exploring new things on their own	75%	70%	87%	59%	89%	100%
4-It is more important that the students are engaged	70%	68%	87%	60%	89%	100%
5-Children learn literacy/language best in small groups	73%	63%	77%	60%	100%	100%
6-If children answer their own questions, they will be more likely to understand	58%	58%	66%	53%	56%	90%
7-It is important for children to be free to talk	50%	49%	55%	40%	78%	70%
8-Children learn best from their peers	61%	42%	70%	44%	78%	80%
9-If children don't understand, they should try to find the answer independently before asking the teacher	19%	20%	26%	16%	22%	60%
10-Children can choose their own learning activities in the classroom	17%	17%	36%	18%	22%	30%

Note: The color scale is a gradient based on the proportion of agreement. Green = strong agreement; yellow = moderate agreement; orange = low agreement; red = poor agreement.

An overall "playfulness" score was further calculated for teachers and head teachers. This score reflects the total number of items on which the respondent selected the statement considered more supportive of LtP. As ***Table 18*** (below) shows, out of a total of 10 items, both teachers and head teachers agreed with the more playful response, on average, one-half of the time. Across all treatment groups, head teachers were marginally more supportive of LtP. The average attitudes score among teachers and head teachers in the control group is slightly lower than the other treatment groups but again, by a marginal amount.

Table 18. Mean Attitudes Scale Scores toward LtP—Teachers and Head Teachers

Respondent	Control	Indirect	Direct	GALOP
Teachers	4.935	5.745	5.632	5.348
Head teachers	5.52	5.78	5.96	5.67

A second, simplified survey about attitudes was given to caregivers, who were also asked to select their preference between two statements. The survey results are presented in ***Table 19.*** The caregivers tended to generally favor more play-based practices in the classroom (Items 1, 3 and 5), but their attitudes were less approving than those of the teachers. The caregivers were divided over the importance of children having "fun" in a friendly classroom environment or a well-disciplined classroom with serious students (Items 2 and 4). The area with the most limited support from caregivers was a willingness to relinquish a focus on discipline and order in the classroom—as shown by the relatively lower agreement with the statements regarding more play that appear in Item 6 (27%).

Table 19. Caregiver Attitudes toward LtP - Ghana

Respondent	Control	Indirect	Direct	GALOP
1- It is more important to teach creativity than academic subjects.	64%	61%	62%	64%
2- Teachers should create a friendly classroom environment.	52%	59%	50%	62%
3- Teachers should use play to promote learning in the classroom.	70%	76%	76%	79%
4- Children learn best when they are having fun.	48%	51%	52%	54%
5- I would be happy if teachers use play to support learning in the classroom.	73%	71%	70%	72%
6- It is important for children to be free to talk.	19%	31%	30%	29%

What are the enabling factors associated with implementing LtP pedagogical methods in classrooms? What are the barriers associated with implementing LtP pedagogical methods in classrooms?

Support to LtP

The sections below present information on the existing environment (at baseline) in which the P3 program will be implemented, broken out by four treatment groups. These sections are organized by stakeholder. At baseline, questions to meeting facilitators were limited to their attitudes toward play, so this group is not included in the findings presented below (they will be incorporated at midline). Additional findings on the supports received by teachers, coaches, trainers, and DEOs are presented in ***Annex A.***

Grade 1 Teachers

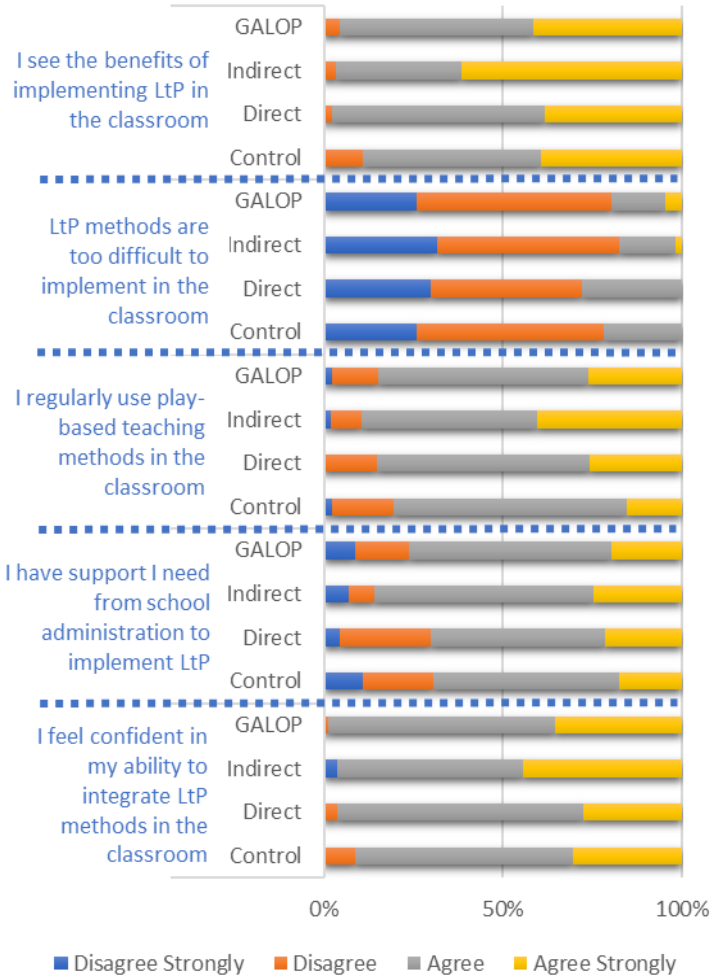
Across all treatment groups, the most common type of support teachers receive is training. The higher rate of training reported by direct implementation schools is to be expected, as the P3 program had already conducted training with these Grade 1 teachers when baseline data were collected. Teachers in direct implementation schools also reported more teacher meetings than those in indirect implementation schools and GALOP schools, who in turn reported more teacher meetings than teachers in control schools. Roughly one-half of teachers in direct implementation and GALOP schools said they received mentoring or coaching—more than teachers in control or indirect implementation schools **(*Table 20*).** Additional information on the frequency and content of these supports are presented in ***Annex A*.**

Table 20. Which of the Following Supports Do You Receive as a Teacher?

	Control	Indirect	Direct	GALOP
Training	67%	57%	95%	83%
Teacher meetings	43%	57%	70%	57%
Mentoring	37%	26%	51%	54%
Coaching (in-person)	39%	38%	51%	52%
Text messages or other remote advice or guidance	11%	4%	4%	17%
None of the above	13%	15%	0%	7%

Grade 1 teachers were asked questions aimed to measure their level of self-efficacy implementing LtP in the classroom **(Figure 7).** The vast majority of teachers agreed that they saw the benefits of implementing LtP and disagreed that LtP methods are too difficult to implement in the classroom. Across all treatment groups, nearly all teachers agreed or strongly agreed that they were confident in their ability to integrate LtP in the classroom, through there was slightly less agreement among teachers that they regularly used play-based teaching methods or had the support from school administration to do so.

Figure 7. Teachers' Self-Efficacy Using LtP, by Treatment Group



Caregivers of Grade 1 Children

To gain a fuller picture of children's exposure to play and play-based learning at home as well as in school, caregivers were asked about things they do to support their children at home. Additional findings on caregiver's involvement in school can be found in ***Annex A.*** As ***Table 21*** shows, when asked about activities they engaged in with their children over the previous 3 days, the highest percentage of caregivers in all treatment groups said they helped children with homework, while roughly one-third of caregivers said they told stories or cooked something. There were some surprising differences between treatment groups, particularly the relatively low percentage of caregivers with Grade 1 children in indirect implementation schools who stated they helped with homework or read storybooks and the relatively high percentage of caregivers in the control group who said they built something together. The same assessors visited schools across the four treatment groups, so it is unlikely an issue of inconsistent coding and rather a product of the open-ended nature of the question coupled with the small sample of respondents (~100 caregivers per group). These findings should be used to gain an understanding of play in students' home environment, and not as a point of comparison for midline findings.

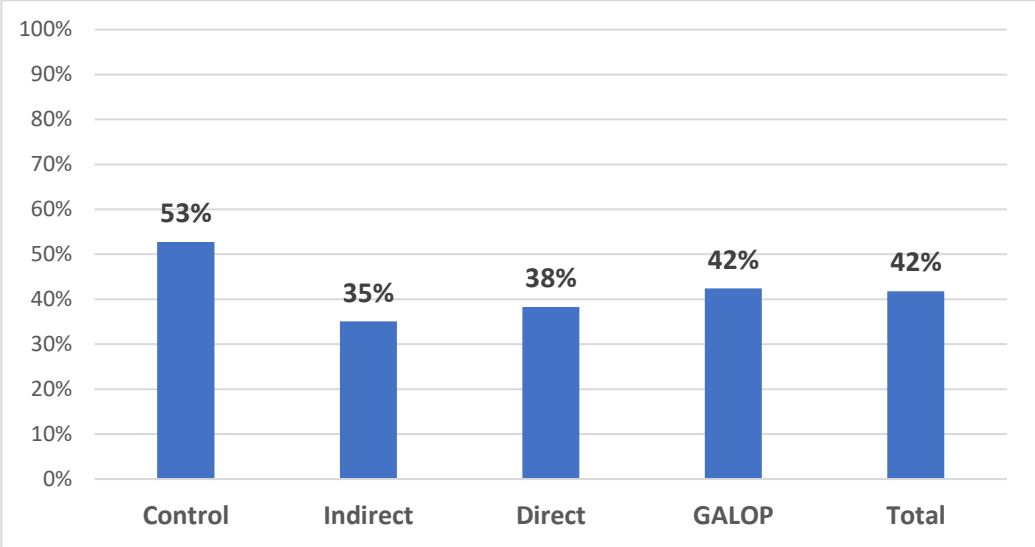
Table 21. Caregiver: In the Past Three Days, Have You or Another Caregiver Engaged in Any of These Activities with Your Child?

	Control	Indirect	Direct	GALOP
Helped with homework	63%	38%	46%	62%
Told stories	34%	26%	35%	33%
Cooked something together	34%	34%	43%	32%
Played a game inside the house	28%	20%	25%	25%

Table 21. Caregiver: In the Past Three Days, Have You or Another Caregiver Engaged in Any of These Activities with Your Child?				
	Control	Indirect	Direct	GALOP
Read storybooks	25%	8%	18%	18%
Played a game outside	18%	11%	12%	14%
Fixed something together	18%	12%	14%	11%
Made up a game together	3%	1%	4%	5%
Did a puzzle or board/card game	4%	3%	6%	3%
Built something together	17%	7%	8%	2%

Only 42% of caregivers replied that they had play materials at home. **Figure 8** shows a breakdown of this question by treatment group.

Figure 8. Caregiver: Do You Have Any Play Materials at Home?



Among caregivers who reported having play materials at home, the types of materials varied and again should be considered as descriptive findings and not used for comparison at midline. Across all treatment groups, the highest proportion of caregivers said they had sports equipment at home. Far fewer caregivers said they had more learning-centered play materials, such as craft supplies, puzzles, blocks, or books (***Table 22*).**

Table 22. Caregiver: What Types of Play Materials Do You Have at Home?				
	Control	Indirect	Direct	GALOP
Balls or sports equipment	48%	65%	55%	50%
Toys from the market	40%	32%	32%	20%
Books	15%	15%	25%	15%
Dolls	29%	24%	25%	10%
Homemade toys	17%	21%	23%	38%
Blocks or other building toys	23%	18%	16%	20%
Puzzles or board games	10%	3%	7%	5%
Craft supplies (paper scissors crayons paint coloring books)	17%	3%	5%	15%

Head Teachers

Questions for head teachers at baseline focused on how they identify and support struggling teachers. As ***Table 23*** shows, most head teachers rely on routine interactions with teachers, as well as student results (including their responses in class) to determine if their teachers are struggling. In response to a qualitative follow-up question, 11 head teachers also cited reviewing lesson plans and notes, and 13 head teachers said they use teacher attendance as an indicator of performance.

Table 23. Head Teacher: How Do You Know if a Teacher in Your School Is Not Performing as Expected?

	Control	Indirect	Direct	GALOP
Regular interaction/discussion with teachers	57%	73%	70%	65%
Student results	76%	65%	70%	78%
Supporting the teacher	2%	6%	19%	11%
Coaching data	2%	4%	11%	11%
Coaching feedback	7%	10%	9%	7%
I have no way of knowing	0	0	0	0

When asked what they do when a teacher is struggling, the vast majority of head teachers said they provide additional support themselves, including “sitting down with the teacher and hearing what the issue is” and “encouraging the teacher.” A higher percentage of head teachers in control, indirect, and direct schools also said they would observe the teacher more frequently, while slightly more head teachers in the GALOP schools said they would ask another teacher to provide mentorship (***Table 24*****).**

Table 24. Head Teacher: What Action Do You Take if a Teacher Is Not Performing as Expected?

	Control	Indirect	Direct	GALOP
Provide additional support to the teacher yourself	70%	77%	90%	85%
Observe the teacher (more frequently)	65%	52%	47%	35%
Request additional coaching support for the teacher	30%	33%	32%	37%
Ask another teacher to mentor the teacher	11%	17%	18%	26%
Provide a report to ministry official	4%	4%	2%	7%

While head teachers in control schools were mostly given a separate questionnaire (not about P3), there was still a small proportion of control school head teachers who indicated they had received P3 training and others who responded that P3 had been successful. The research team will work with the program, our in-country consultant, and assessors to refine these items and the head teacher interview administration to ensure more clarity as well as additional triangulation of responses.

Coaches

Coach interviews were conducted with School Improvement Support Officers (SISOs), formerly called Circuit Supervisors. Because SISOs serve schools from multiple treatment groups, findings from coach interviews are not disaggregated by treatment designation. Of the 47 coaches interviewed, however, 9 indicated they did not, at the time of the baseline

data collection, serve as a coach or mentor for P3 and therefore were not asked questions about the program specifically. Additional findings on the supports coaches say they provide to support LtP can be found in ***Annex A.***

Virtually all coaches (96%) reported feeling confident or very confident in their ability to support teachers with playful and interactive or student-directed strategies in the classroom. Based on their experiences, the majority of coaches felt that the biggest challenge to teachers' increasing their use of play-based approaches was a lack of the appropriate materials or support, followed by a lack of the necessary knowledge or skills. Interestingly, one-third of coaches also noted a lack of incentives for teachers to increase play-based approaches in the classroom ***Table 25.*** In response to a related qualitative follow-up question, five coaches also mentioned large class sizes and limited physical space to move within the classroom as a major barrier to LtP.

Table 25. Coaches: What Do You Think Are the Biggest Challenges to Increasing the Amount of Playful or Interactive Activities in the Classroom?

Challenge	Percent
Teachers do not have the appropriate materials/support	72%
Teachers do not have the necessary knowledge/skills	47%
There is no incentive to increase play	34%
Teachers do not support play in the classroom	11%
It is difficult to find time for play	6%
Head teachers do not support play in the classroom	2%

<u>Trainers</u>

District Training Officers (DTOs) serve as P3 trainers, so similar to the findings from coaches, findings from DTO interviews are presented in aggregate. Baseline interviews with DTOs focused on their confidence as trainers and their perspective on effective training approaches. Additional findings on trainings provided, training methods used, and suggested ways to improve trainings can be found in **Annex A.**

When asked (without prompting) what LtP-related content they train on under P3, trainers cited the multiple content areas listed in ***Table 26*** equally, with the exception of "using playful or interactive materials" – which seven of the nine DTOs mentioned. Trainers were then asked to select which of these content areas was most effective or useful to teachers; four of the nine DTOs identified content on how to use playful strategies in the classroom as the most effective for teachers.

Table 26. Trainers: What playful, Interactive, or Student-Directed Learning Content Areas Are Covered in Program Trainings?

	Element covered in program trainings	Element is the most effective training component
Use of playful/interactive materials	78%	22%
How to use playful strategies in the classroom	67%	44%
Development of playful/interactive materials	67%	11%
Working with parents/caregivers to support learning through play	67%	0%
None of the above	0%	22%

Trainers were also asked to rate their confidence in their ability to train teachers on playful, interactive, or student-directed strategies in the classroom, where 1 is not at all confident and 10 is very confident. All of the trainers interviewed ranked themselves a "7" or higher, and five of the nine trainers interviewed ranked themselves a "9" or "10."

District Education Officials

DEO interviews were conducted with District Education Directors (DEDs) in districts comprised of schools from all treatment groups. As such, findings from DEO interviews are also presented in aggregate. Baseline interviews with DEOs focused on the support they provide to schools. Additional findings on DEDs' roles and communication provided at the district and sub-district levels can be found in ***Annex A.***

When asked what their offices' role is in supporting P3, DEOs noted a myriad of responsibilities, most frequently citing monitoring, communicating expectations, and training (***Table 27)*.**

Table 27. DEOs: What Is Your (or Your Office's) Role in Supporting P3?

Role	Percent
Providing monitoring of program implementation	60%
Communicating program expectations to sub-district and school-level staff	50%
Training sub-district or school-level staff	50%
Ensuring that teachers participate in program activities	40%
Providing coaching and/or school supervision	40%
Overseeing or providing funds for mentoring and/or teacher meetings	10%
Working with parents/communities on program components	10%
No role in supporting the program	0%

DEOs were also asked questions about access to play-based learning supports in their district. When asked "Does the curriculum used in your primary schools include play-based learning activities and teaching methods?" all 10 DEOs said yes. All DEOs also said that they have received training or orientation on "learning through play" or play-based learning. When asked "Do primary teachers in your schools have materials for play-based learning in the classroom?" DEOs were divided in their response; four said "Yes", while six said "No."

What are the baseline levels of student literacy, numeracy and SEL among Grade 1 students in treatment and control schools?

Assessment of SEL

Observed Student Behavior During Assessments

During the SEL assessment, six student behaviors were observed by the assessor (***Table 28***), providing a measure of student confidence and positive emotions. There was a positive skew to the distribution in that a significant number, around 9%, of children received the maximum score on the scale. ***Table 29*** shows that the mean score was similar across treatment groups, with an overall mean of 7.8 (see **Annex A** for distributions of SEL scores).

Table 28. Means Scores for Six Observed Student Behaviors

Behavior	Mean	Max
Does the child greet you confidently?	1.09	2

Table 28. Means Scores for Six Observed Student Behaviors		
Behavior	**Mean**	**Max**
Does the child respond to questions confidently?	1.3	2
Does the child look you in the eye when listening to or responding to questions?	1.34	2
Actively attempts to engage interviewer	1.42	3
Shows pleasure in accomplishment and active task mastery	1.48	3
Child shows frequent positive emotions and behaviors	1.16	3
Total	7.78	15

Table 29. Observed Behavior by Treatment Group	
Heading	**Heading**
Control	7.71
Indirect	7.87
Direct	7.83
GALOP	7.69
Overall	7.78

Relationships

Table 30 shows that most children—around 60% to 80%—have a friend they can turn to for help in difficult situations or when learning and with whom they can share secrets. The distribution of the relationship score is skewed to the right, with 42% of children agreeing with all 5 questions. The mean value is 3.6. ***Table 31*** shows that the mean percentage of children responding positively to relationship questions was similar across treatment groups, with an overall mean of 72.7%. In an additional question, children named up to 20 friends by name with a median number of 4 friends.

Table 30. Relationship Item Scores	
Item	**% Agreement**
Do you talk to any of your friends when you feel sad? If Yes, whom do you talk to?	73.2
When you have an argument/fight with another child, do you go to any of your friends?	69.4
If you are working on something difficult, do you ask any of your friends for help?	76.9
Do you have a friend who helps you learn? If yes, who helps you learn?	80.9
Do you have a friend whom you tell your secrets to? If yes, who?	63.2

Table 31. Average Percent of Relationship Items	
Group	**Average Percent of "Yes" Responses**
Control	74.2
Indirect	75.3
Direct	72.8
GALOP	68.4
Overall	72.7

Conflict Resolution

Children listened to five scenarios in which there was potential conflict with another student and were asked what they would do in that situation. For each scenario their response was placed into one of eight categories (see ***Table 32***). The most common responses across all scenarios were disengagement from the situation or appealing to authority (i.e., asking the teacher to intervene). ***Table 33*** shows that the average number of children giving each type of response was similar across treatment groups.

Table 32. Percentage of Responses to Conflict Resolution Scenarios in Ghana

	Level 0		Level 1	Level 2			Level 3		
Scenario Number	Verbal aggression	Physical aggression	Disengage	Command	Appeal to Authority	Ask for Reasons	Influence, give reasons	Help the other child	Missing/No Response
1	9.2	38.4	27.4	4.5	13.4	1.3	2.8	0.7	2.4
2	3.1	5.5	37.0	7.8	35.2	1.1	4.6	2.4	3.2
3	2.9	4.0	41.9	5.2	19.7	1.4	8.9	12.1	3.9
4	6.0	13.0	29.5	9.2	28.2	2.5	4.5	4.2	3.0
5	10.4	28.4	18.6	8.6	26.6	1.8	2.8	0.2	2.6
Mean	6.3	17.9	30.9	7.1	24.6	1.6	4.7	3.9	3.0

Table 33. Average Percent of Responses to Conflict Resolution Scenarios by Treatment Group

	Level 0		Level 1	Level 2			Level 3		
Treatment Group	Verbal aggression	Physical aggression	Disengage	Command	Appeal to Authority	Ask for Reasons	Influence, give reasons	Help the other child	Missing/No Response
Control	6.4	18.8	28.6	6.6	24.5	1.8	5.2	4.3	3.8
Indirect	7.0	18.1	31.8	6.9	24.0	1.5	4.4	3.8	2.5
Direct	6.7	17.2	30.9	7.1	24.0	1.8	5.2	3.6	3.5
GALOP	5.0	17.5	32.1	7.6	26.1	1.4	4.1	4.2	2.1
Overall	**6.3**	**17.9**	**30.9**	**7.1**	**24.6**	**1.6**	**4.7**	**3.9**	**3.0**

Each response was classified as Level 0, Level 1, Level 2, or Level 3 and received a corresponding score of 0, 1, 2, or 3. The scores are normally distributed around the mean of 5.5

For each scenario, students were asked why the other child in the scenario behaved the way they did. The explanation was coded according to whether the motive of the other student was seen as being negative toward the student being interviewed (e.g., “she doesn’t like me”), neutral (e.g., “she was distracted”), or positive (e.g., “she was trying to help me”). ***Table 34*** shows the percentage of students with a negative attribute, referred to as ‘hostile attribution bias.’ ***Table 35*** shows that the average percentage was similar across treatment groups.

Table 34. Percentage Negative Attributions by Students

Scenario question	%
Why do you think the child took the toy away?	43.8
Why do you think the child wasn’t helping to sweep the room?	42.7
Why do you think the child is building the tower/building a different way from you?	33.6
Why do you think your classmate hasn't given you a turn yet?	43.7
Why do you think your classmate stood in front of you and took your place?	48.1
Total	42.4

Table 35. Percentage Negative Attributions by Treatment Group

Treatment Group	%
Control	45.5
Indirect	43.5
Direct	40.9
GALOP	39.9
Overall	42.4

Confidence

Students were read descriptions of two children—a child who shows confidence and a child who does not show confidence—and were asked which one they identified with. ***Tables 36*** and ***37*** shows the percentage of children identifying with the confident child across the seven questions. The average percentage of confident responses did not differ significantly across treatment groups.

Table 36. Percentage Students Identifying with Confident Person

Question	Average percentage identifying with confident person
X often feels confident (not afraid) when going into a room full of people.	53.2
When children in the class have a question, they get X to ask the teacher.	57.1
X can lead other children to sing a song.	54.3
X starts conversations with adults he/she doesn’t know well.	40.2
X likes to read stories out to the whole class.	57.8

X is trusted to look after younger children.	72.3
X shares his/her thoughts even if they are very different from others.	54.1
Total	55.6

Table 37. Percentage Students Identifying with Confident Person by Treatment Group

Treatment Group	%
Control	57.3
Indirect	54.6
Direct	55.9
GALOP	54.6
Overall	55.6

Creativity

Children were asked to imagine different uses for two everyday objects – a spoon and a pencil. On average, children in Ghana named 4.9 different ways these objects could be used. ***Table 38*** shows students' average creativity scores (the number of different uses they name), by treatment group, with a mean of 4.9. The distribution of creativity scores skews left, with most children suggesting very few (between one and five) different uses for the two items.

Table 38. Average Creativity Score by Treatment Group

Treatment Group	Average
Control	4.9
Indirect	4.6
Direct	4.9
GALOP	5.1
Overall	4.9

Student Literacy and Numeracy Outcomes

Shortened versions of traditional EGRA and EGMAs were administered to nearly 3,000 Grade 1 students across treatment groups in Ghana. The EGRA consisted of four subtasks administered in the local language of instruction for each school (with Asante Twi and Ga used in Greater Accra, Dagbani in the Northern region, and Ewe in Volta). The EGMA consisted of six subtasks (also administered in the school's LOI). Results are presented by treatment group (combined across languages).

As previously stated, the purpose of the EGRA and EGMA measures is simply to set a baseline against which program impact will be assessed at midline and endline. Therefore, it should not be concerning that performance was low across the majority of subtasks (as shown in ***Table 39***). This is to be expected given that students were assessed in Grade 1, at which point they are only just beginning to be introduced to these skills. For example, nearly all students (across all groups) scored zero on the invented word, oral reading passage, and reading comprehension subtasks, but performance was stronger in the letter sound identification task, which tests a lower order skill that is a starting point for the higher order skills. EGMA scores followed an even more conspicuous pattern of decreasing accuracy and

fluency scores (with increasing proportions of zero scores) from the first to the fourth subtask.

While scores were relatively consistent across treatment groups, there were some statistically significant differences, which will be accounted for at midline and endline.

Table 39. EGRA and EGMA Results for Grade 1 Students in Ghana, by Subtask and Treatment Group

Reading Subtasks	Measure	Control	Indirect	Direct	GALOP
		n= 677	*n*= 741	*n*= 859	*n*= 699
Letter Sounds	Correct letters per minute	7.4 [±.0]	5.4 [±.8]	6.3 [±.8]	5.7 [±.9]
	Percent correct of items attempted	22.8 [±2.2]	16.1 [±1.8]	19.2 [±1.7]	18 [±2]
	Zero scores	47.3 [±3.8]	59.5 [±3.5]	52.9 [±3.3]	58.2 [±3.7]
Invented Words	Correct invented words per minute	.4 [±.2]	.4 [±.2]	.3 [±.1]	.3 [±.2]
	Percent of correct items attempted	1.7 [±.7]	1.7 [±.7]	1.2 [±.5]	1.2 [±.5]
	Zero scores	95.4 [±1.6]	95.4 [±1.5]	96.2 [±1.3]	96.7 [±1.4]
Oral Reading Passage	Oral reading fluency (correct words per minute [cwpm])	.4 [±.2]	.3 [±.2]	.2 [±.1]	.2 [±.1]
	Zero scores	1.5 [±.7]	1.1 [±.5]	.9 [±.5]	.8 [±.4]
	Percentage correct	94.4 [±1.8]	95.5 [±1.5]	95.9 [±1.3]	95.7 [±1.5]
Reading Comprehension	Percentage correct	.2 [±.3]	.2 [±.2]	.1 [±.1]	.1 [±.1]
	Percent of correct items attempted	.2 [±.3]	.3 [±.3]	.1 [±.1]	.1 [±.1]
	Percent of students with >=80% comprehension	.1 [±.5]	0 [±.]	0 [±.]	0 [±.]
	Percent of students with zero scores	99.6 [±.6]	99.3 [±.7]	99.8 [±.4]	99.7 [±.5]
Quantity Discrimination	Percentage correct	48.1 [±2.2]	39.5 [±2.0]	41.9 [±1.8]	39.5 [±2.1]

Table 39. EGRA and EGMA Results for Grade 1 Students in Ghana, by Subtask and Treatment Group

Reading Subtasks	Measure	Control	Indirect	Direct	GALOP
		n= 677	*n*= 741	*n*= 859	*n*= 699
	Percent of correct items attempted	51.7 [±2.1]	43.7 [±1.9]	46.1 [±1.7]	43.8 [±2.1]
	Percent of students with zero scores	10.5 [±2.3]	14.6 [±2.6]	11.6 [±2.2]	17.3 [±2.8]
Missing Number	Percentage correct	16.9 [±1]	13.7 [±.9]	14.9 [±.8]	13.6 [±.0]
	Percent of correct items attempted	25.1 [±1.2]	21 [±1.1]	22.8 [±1.0]	20.7 [±1.2]
	Percent of students with zero scores	19.6 [±3]	28.7 [±3.3]	23.5 [±2.8]	30.5 [±3.4]
Basic Addition	Correct addition items per minute	2.9 [±.3]	2.1 [±.2]	2.3 [±.2]	2.5 [±.3]
	Percent correct of items attempted	21 [±2.1]	14.3 [±1.7]	15.5 [±1.6]	17.6 [±2.0]
	Percent of students with zero scores	48.2 [±3.8]	60.6 [±3.5]	58.9 [±3.3]	54.9 [±3.7]
Basic Subtraction	Correct subtraction items per minute	1 [±.2]	.9 [±.2]	1.2 [±.2]	1 [±.2]
	Percent correct of items attempted	7.5 [±1.4]	5.8 [±1.1]	7.4 [±1.1]	7.4 [±1.5]
	Percent of students with zero scores	79.3 [±3.1]	79.6 [±2.9]	75.1 [±2.9]	81.7 [±2.9]
Addition Level 2	Percentage correct	6.9 [±1.3]	3.8 [±.0]	6.5 [±1.2]	4.9 [±1.2]
	Percent correct of items attempted	6.9 [±1.3]	3.8 [±.0]	6.5 [±1.2]	4.9 [±1.2]

Table 39. EGRA and EGMA Results for Grade 1 Students in Ghana, by Subtask and Treatment Group

Reading Subtasks	Measure	Control	Indirect	Direct	GALOP
		n= 677	*n*= 741	*n*= 859	*n*= 699
	Percent of students with zero scores	83.2 [±2.8]	90 [±2.2]	86.4 [±2.3]	87.8 [±2.4]
Subtraction Level 2	Percentage correct	1.7 [±.6]	1 [±.5]	1.7 [±.6]	1.6 [±.6]
	Percent correct of items attempted	1.7 [±.6]	1 [±.5]	1.7 [±.6]	1.6 [±.6]
	Percent of students with zero scores	94.8 [±1.7]	96.9 [±1.3]	95.5 [±1.4]	95.4 [±1.6]

2.2 Tucheza Kujifunza (TuKu) - Kenya

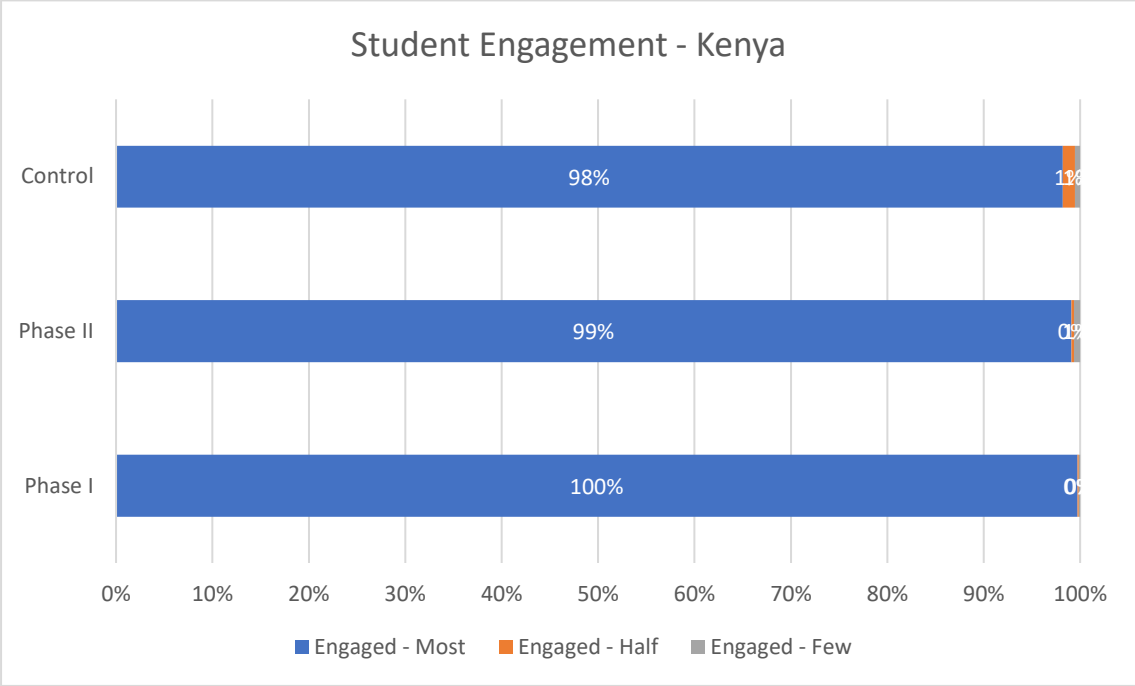
What is the baseline state of teachers' pedagogical practices in the classroom (including LtP) among Grade 1 teachers in treatment and control schools?

As in Ghana and Rwanda, classroom observations in Kenya were conducted in Grade 1 classrooms for both reading and mathematics lessons. This decision was made in order to ensure variation in pedagogical practices (without being limited by approaches used for a particular subject). Accordingly, results for reading and mathematics lessons are combined throughout this section. Results are averaged across classroom observations at the school level and are disaggregated by treatment group throughout this section, with observations occurring in 40 Phase I (innovation) implementation schools, 50 Phase II (network), implementation schools, and 50 control schools.

Timed Classroom Observations

Across all observed lessons, assessors in Kenya reported that most students were engaged for virtually all of observed class time in all schools (**Figure**). This strong level of engagement is a promising starting point for the program to continue to improve the depth and breadth of student engagement through implementation.

Figure 9. Student Engagement During Lessons (Percentage of Observed Class Time), by Treatment Group in Kenya



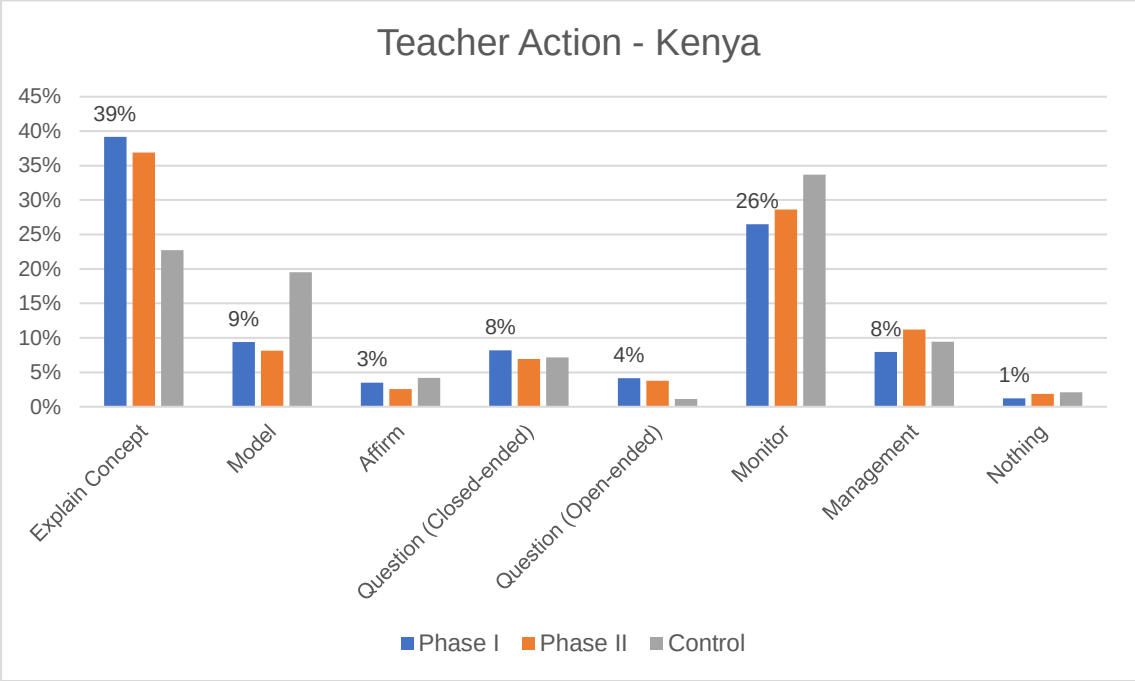
Assessors measured the classroom climate during observed lessons, focusing on four main criteria: perseverance (e.g., recognition of effort or encouragement of goal setting), positive verbal/physical (e.g., use of positive language of physical evidence of joy), negative verbal/physical (e.g., use of negative language or physical punishment), and ordinary (e.g., tone is neither noticeably positive nor negative). Assessors found an ordinary climate for more than 95% of observed class time across groups (***Table 40***). There was no negative climate but also nearly no perseverance and only limited positive climate.

Table 40. Observed Classroom Climate (Percentage of Observed Class Time), by Treatment Group in Kenya

	Phase I	Phase II	Control
Ordinary	95%	95%	98%
Positive verbal/physical	4%	4%	2%
Negative verbal/physical	0%	0%	0%
Perseverance	0%	1%	0%
Number of schools	40	50	50

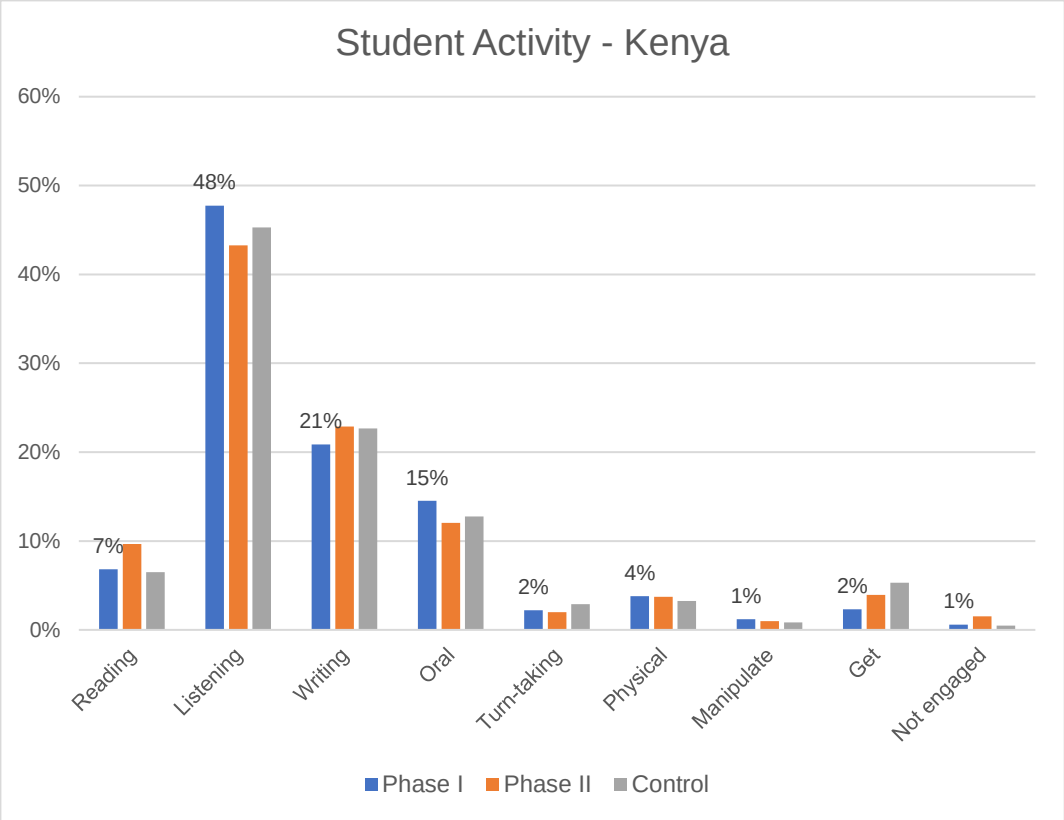
During the observed lessons in Kenya, teachers spent the largest proportions of time explaining concepts and monitoring students' understanding (or tracking participation in an activity) (**Figure 10**). A significantly smaller amount of time was spent modeling activities. More interactive approaches were used even less frequently in the classroom, which is a viable area of improvement for the program. For example, open-ended questioning was rarely used (approximately 4% of class time across groups)—though closed-ended questioning was more prevalent (approximately 8% across groups). Teacher affirmations or recognition of student input made up a small proportion of time, while management or an unengaged teacher (i.e., nothing) accounted for around 10% of class time.

Figure 10. Teacher Actions During Lessons (Percentage of Observed Class Time), by Treatment Group in Kenya



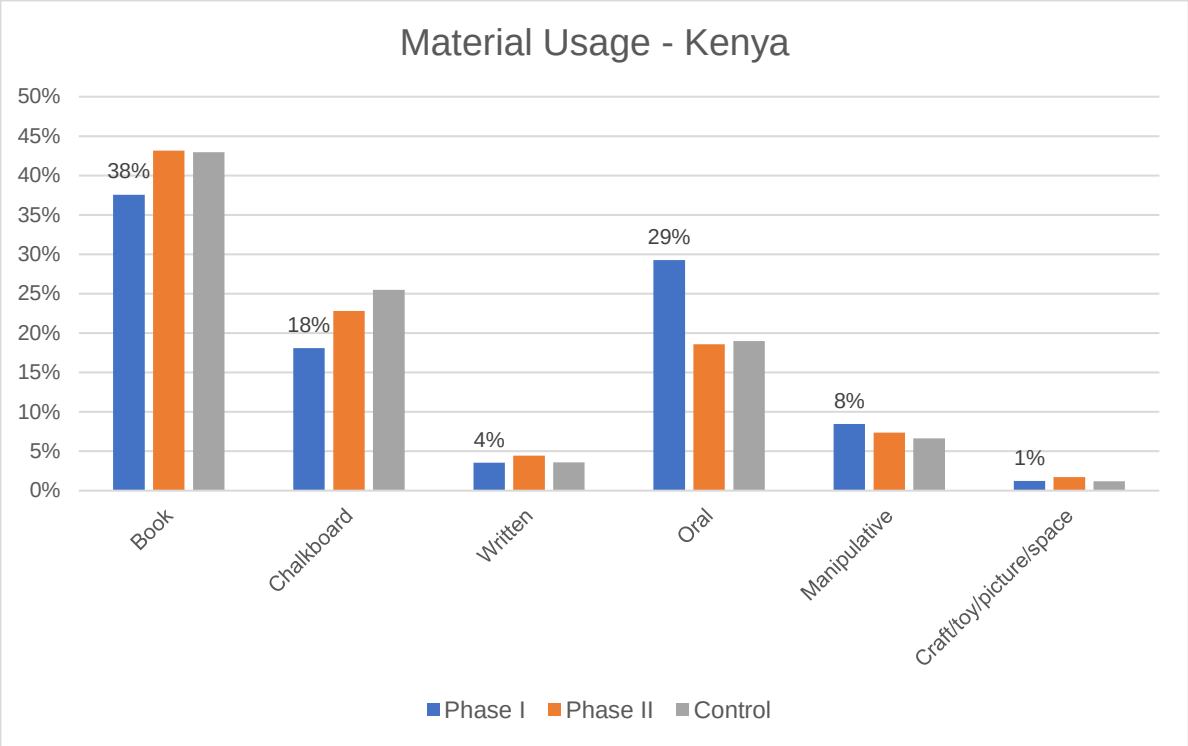
Students in observed classrooms spent nearly half of their time listening to teachers (**Figure 11**). Interestingly, even though these are Grade 1 classrooms, approximately one-fifth of time was spent writing, while less than 10% was dedicated to reading. Less than 10% of class time overall was also dedicated to students either discussing/responding to one another (turn-taking), having a physical response, or manipulating something with their hands. Students spent less than 5% of class getting something to work on, waiting for the teacher, or disengaged (i.e., 'get' and 'not engaged' in Figure 14).

Figure 11. Student Activity During Lessons (Percentage of Observed Class Time), by Treatment Group in Kenya



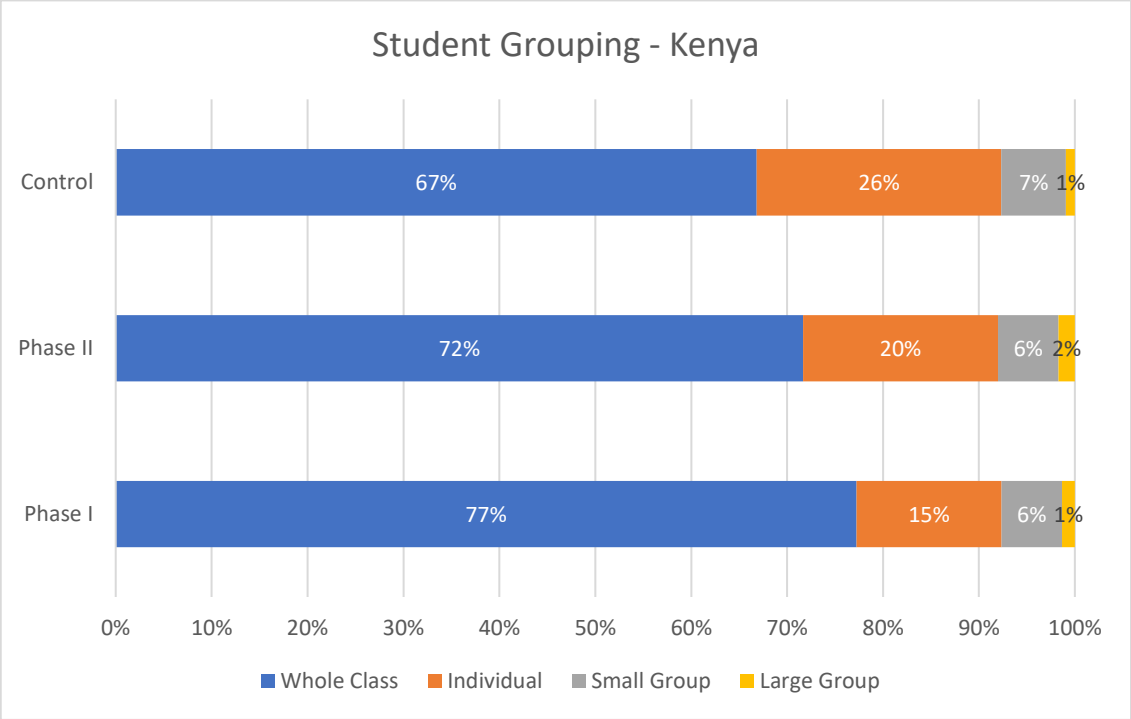
Materials provide a great opportunity for programs to increase the use of playful practices in the classroom. Despite the fact that students had limited reading activities the classroom, books were the most prominently used material during lessons (**Figure 12**). The use of chalkboards and oral language were also common. This left little time for the use of other written materials (such as letter/number cards or pocket charts), manipulatives, and alternative materials (e.g., crafts, toys, illustrations, space, etc.).

Figure 12. Use of Materials During Lessons (Percentage of Observed Class Time), by Treatment Group in Kenya



Although the majority of teachers in Kenya said that children learn language/literacy best through group work and that student exploration is important (on the attitude survey), these are still areas for improvement in actual instruction. The results of the classroom observations showed that two-thirds to three-quarters of class time was spent on whole class instruction, with 15% to 26% of the time used for individual work (**Figure 13**). Less than 10% of the total class time was spent with students working in large or small groups.

Figure 13. Student Grouping During Lessons (Percentage of Observed Class Time), by Treatment Group in Kenya



Note: "Corners" omitted from the table because <1% of time was spent using learning corners in observed classrooms.

There was also evidence of limited student agency found in the classroom observations in Kenya, with nearly all of class time dedicated to activities that were initiated and directed by teachers (***Table 41***). Phase II schools had the higher proportion of time spent on student-directed or student-initiated activities (7%), while the other groups had 3% or less.

Table 41. Student Agency (Percentage of Observed Class Time), by Treatment Group in Kenya

	Phase I	Phase II	Control
Teacher directed	97%	93%	98%
Teacher initiated; Student directed	1%	5%	1%
Student initiated	2%	2%	1%

Post-Classroom Observation Indicators

At the end of each timed classroom observation, assessors completed a checklist to determine whether a range of activities occurred (at least once) during the lesson. These indicators are divided into six constructs: 1) integrated pedagogies, 2) application (use of games), 3) responsiveness (to student behaviors), 4) resources (availability and use), 5) efficiency (maximized instructional time), and 6) student-directed (ownership or autonomy).

As shown in ***Table 42*** there was significant variation in the prevalence of activities occurring throughout the observed lessons (with high prevalence in green; moderate prevalence in yellow; low prevalence in red). In terms of constructs, teachers tended to perform best on responsiveness and efficiency indicators. Integrated pedagogies, as well as resources, included some high prevalence and some low prevalence items, while application tended to

be low (aside from using songs or energizers in the classroom). The most limited use was with regard to student-directed opportunities. Overall, basic best practices (e.g., student monitoring and support, material preparation, good instructional pace, making connections and sharing ideas, inclusionary practices) were used throughout lessons but there is significant room for improvement in including more playful pedagogies.

Table 42. Post-observation checklist (percentage of observed lessons), by treatment group in Kenya

	Phase I	Phase II	Control
Integrated Pedagogies			
The teacher asked questions or raised topics that got students to respond to each other, not to her (i.e., active).	79%	75%	58%
Student(s) persisted with the activity after the rest of the class has started to do something else (i.e., engage).	19%	10%	17%
Student(s) started the task again or adjusted when they thought of a better idea/approach (i.e., iterative).	16%	24%	18%
Student(s) had a chance to make connections between the new experiences/content to their prior knowledge or experiences (i.e., meaningful).	63%	49%	59%
Students problem solved or had a chance to think critically or reflect about a topic.	53%	42%	69%
Application: Students have generalized games			
Songs or energizers to start or divide activities.	76%	65%	67%
Students played games with obvious rules.	26%	34%	22%
Students had physical play.	14%	18%	17%
Students used objects to represent something else.	23%	18%	12%
Students had symbolic play.	9%	14%	6%
Responsiveness: The teacher adapts to student behaviors			
The teacher gave students hints, suggestions, or redirection when they needed it.	81%	77%	96%
The teacher circulated, leaned in, redirected, to support students as needed.	86%	87%	84%
The teacher asked the class or a group for a solution to a student question or problem.	65%	61%	43%
The teacher included students who did not volunteer to answer (i.e., Inclusion).	74%	72%	75%
Resources: Materials that are needed for implementation are used			
Students used a resource that was dedicated for their use (e.g., book, workbook, slate, tablet).	96%	96%	92%
Students used something to create something else (e.g., art supplies, items in nature, objects repurposed).	11%	8%	9%
Students used a resource to solve a problem (e.g., manipulatives, math charts, dictionary).	60%	52%	35%

Instruction used an item designed for interaction between teacher and students or between students (e.g., read aloud, pocket chart, counters, puzzle).	66%	55%	49%
Students used physical space for learning activities (e.g., counting large steps for measurement, sit in a circle for sharing).	18%	29%	16%
Students used stations/corner work/learning centers.	0%	11%	6%
Efficiency: Instructional time is maximized			
The teacher had materials ready at the start of an activity or the class.	98%	94%	93%
The pace of the instruction was lively.	96%	92%	87%
Student-Directed (Ownership / Autonomy)			
Students developed and shared ideas with either classmates or the teacher.	48%	31%	29%
Students had a chance to select from choices (state of play: exploring).	15%	11%	1%
Students chose collaborators or roles.	11%	12%	2%
Students made or changed rules for an activity.	4%	0%	1%
Students chose how long to do an activity (state of play: owning).	4%	1%	2%
Students set goals or helped to plan how to use time.	1%	0%	0%
Students chose when to move around.	1%	2%	4%

In addition to the item level estimates in Table 42, construct level scores were calculated (***Table 43***). These scores will be used as the basis for measuring change/impact across treatment groups, over time. The scores are calculated as simple averages of the number of items occurring per construct for each treatment group. At baseline, there was only one statistically significant difference between Phase II treatment schools and their controls (i.e., Resources). These differences will be accounted for in midline/endline impact estimates.

Table 43. Post-Observation Checklist Scores, by Treatment Group in Kenya

	Phase I	Phase II	Control
Integrated Pedagogies (0–5)	2.3	2.0	2.2
Application: Students have generalized games (0–5)	1.5	1.5	1.2
Responsiveness: The teacher adapts to student behaviors (0–4)	3.1	3.0	3.0
Resources: Materials that are needed for implementation are used (0–6)	2.5	2.5*	2.1
Efficiency: Instructional time is maximized (0–2)	1.9	1.9	1.8
Student-Directed (Ownership/Autonomy) (0–7)	0.8	0.6	0.4
Overall (0–29)	12.1	11.4	10.7

*$p<0.05$ (comparison between Phase II treatment and control)

The post-observation component of the data collection also included a classroom inventory. The inventory was focused on determining the availability of basic materials in each classroom. For simplicity, the results displayed in ***Table 44*** are for reading lessons only. Overall, the availability of materials in the observed classrooms in Kenya was high. More than 90% of classrooms had teacher's guides and textbooks available (although only about

three-quarters of individual students had textbooks). Fewer classrooms had workbooks (58%), teaching aids (50%), or supplementary readers (20%).

Table 44. Post-Observation Reading Classroom Inventory (Availability in Percentage of Classrooms), by Treatment Group in Kenya

	Phase I	Phase II	Control	Overall
Teacher's guide	98%	90%	100%	96%
Teaching aids	65%	48%	40%	50%
Textbooks (present in classroom)	95%	96%	96%	96%
Textbook Ratio (percentage of students with textbook)	73%	69%	78%	73%
Workbooks	60%	70%	44%	58%
Supplementary readers	15%	6%	38%	20%
N	40	50	50	140

Learning Environment (Student Survey)

Students were asked 10 questions about the learning environment (see ***Table 45***) with each response giving a score of 0,1, or 2. Scores were normally distributed around the mean of 11.8. ***Table 46*** shows that the learning environment score was similar across treatment groups.

Table 45. Responses to Learner Environment Questions by Students in Kenya

Item	No/Never	Sometimes	Yes/Often
Does your teacher praise you when you give the right answer?	12	38	49
Can you talk to at least one teacher in the school about things that are bothering you?	26	12	63
Last week, did children in your class ask the teacher questions in lessons?	40	31	29
Last week, did some children in you class ask the teachers for help with their classwork?	31	38	31
Do your teachers encourage you to try out new things in class?	19	36	44
Do your teachers get angry or yell when you get the answer wrong?	45	37	18
Do your teachers help those who do not understand?	19	35	45
In the last week, did an adult in your school scream or yell at a child?	66	21	12
Can students share their thoughts in front of the class?	59	22	18
In the last week, did an adult in your school hit a child or use harsh punishments?	46	34	20

Table 46. Learning Environment Score by Treatment Status

Treatment Status	Mean Score
Control	11.8
Phase 1	12.0
Phase 2	11.7
Total	11.8

What are the baseline perceptions of LtP among school, district, and community-level stakeholders?

Attitudes toward LtP

Attitudes towards LtP were similar across stakeholders, with trainers and DEOs indicating slightly lower support for LtP practices than head teachers, teachers, coaches, and teacher meeting facilitators (see ***Table 47***).

Nearly all of the interview participants strongly supported play as an important component of learning for children in primary school (Item 1). In addition, most interview participants highly favored work in small groups (Items 2 and 6) and prioritized student engagement and exploration (Items 3 through 5). The areas that need the most improvement are the views on student self-directed work, as shown by the low agreement in Items 9 and 10.

Table 47. Stakeholder Attitudes toward LtP - Kenya

Items	Head teacher	Teacher	Teacher meeting facilitator	Trainer	DEO
1—Play is an important part of learning for children in primary school.	94%	95%	94%	86%	100%
2—Children learn literacy/language best in small groups.	90%	86%	87%	57%	64%
3—It is more important that the students are engaged.	83%	86%	80%	57%	91%
4—Children learn best by exploring new things on their own.	91%	74%	86%	71%	73%
5—If children answer their own questions, they will be more likely to understand.	80%	70%	76%	57%	64%
6—Children learn best from their peers.	77%	55%	68%	50%	64%
7—It is important for children to be free to talk.	64%	55%	62%	71%	73%
8—Parents and caregivers welcome more play-based approaches in the classroom.	56%	55%	52%	71%	46%
9—If children do not understand, then they should try to find the answer independently before asking the teacher.	50%	37%	49%	14%	27%
10—Children can choose their own learning activities in the classroom.	23%	23%	22%	14%	27%
Number of Interviews Conducted	**133**	**139**	**136**	**14**	**11**

Note: The color scale is a gradient based on the proportion of agreement. Green = strong agreement; yellow = moderate agreement; orange = low agreement; red = poor agreement.

An overall "playfulness" score was further calculated for teachers and head teachers. This score reflects the total number of items on which the respondent selected the statement considered more supportive of LtP. As ***Table 48*** (below) shows, out of a total of 10 items, both teachers and head teachers agreed with the more playful response, on average, roughly two-thirds of the time. Across all treatment groups, head teachers were marginally more supportive of LtP than teachers. The average attitudes score among teachers and head teachers is consistent across treatment groups.

Table 48. Mean Attitudes Scale Scores toward LtP—Teachers and Head Teachers

Respondent	Control	Phase 1	Phase 2
Teachers	6.1	6.7	6.4

Table 48. Mean Attitudes Scale Scores toward LtP—Teachers and Head Teachers			
Respondent	**Control**	**Phase 1**	**Phase 2**
Headteachers	7.1	6.9	7.1

A second, simplified survey about attitudes was given to caregivers, who were also asked to select their preference between two statements. The survey results are presented in ***Table 49.*** The caregivers tended to generally favor more play-based practices in the classroom (Items 1, 3 and 5), but their attitudes were less approving than those of the teachers. Particularly in the control group, caregivers were divided over the importance of children having "fun" in a friendly classroom environment over a well-disciplined classroom with serious students (Items 2 and 4). The area with the most limited support from caregivers was a willingness to relinquish a focus on discipline and order in the classroom—as shown by the relatively lower agreement with the statements regarding more play that appear in Item 6 (27%). Across all attitude questions, caregivers in Phase 2 schools were slightly more "pro-play" than Phase 1 caregivers, and even more so than caregivers of students in control schools.

Table 49. Caregiver Attitudes toward LtP—Kenya			
Respondent	**Control**	**Phase 1**	**Phase 2**
1- It is more important to teach creativity than academic subjects.	60%	69%	74%
2- Teachers should create a friendly classroom environment.	40%	66%	51%
3- Teachers should use play to promote learning in the classroom.	53%	64%	65%
4- Children learn best when they are having fun.	47%	59%	65%
5- I would be happy if teachers use play to support learning in the classroom.	64%	69%	67%
6- It is important for children to be free to talk.	27%	31%	48%

What are the enabling factors associated with implementing LtP pedagogical methods in classrooms? What are the barriers associated with implementing LtP pedagogical methods in classrooms?

Support to LtP

The sections below present information on the existing environment in which the TuKu program will be implemented, broken out by four treatment groups. These sections are organized by stakeholder. At baseline, questions to meeting facilitators were limited to their attitudes toward play, so this group is not included in the enabling environment findings presented below (they will be incorporated at midline). Additional findings on the supports received by teachers, trainers, and DEOs are presented in ***Annex A*****.**

Grade 1 Teachers

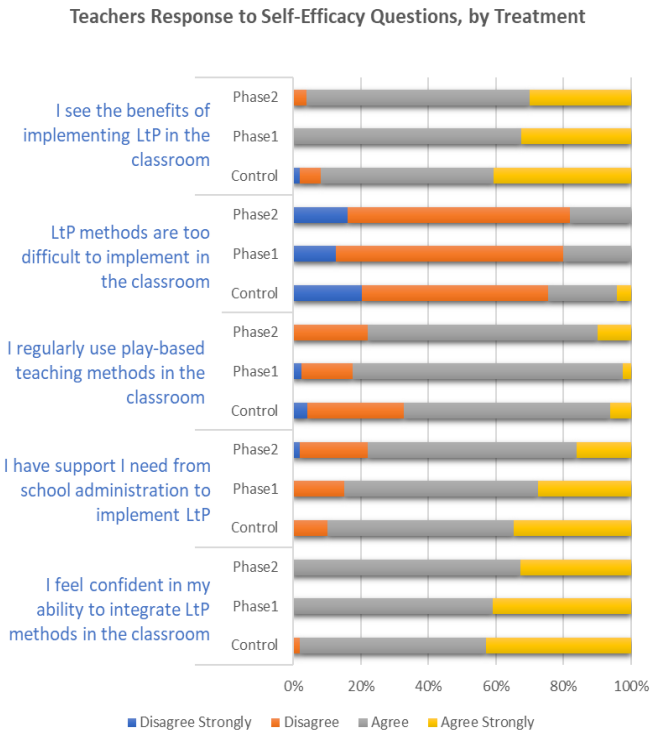
Across all treatment groups, teachers most commonly received support through teacher meetings or training. **(*Table 50*).** Additional information on the frequency and content of these supports is presented in ***Annex A*****.**

Table 50. Which of the Following Supports Do You Receive as a Teacher?

	Control	Phase 1	Phase 2
Teacher meetings	63%	55%	60%
Training	67%	68%	46%
None of the above	37%	48%	38%
Mentoring	22%	18%	36%
Coaching (in-person)	31%	23%	30%
Text message or other remote advice or guidance	24%	20%	24%

Grade 1 teachers were asked questions aimed to measure their level of self-efficacy implementing LtP in the classroom **(Figure).** A large majority of teachers agreed that they see the benefits of implementing LtP and disagreed that LtP methods are too difficult to implement in the classroom. To triangulate this, virtually all teachers agreed or strongly agreed that they are confident in their ability to integrate LtP in the classroom, though there was less agreement among teachers that they regularly use play-based teaching methods and slightly less agreement that they have the support from school administration to do so. Teacher's self-efficacy responses were similar across the three treatment groups.

Figure 14. Teachers' Self-Efficacy Using LtP, by Treatment Group



Caregivers of Grade 1 Children

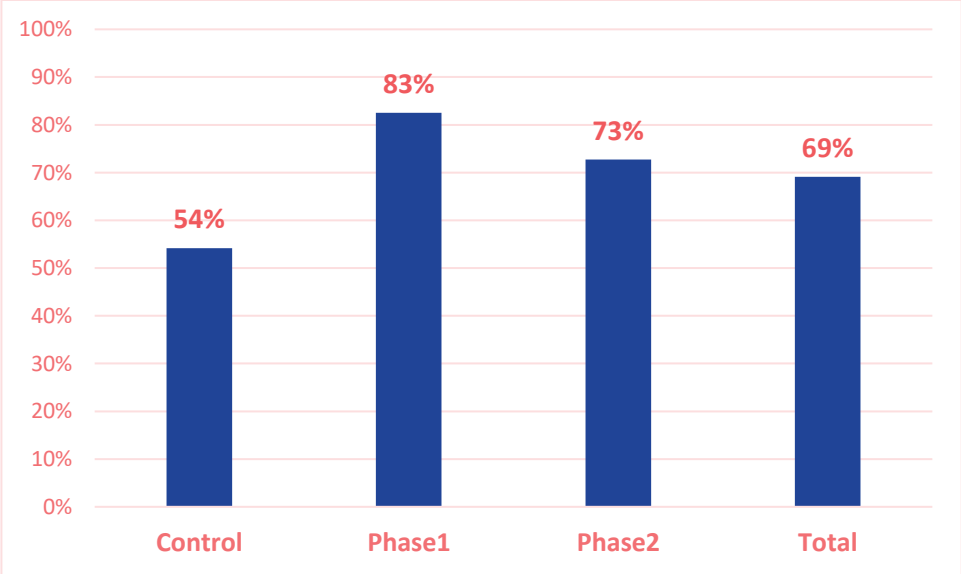
To gain a fuller picture of children's exposure to play and play-based learning at home as well as in school, caregivers were asked about things they do to support their children at home. Additional findings on caregiver's involvement in school can be found in ***Annex A.*** As ***Table 51*** shows, when asked about activities they engaged in with their children over the previous 3 days, the highest percentage of caregivers in all treatment groups said they helped children with homework, while roughly one-half of caregivers said they told stories or cooked something together. Almost no caregivers mentioned making up a game with their children, though one-third to one-half of caregivers mentioned playing games. Findings were consistent across treatment groups, although these findings should be used to gain an understanding of play in students' home environment, and not as a point of comparison for midline findings.

Table 51. Caregiver: In the Past 3 Days, Have You or Another Caregiver Engaged in Any of These Activities with Your Child?

	Control	Phase 1	Phase 2
Helped with homework	72%	90%	84%
Cooked something together	52%	48%	52%
Told stories	53%	51%	47%
Played a game inside the house	42%	42%	45%
Read storybooks	32%	40%	38%
Built something together	49%	27%	38%
Played a game outside	27%	27%	32%
Fixed something together	28%	10%	17%
Other	5%	9%	11%
Made up a game together	4%	1%	6%

Caregivers with students in Phase 1 and 2 schools reported they have play materials at home at a higher rate than caregivers with students in control schools (**Figure).**

Figure 15. Caregiver: Do You Have Any Play Materials at Home?



Among caregivers who reported having play materials at home, the most common responses were homemade toys, balls or sports equipment and dolls. Caregivers in the Phase 2 treatment group reported having books at home more than did control parents, but this may be in part due to the open-ended nature of the question (***Table 52*).**

Table 52. Caregiver: What Types of Play Materials Do You Have at Home?

	Control	Phase 1	Phase 2
Homemade toys	27%	38%	46%
Dolls	19%	20%	31%
Balls or sports equipment	29%	43%	24%
Craft supplies (paper scissors, crayons paint coloring books)	12%	14%	22%
Toys from the market	15%	20%	18%
Books	4%	11%	17%

Blocks or other building toys	17%	15%	11%
Puzzles or board games	1%	14%	4%

Head Teachers

Questions for head teachers at baseline focused on how they identify and support struggling teachers. As ***Table 53*** shows, most head teachers rely on student results, followed by routine interactions with teachers to determine whether their teachers are struggling.

Table 53. Head Teacher: How Do You Know if a Teacher in Your School Is Not Performing as Expected?

	Control	Phase 1	Phase 2
Student results	82%	85%	70%
Regular interaction/discussion with teachers	60%	55%	52%
Supporting the teacher	34%	25%	22%
Other (specify)	46%	28%	18%
Coaching feedback	6%	23%	14%
Coaching data	14%	20%	6%

When asked what they do to when a teacher is struggling, the majority of head teachers said they provide additional support themselves, while roughly half said they would request additional coaching for the teacher—an interesting response, given that head teachers said they relied on coaching feedback or data very little to identify struggling teachers. (***Table 54*****).**

Table 54. Head Teacher: What Action Do You Take if a Teacher Is Not Performing as Expected?

	Control	Phase 1	Phase 2
Provide additional support to the teacher yourself	86%	80%	74%
Request additional coaching support for the teacher	46%	50%	46%
Ask another teacher to mentor the teacher	40%	50%	36%
Observe the teacher (more frequently)	50%	30%	32%
Other	28%	10%	12%

Trainers

Baseline interviews with TuKu program trainers—selected from local civil society organizations—focused on their confidence as trainers and their perspective on effective training approaches. Additional findings on trainings provided, training methods used, and suggested ways to improve trainings can be found in ***Annex A*****.**

When asked what LtP-related content they are prepared to train on under TuKu, most trainers pointed to using playful strategies in the classroom, followed by working with caregivers to support LtP. Over one-half of the trainers also mentioned the development or use of playful materials (***Table 55*****).** Trainers were then asked to select which of these content areas was most effective or useful to teachers. Guidance on using playful strategies and the development of playful materials were identified most frequently (30%) as the most effective component of the trainings.

Table 55. Trainers: What Playful, Interactive, or Student-Directed Learning Content Areas Are Covered in Program Trainings?

	Element covered in program trainings	Element is the most effective training component
How to use playful strategies in the classroom	80%	30%
Working with parents/caregivers to support LtP	70%	10%
Use of playful/interactive materials	60%	10%
Development of playful/interactive materials	60%	30%
None of the above	0%	20%

Trainers were also asked to rate their confidence in their ability to train teachers on playful, interactive, or student-directed strategies in the classroom, where 1 is not at all confident and 10 is very confident. All of the trainers interviewed ranked themselves a "6" or higher, and 8 of the 14 trainers interviewed ranked themselves a "9" or "10."

District Education Officials

DEO interviews were conducted with 11 Sub-County Directors and are presented in aggregate. Baseline interviews with DEOs focused on the support they provide to schools. Additional findings on Sub-County Director's roles and communication provided at the district and sub-district levels can be found in ***Annex A*.**

When asked what their office's role is in supporting TuKu, DEOs noted a myriad of responsibilities, most frequently citing monitoring, communicating expectations, and training (***Table 56*).**

Table 56. DEOs: What Is Your (or Your Office's) Role in Supporting TuKu?

Role	Percent
Communicating program expectations to sub-district and school-level staff	27%
Providing monitoring of program implementation	27%
Ensuring that teachers participate in program activities	18%
Providing coaching and/or school supervision	18%
Training sub-district or school-level staff	9%
Overseeing or providing funds for mentoring and/or teacher meetings	0%
Working with parents/communities on program components	0%
No role in supporting the program	0%

Sub-county directors were also asked questions about access to play-based learning supports in their district. When asked "Does the curriculum used in your primary schools include play-based learning activities and teaching methods?" 9 of the 11 Sub-County Directors said yes. One-half of the Sub-County Directors also said that they received training or orientation on "learning through play" or play-based learning. When asked "Do primary teachers in your schools have materials for play-based learning in the classroom?" 7 said "Yes", while 4 said "No." Three of the Sub-County Directors who said "No" also said they did not receive training on LtP.

What are the baseline levels of student literacy, numeracy, and SEL among Grade 1 students in treatment and control schools?

Assessment of SEL

Observed Student Behavior During Assessments

During the SEL assessment, six student behaviors were observed by the assessor (see ***Table 57***), providing a measure of student confidence and positive emotions. **Th**e total score was well distributed around the mean of 9.3. There was a positive skew to the distribution in that a significant number—around 13%—of children received the maximum score on the scale. ***Table 58*** shows that the mean score was similar across treatment groups (see **Annex A** for distributions of SEL scores).

Table 57. Student Behaviors Observed

Behavior	Mean	Max
Does the child greet you confidently?	1.41	2
Does the child respond to questions confidently?	1.39	2
Does the child look you in the eye when listening to or responding to questions?	1.46	2
Actively attempts to engage interviewer	1.73	3
Shows pleasure in accomplishment and active task mastery	1.79	3
Child shows frequent positive emotions and behaviors	1.57	3
Total	9.35	15

Table 58. Observed Behavior Score by Treatment Group

Treatment Group	Average Score
Control	9.2
Phase 1	9.6
Phase 2	9.2
Overall	9.3

Relationships

Table 59 shows that most children—around 70% to 90%—have a friend they can turn to for help in difficult situations or when learning and with whom they can share secrets. The distribution of the relationship score is skewed to the right, with 47% of the children agreeing with all five questions. The mean was 4.0. ***Table 60*** shows that the mean percentage of children responding positively to relationship questions was similar across treatment groups, with an overall mean of 79.7%. In an additional question, children named up to 15 friends by name with a median number of 4 friends.

Table 59. Relationship Item Scores in Kenya

Item	Percentage in agreement
Do you talk to any of your friends when you feel sad? If Yes, whom do you talk to?	76.6
When you have an argument/fight with another child, do you go to any of your friends?	76.8
If you are working on something difficult, do you ask any of your friends for help?	86.6
Do you have a friend who helps you learn? If yes, who helps you learn?	84.8

Do you have a friend whom you tell your secrets to? If yes, who?	73.9

Table 60. Average Percent of Relationship Items

Treatment Group	Average Percent of "Yes" responses
Control	84.8
Phase 1	76.2
Phase 2	77.5
Overall	79.7

Conflict Resolution

Children listened to five scenarios in which there was potential conflict with another student and were asked what they would do in that situation. For each scenario their response was placed into one of eight categories (see ***Table 61***). The most common responses across all scenarios were disengagement from the situation or appealing to authority (i.e., asking the teacher to intervene). ***Table 62*** shows that the average number of children giving each type of response was similar across treatment groups.

Table 61. Percentage of Responses to Conflict Resolution Scenarios in Kenya

	Level 0		Level 1	Level 2		Level 3			
Scenario number	Verbal aggression	Physical aggression	Disengage	Command	Appeal to authority	Ask for reasons	Influence, give reasons	Help the other child	Missing/No response
1	0.9	22.7	21.5	9.0	28.4	3.7	3.3	3.5	7.1
2	0.2	1.5	33.4	5.0	41.9	1.5	3.7	4.7	8.1
3	0.3	1.7	34.3	3.5	26.9	2.0	3.9	12.6	14.9
4	0.7	5.9	28.8	6.7	30.0	3.5	4.1	6.5	13.8
5	2.5	17.0	17.8	9.3	42.2	1.3	1.7	0.6	7.7
Total	0.9	9.8	27.2	6.7	33.9	2.4	3.3	5.6	10.3

Table 62. Average Percent of Responses Conflict Resolution Scenarios by Treatment Group

	Level 0		Level 1	Level 2		Level 3			
Treat-ment group	Verbal aggression	Physical aggres-sion	Disengage	Command	Appeal to authority	Ask for reasons	Influence, give reasons	Help the other child	Missing/No response
Control	0.7	6.1	29.9	7.4	36.3	1.2	1.7	7.0	9.8
Phase 1	1.1	11.7	26.7	7.6	32.3	2.9	3.2	4.8	9.6
Phase 2	1.0	8.7	26.5	6.0	35.0	2.4	4.3	5.3	10.7
Overall	1.0	9.0	27.4	5.7	34.6	2.2	3.6	5.8	10.7

Each response was classified as Level 0, Level 1, Level 2, or Level 3 and received a corresponding score of 0, 1, 2, or 3. The distribution of the score, totaled across the five scenarios. The scores are normally distributed around the mean of 6.1.

For each scenario, students were asked why the other child in the scenario behaved the way they did. The explanation was coded according to whether the motive of the other student was seen as being negative toward the student being interviewed (e.g., “she doesn’t like me”), neutral (e.g., “she was distracted”), or positive (e.g., “she was trying to help me”). ***Table 63*** shows the percentage of students with a negative attribute, referred to as ‘hostile attribution bias.’ ***Table 64*** shows that the average percentage was similar across treatment groups.

Table 63. Percentage Negative Attributions by Students in Kenya

Scenario question	%
Why do you think the child took the toy away?	50.4
Why do you think the child wasn’t helping to sweep the room?	63.6
Why do you think the child is building the tower/building a different way from you?	61.6
Why do you think your classmate hasn't given you a turn yet?	61.6
Why do you think your classmate stood in front of you and took your place?	54.8
Total	58.4

Table 64. Average Percentage of Negative Perception by Treatment Group

Treatment group	Average percent of negative perception
Control	56.9
Phase 1	56.6
Phase 2	61.3
Overall	58.4

Confidence

Students were read descriptions of two children, a child who shows confidence and a child who does not show confidence, and were asked which one they identified with. ***Table 65*** shows the percentage of children identifying with the confident child across the seven questions. ***Table 66*** shows that the average percentage of confident responses did not differ significantly across treatment groups. Scores were well distributed with a mean of 4.4.

Table 65. Percentage of Students Identifying with Confident Person

Question	%
X often feels confident (not afraid) when going into a room full of people.	58.9
When children in the class have a question, they get X to ask the teacher.	59.0
X can lead other children to sing a song.	67.0
X starts conversations with adults he/she doesn’t know well.	48.3
X likes to read stories out to the whole class.	67.8
X is trusted to look after younger children.	75.5
X shares his/her thoughts even if they are very different from others.	69.3
Overall	63.7

Table 66. Percentage Students Identifying with Confident Person by Treatment Group	
Treatment group	**Average percentage identifying with confident person**
Control	62.9%
Phase 1	62.7%
Phase 2	65.3%
Overall	63.7%

Creativity

Children were asked to imagine different uses for two everyday objects – a spoon and a pencil. On average, children in Kenya named 4 different ways these objects could be used. ***Table 67*** shows students' average creativity scores (the number of different uses they name), by treatment group. The distribution of creativity scores skews left, with most children suggesting very few (between one and five) different uses for the two items.

Table 67. Average Creativity Score by Treatment Group	
Treatment Group	**Average**
Control	4.1
Phase1	3.9
Phase2	4.0
Total	4.0

Student Literacy and Numeracy Outcomes

Shortened versions of both a traditional EGRA and EGMA were administered to 2,155 Grade 1 students across treatment groups in Kenya. These assessments consisted of the same four EGRA subtasks administered in Kiswahili and English, as well as six subtasks on the EGMA.

As previously stated, the purpose of these measures is simply to set a baseline against which program impact will be assessed at midline and endline. Therefore, it should not be surprising that performance was low across the majority of subtasks (as shown in ***Table 68***). This is to be expected given that students were assessed in Grade 1, at which point they are only just beginning to be introduced to these skills. For example, upwards of half of students (across all groups) scored zero on the invented word and oral reading passage subtasks in Kiswahili and English (with slightly higher English performance on all EGRA tasks except for reading comprehension). Performance was stronger in the letter sound identification task, which tests a lower order skill that is a starting point for the higher order EGRA skills. EGMA scores followed an even more conspicuous pattern of decreasing accuracy and fluency scores (with increasing proportions of zero scores) from the first to the fourth subtask.

Innovation and network treatment schools tended to outperform control schools on all subtasks, with larger differences in literacy than numeracy. These differences will be accounted for at midline and endline.

Table 68. EGRA and EGMA Results for Grade 1 Students in Kenya, by Subtask and Treatment Group

Subtasks	Measure	Control	Phase 1	Phase 2
		n= 766	*n*= 610	*n*= 779
Kiswahili Reading Subtasks				
Letter Sounds	Fluency (correct letters per minute)	14.8 [±.9]	23.4 [±1.4]	24.2 [±1.2]
	Percentage correct of items attempted	42.4 [±2.2]	56.1 [±2.4]	58.3 [±2.0]
	Percentage of students with zero scores	26 [±3.1]	12.6 [±2.7]	12.6 [±2.3]
Invented Words	Fluency (correct letters per minute)	1.9 [±.4]	3.6 [±.6]	2.9 [±.5]
	Percentage correct of items attempted	10.2 [±1.7]	17.8 [±2.4]	15.8 [±2.0]
	Percentage of students with zero scores	81.2 [±2.8]	69.6 [±3.7]	71.1 [±3.2]
Oral Reading Passage	Reading Fluency (correct words per minute)	2.4 [±.5]	3.8 [±.7]	3 [±.5]
	Percentage correct of items attempted	12.4 [±2.0]	18.4 [±2.6]	15.8 [±2.1]
	Percentage of students with zero scores	80.4 [±2.8]	73 [±3.5]	74.8 [±3.1]
Reading Comprehension	Percentage correct	6.8 [±1.3]	11.4 [±2.0]	9 [±1.5]
	Percentage correct of items attempted	7.5 [±1.4]	12.4 [±2.1]	10.2 [±1.6]
	Percentage of students with 80% comprehension	1.6 [±.9]	6.6 [±2.0]	3 [±1.2]
	Percentage of students with zero scores	85.9 [±2.5]	79.5 [±3.2]	80.9 [±2.8]
English Reading Subtasks				
Letter Sounds	Fluency (correct letters per minute)	19.1 [±1.2]	29.8 [±1.8]	29.7 [±1.4]
	Percentage correct of items attempted	42.4 [±2.2]	57 [±2.5]	58.5 [±2]
	Percentage of students with zero scores	26.4 [±3.1]	12 [±2.6]	9.8 [±2.1]
Invented Words	Fluency (correct letters per minute)	5.6 [±.6]	7.9 [±.9]	7.2 [±.7]
	Percentage correct of items attempted	27.8 [±2.5]	35.1 [±3.0]	34.2 [±2.7]
	Percentage of students with zero scores	53.8 [±3.5]	49.8 [±4.0]	49.9 [±3.5]
Oral Reading Passage	Reading Fluency (correct words per minute)	3.1 [±.6]	6.5 [±1.1]	4.3 [±.7]
	Percentage correct of items attempted	11.6 [±1.6]	19 [±2.4]	15.2 [±1.9]
	Percentage of students with zero scores	72.5 [±3.2]	66.4 [±3.8]	69.2 [±3.2]
Reading Comprehension	Percentage correct	.9 [±.4]	7.6 [±1.7]	3.5 [±.0]

Table 68. EGRA and EGMA Results for Grade 1 Students in Kenya, by Subtask and Treatment Group

Subtasks	Measure	Control	Phase 1	Phase 2
		n= 766	*n*= 610	*n*= 779
	Percentage correct of items attempted	.9 [±.4]	7.7 [±1.7]	3.6 [±.0]
	Percentage of students with 80% comprehension	.1 [±.5]	4.6 [±1.7]	1.4 [±.9]
	Percentage of students with zero scores	97.3 [±1.2]	86 [±2.8]	91.9 [±1.9]
Math Subtasks				
Quantity Discrimination	Percentage correct	34.4 [±1.7]	38.8 [±2.2]	42.6 [±1.8]
	Percentage correct of items attempted	40.3 [±1.7]	43.6 [±2.0]	47.2 [±1.7]
	Percentage of students with zero scores	10.4 [±2.2]	11.5 [±2.5]	8 [±1.9]
Missing Number	Percentage correct	18.8 [±.8]	20.6 [±1.3]	18.7 [±.9]
	Percentage correct of items attempted	28.3 [±.9]	29.7 [±1.3]	28.2 [±.9]
	Percentage of students with zero scores	9.7 [±2.1]	9.3 [±2.3]	9.6 [±2.1]
Basic Addition	Fluency (correct addition problems per minute)	3.4 [±.2]	3.7 [±.3]	3.2 [±.2]
	Percentage correct of items attempted	35.6 [±2.5]	44.4 [±3.1]	37.2 [±2.7]
	Percentage of students with zero scores	33.7 [±3.4]	28.4 [±3.6]	34.3 [±3.3]
Basic Subtraction	Fluency (correct subtraction problems per minute)	1.6 [±.2]	1.9 [±.2]	1.5 [±.2]
	Percentage correct of items attempted	17.6 [±2.0]	24.5 [±2.9]	18.8 [±2.2]
	Percentage of students with zero scores	63.6 [±3.4]	59.5 [±3.9]	64.2 [±3.4]
Addition Level 2	Percentage correct	5.2 [±1.1]	7.4 [±1.3]	7.2 [±1.3]
	Percentage correct of items attempted	5.2 [±1.1]	7.4 [±1.3]	7.2 [±1.3]
	Percentage of students with zero scores	87.9 [±2.3]	79.9 [±3.1]	81.7 [±2.8]
Subtraction Level 2	Percentage correct	1.8 [±.6]	2.6 [±.8]	2.8 [±.8]
	Percentage correct of items attempted	1.8 [±.6]	2.6 [±.8]	2.8 [±.8]
	Percentage of students with zero scores	94.8 [±1.6]	92 [±2.1]	92.7 [±1.9]

2.3 Scaling Learning through Play - Rwanda

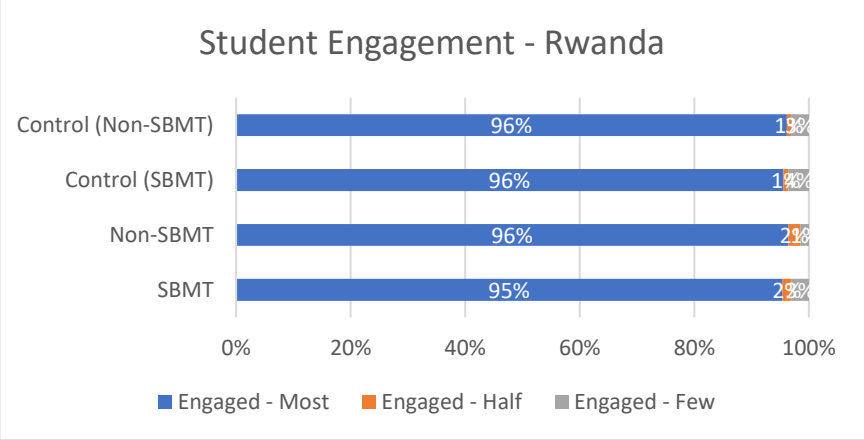
What is the baseline state of teachers' pedagogical practices in the classroom (including LtP) among Grade 1 teachers in treatment and control schools?

As in Ghana, classroom observations in Rwanda were conducted in Grade 1 classrooms for both reading and mathematics lessons. This decision was made in order to ensure variation in pedagogical practices (without being limited by approaches used for a particular subject). Accordingly, results for reading and mathematics lessons are combined throughout this section. Results are averaged across classroom observations at the school level and are disaggregated by treatment group throughout this section, with observations occurring in 52 SBMT implementation schools, 51 non-SBMT implementation schools, 33 SBMT control schools, and 50 non-SBMT control schools.

Timed Classroom Observations

Across all observed lessons, assessors in Rwanda reported that most students were engaged for nearly all of observed class time, with little variation across treatment groups **(Figure 16**). This strong level of engagement is a promising starting point for the program to continue to improve the depth and breadth of student engagement through implementation.

Figure 16. Student Engagement During Lessons (Percentage of Observed Class Time), by Treatment Group in Rwanda



Assessors measured classroom climate during observed lessons, focusing on four main criteria: perseverance (e.g., recognition of effort or encouragement of goal setting), positive verbal/physical (e.g., use of positive language of physical evidence of joy), negative verbal/physical (e.g., use of negative language or physical punishment), and ordinary (e.g., tone is neither noticeably positive nor negative). Assessors found an ordinary or positive climate for nearly all of observed class time (***Table 69***). There was very little negative climate (which is good), but there was also no perseverance.

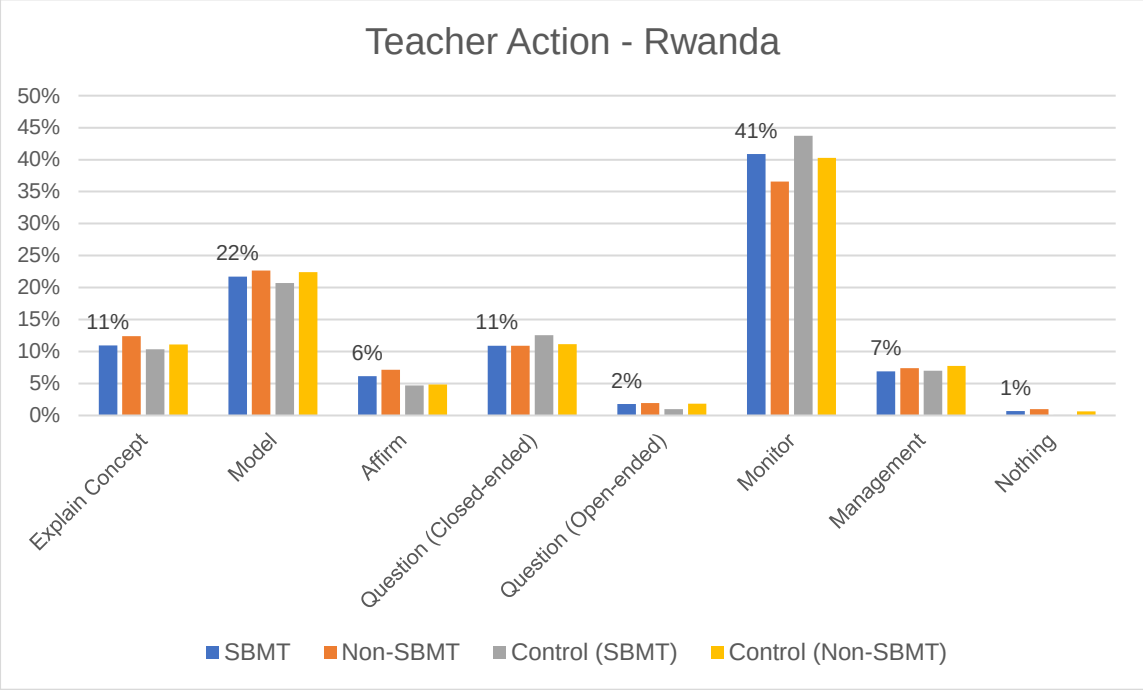
Table 69. Observed Classroom Climate (Percentage of Observed Class Time), by Treatment Group in Ghana

	SBMT	Non-SBMT	Control (SBMT)	Control (Non-SBMT)
Ordinary	91%	90%	95%	89%
Positive verbal/physical	8%	10%	5%	10%
Negative verbal/physical	1%	0%	0%	0%

Perseverance	0%	0%	0%	0%
Number of schools	52	51	33	50

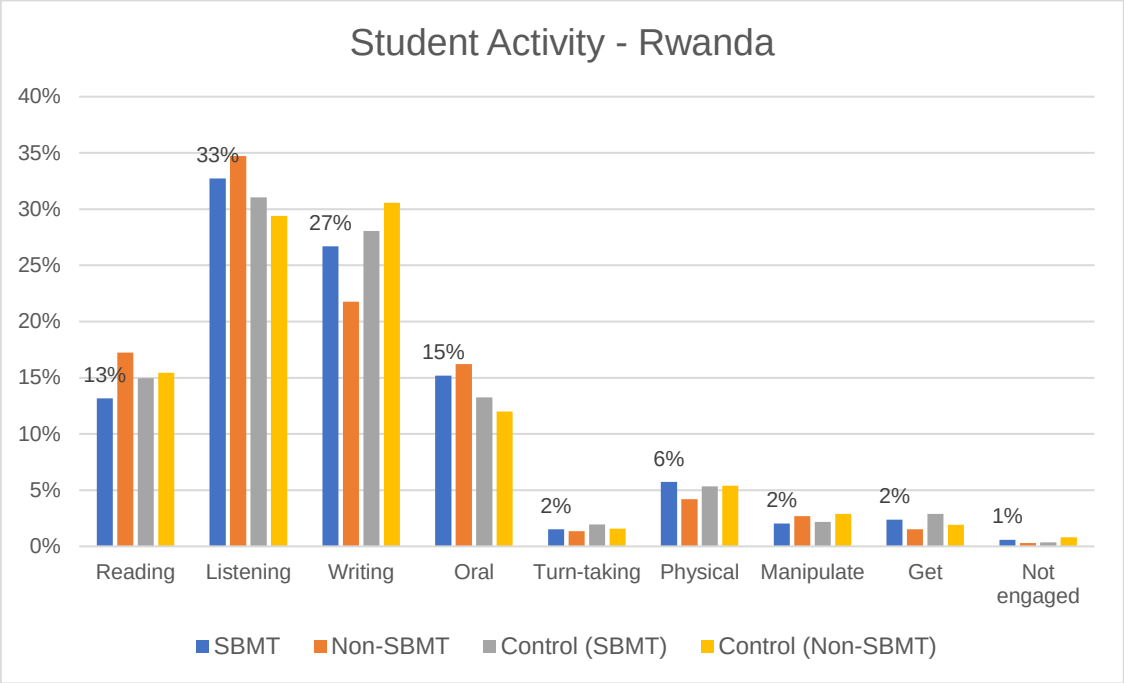
During the observed lessons, teachers spent the largest proportion of time monitoring students' understanding (or tracking participation in an activity) (**Figure 17**). They spent another third of their time explaining concepts or modeling activities. More interactive approaches were used less frequently in the classroom, which is a viable area of improvement for the program. For example, open-ended questioning was rarely used (approximately 2% of class time across groups)—and closed-ended questioning was more prevalent. Teacher affirmations or recognition of student input made up a small proportion of time, while management or an unengaged teacher (i.e., nothing) accounted for under 10% of class time.

Figure 17. Teacher Actions During Lessons (Percentage of Observed Class Time), by Treatment Group in Rwanda



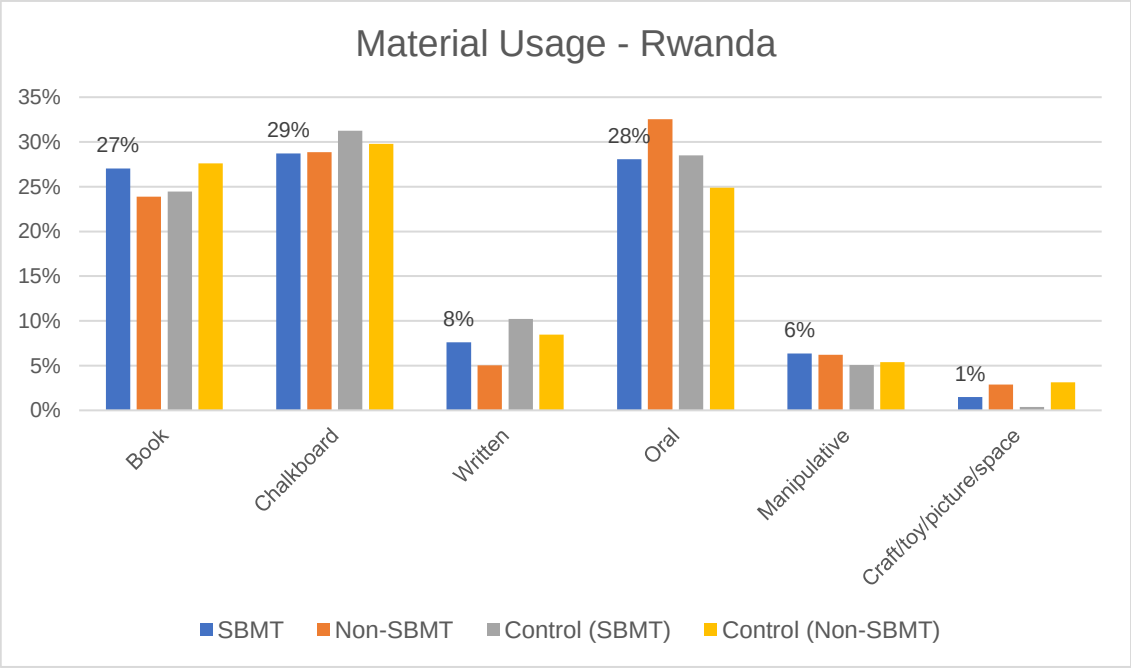
With regard to student activity, the largest amount of class time (about one-third) was spent listening (**Figure 18**). Interestingly, even though these are Grade 1 classrooms, approximately 25%–30% of time was spent writing, while only about 15% was dedicated to reading (with the same dedicated to students responding orally). About 10% of class time overall was dedicated to students either discussing/responding to one another (turn-taking), having a physical response, or manipulating something with their hands. Students spent less than 5% of class getting something to work on, waiting for the teacher, or disengaged (i.e., 'get' and 'not engaged' in **Figure 18**).

Figure 18. Student Activity During Lessons (Percentage of Observed Class Time), by Treatment Group in Rwanda



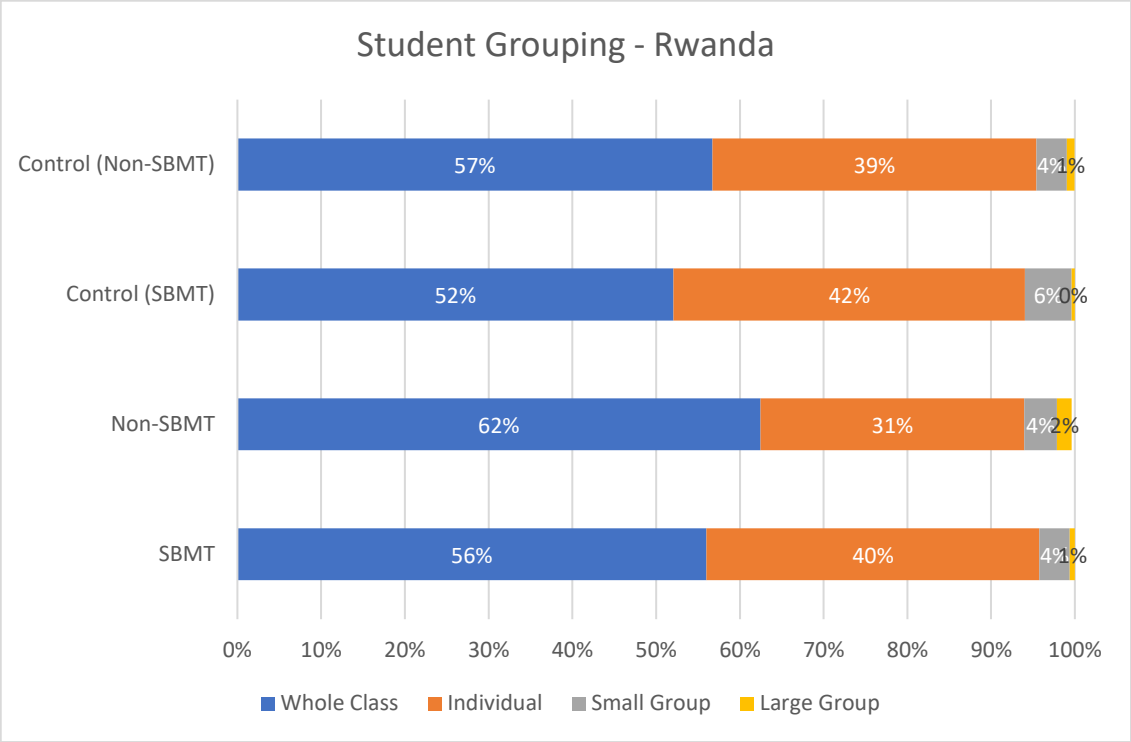
Materials provide a great opportunity for programs to increase the use of playful practices in the classroom. At baseline, more than half of class time across treatment groups was spent either using a chalkboard or with material being presented orally (**Figure 19**). The use of books also accounted for approximately one-quarter of class time. This left little time for the use of other written materials (such as letter/number cards or pocket charts), manipulatives, and alternative materials (e.g., crafts, toys, illustrations, space, etc.).

Figure 19. Use of Materials During Lessons (Percentage of Observed Class Time), by Treatment Group in Rwanda



Although the majority of teachers in Rwanda said that children learn language/literacy best through group work and that student exploration is important (on the attitude survey), these are still areas for improvement in actual instruction. The results of the classroom observations showed that more than half of class time was spent on whole class instruction, with 31% to 42% of the time used for individual work (**Figure 20**). Less than 10% of the total class time was spent with students working in large or small groups.

Figure 20. Student Grouping During Lessons (Percentage of Observed Class Time), by Treatment Group in Rwanda



Note: "Corners" omitted from the table because <1% of time was spent using learning corners in observed classrooms.

There was also evidence of limited student agency found in the classroom observations in Rwanda, with nearly all of class time dedicated to activities that were initiated and directed by teachers (***Table 70***). Interestingly, while 8% to 11% of time was used for student-directed work in SBMT schools (treatment and control), only 2% to 3% was used for the same activities in non-SBMT schools.

Table 70. Student Agency (Percentage of Observed Class Time), by Treatment Group in Rwanda

	SBMT	Non-SBMT	Control (SBMT)	Control (Non-SBMT)
Teacher directed	91%	97%	88%	96%
Teacher initiated; Student directed	8%	2%	11%	3%
Student initiated	1%	0%	0%	1%

Post-Classroom Observation Indicators

At the end of each timed classroom observation, assessors completed a checklist to determine whether a range of activities occurred (at least once) during the lesson. These indicators are divided into six constructs: 1) integrated pedagogies, 2) application (use of games), 3) responsiveness (to student behaviors), 4) resources (availability and use), 5) efficiency (maximized instructional time), and 6) student-directed (ownership or autonomy).

As shown in ***Table 71***, there was significant variation in the prevalence of activities occurring throughout the observed lessons (with high prevalence in green, moderate prevalence in yellow, low prevalence in red). In terms of constructs, teachers tended to perform best on integrated pedagogy, responsiveness and efficiency indicators. Resources included some high prevalence and some low prevalence items, while application tended to be low (aside from using songs or energizers in the classroom) and the most limited use was with regard to student-directed opportunities. Overall, basic best practices (e.g., student monitoring and support, material preparation, good instructional pace, making connections and sharing ideas, inclusionary practices) were used throughout lessons, but there is significant room for improvement in including more playful pedagogies.

Table 71. Post-Observation Checklist (Percentage of Observed Lessons), by Treatment Group in Rwanda

	SBMT	Non-SBMT	Control (SBMT)	Control (Non-SBMT)
Integrated Pedagogies				
The teacher asked questions or raised topics that got students to respond to each other, not to her (i.e., active).	75%	64%	49%	58%
Student(s) persisted with the activity after the rest of the class has started to do something else (i.e., engage).	64%	64%	68%	52%
Student(s) started the task again or adjusted when they thought of a better idea/approach (i.e., iterative).	56%	75%	74%	63%
Student(s) had a chance to make connections between the new experiences/content and their prior knowledge or experiences (i.e., meaningful).	74%	66%	77%	73%
Students problem solved or had a chance to think critically or reflect about a topic.	53%	57%	61%	72%
Application: Students have generalized games				
Songs or energizers to start or divide activities.	63%	45%	35%	55%
Students played games with obvious rules.	16%	16%	9%	14%
Students had physical play.	22%	29%	29%	25%
Students used objects to represent something else.	11%	17%	9%	10%
Students had symbolic play.	22%	21%	15%	22%
Responsiveness: The teacher adapts to student behaviors				
The teacher gave students hints, suggestions, or redirection when they needed it.	94%	91%	86%	92%
The teacher circulated, leaned in, redirected, to support students as needed.	87%	92%	88%	88%
The teacher asked the class or a group for a solution to a student question or problem.	43%	49%	30%	40%
The teacher included students who did not volunteer to answer (i.e., Inclusion).	92%	83%	91%	89%
Resources: Materials that are needed for implementation are used				
Students used a resource that was dedicated for their use (e.g., book, workbook, slate, tablet).	86%	81%	85%	88%
Students used something to create something else (e.g., art supplies, items in nature, objects repurposed).	12%	10%	8%	14%

Table 71. Post-Observation Checklist (Percentage of Observed Lessons), by Treatment Group in Rwanda

	SBMT	Non-SBMT	Control (SBMT)	Control (Non-SBMT)
Integrated Pedagogies				
Students used a resource to solve a problem (e.g., manipulatives, math charts, dictionary).	50%	52%	38%	53%
Instruction used an item designed for interaction between teacher and students or between students (e.g., read aloud, pocket chart, counters, puzzle).	85%	74%	82%	81%
Students used physical space for learning activities (e.g., counting large steps for measurement, sit in a circle for sharing).	6%	20%	9%	7%
Students used stations/corner work/learning centers.	14%	24%	6%	14%
Efficiency: Instructional time is maximized				
The teacher had materials ready at the start of an activity or the class.	89%	81%	88%	89%
The pace of the instruction was lively.	87%	82%	88%	85%
Student-directed (ownership / autonomy)				
Students developed and shared ideas with either classmates or the teacher.	27%	21%	27%	17%
Students had a chance to select from choices (state of play: exploring).	14%	7%	17%	19%
Students chose collaborators or roles.	1%	0%	0%	4%
Students made or changed rules for an activity.	2%	0%	0%	2%
Students chose how long to do an activity (state of play: owning).	4%	0%	3%	4%
Students set goals or helped to plan how to use time.	3%	2%	6%	6%
Students chose when to move around.	3%	0%	0%	2%

Note: The color scale is a gradient based on the proportion of lessons with observed indicators. Green = high prevalence; yellow = moderate prevalence; orange = low prevalence; red = extremely low prevalence.

In addition to the item level estimates in Table 78, construct level scores were calculated (***Table 72***). These scores will be used as the basis for measuring change/impact across treatment groups, over time. The scores are calculated as simple averages of the number of items occurring per construct for each treatment group. At baseline, there were some statistically significant differences between treatment schools and their respective controls (as shown in Table 79). These differences will be accounted for in midline/endline impact estimates.

Table 72. Post-Observation Checklist Scores, by Treatment Group in Rwanda

	SBMT	Non-SBMT	Control SBMT	Control Non-SBMT
Integrated Pedagogies (0–5)	3.2	3.2	3.3	3.2
Application: Students have generalized games (0–5)	1.3*	1.3	1.0	1.3
Responsiveness: The teacher adapts to student behaviors (0–4)	3.2*	3.2	3.0	3.1
Resources: Materials that are needed for implementation are used (0–6)	2.5*	2.6	2.3	2.6
Efficiency: Instructional time is maximized (0–2)	1.8	1.6*	1.8	1.7
Student-Directed (Ownership/Autonomy) (0–7)	0.5	0.3*	0.5	0.5
Overall (0–29)	12.5*	12.2	11.8	12.4

*$p<0.05$ (comparison between SBMT treatment and control or non-SBMT treatment and control)

The post-observation component of the data collection also included a classroom inventory. The inventory was focused on determining the availability of basic materials in each classroom. For simplicity, the results displayed in ***Table 73*** are for reading lessons only. Overall, the availability of materials in the observed classrooms was low. The most readily available materials were teaching aids, which were present in 59% of classrooms, on average. Teacher's guides and textbooks were the least prevalent—though they were more available in SBMT treatment schools than in the other three groups. Lastly, just over a quarter of classrooms had supplementary readers, while only about 1 in 10 had workbooks.

Table 73. Post-Observation Reading Classroom Inventory (Availability in Percentage of Classrooms), by Treatment Group in Rwanda

	SBMT	Non-SBMT	Control SBMT	Control Non-SBMT	Overall
Teacher's guide	14%	6%	3%	2%	6%
Teaching aids	56%	60%	61%	64%	59%
Textbooks (present in classroom)	12%	4%	3%	0%	5%
Textbook ratio	14%	2%	14%	13%	11%
Workbooks	12%	8%	9%	12%	11%
Supplementary readers	21%	31%	30%	30%	28%
N	52	48	33	50	185

Learning Environment (Student Survey)

Students were asked 10 questions about the learning environment (see ***Table 74***) with each response giving a score of 0,1, or 2. Scores were normally distributed around the mean of 14.9. **Table 75** shows that the learning environment score was similar across treatment groups.

Table 74. Responses to Learner Environment Questions by Students in Rwanda

Item	No/Never	Sometimes	Yes/Often	Missing/No Reply
Does your teacher praise you when you give the right answer?	5	16	78	1
Can you talk to at least one teacher in the school about things that are bothering you?	11	12	75	1
Last week, did children in your class ask the teacher questions in lessons?	19	31	48	1
Last week, did some children in you class ask the teachers for help with their classwork?	20	32	47	1
Do your teachers encourage you to try out new things in class?	10	31	59	1
Do your teachers get angry or yell when you get the answer wrong?	53	27	18	1
Do your teachers help those who do not understand?	20	32	47	1
In the last week, did an adult in your school scream or yell at a child?	77	15	7	1
Can students share their thoughts in front of the class?	22	16	60	1

Table 74. Responses to Learner Environment Questions by Students in Rwanda				
Item	No/Never	Sometimes	Yes/Often	Missing/No Reply
In the last week, did an adult in your school hit k a child or use harsh punishment	78	13	8	1

Table 75. Learning Environment Score	
Learning Environment Score—Rwanda	
Treatment Status	Mean Score
Control-Non-SBMT	15.0
Control-SBMT	15.2
Treatment-Non-SBMT	14.8
Treatment-SBMT	14.6
Overall	14.9

What are the baseline perceptions of LtP among school, district, and community-level stakeholders?

Attitudes toward LtP

In Rwanda, attitudes towards LtP were similar across stakeholders, with trainers and DEOs having slightly lower support for LtP practices than head teachers, teachers, coaches, and teacher meeting facilitators (see ***Table 76*).**

Nearly all of the interview participants strongly supported play as an important component of learning for children in primary school (Item 1). In addition, most participants agreed that parents and caregivers support play-based approaches in the classroom (Item 2), that children learn best through exploration (Items 3 and 5), and that group work is beneficial for students (Item 4). Despite the support for exploration, however, the respondents highly favored teacher-led activities and tight teacher control of the classroom (Items 7 through 10).

Table 76. Stakeholder Attitudes toward LtP - Rwanda						
Items	Head Teacher	Teacher	Coach	Teacher Meeting Facilitator	Trainer	DEO
1—Play is an important part of learning for children in primary school.	88%	93%	86%	75%	90%	100%
2—Parents and caregivers welcome more play-based approaches in the classroom.	92%	87%	89%	63%	95%	73%
3—Children learn best by exploring new things on their own.	89%	79%	86%	38%	91%	73%
4—Children learn literacy/language best in small groups.	73%	64%	69%	38%	75%	91%
5—If children answer their own questions, then they will be more likely to understand.	64%	52%	60%	38%	72%	73%
6—It is important for children to be free to talk.	40%	45%	43%	63%	50%	82%
7—If children do not understand, then they should try to find the answer	36%	24%	45%	38%	50%	55%

Table 76. Stakeholder Attitudes toward LtP - Rwanda

Items	Head Teacher	Teacher	Coach	Teacher Meeting Facilitator	Trainer	DEO
independently before asking the teacher.						
8—It is more important that the students are engaged.	34%	24%	47%	38%	46%	36%
9—Children can choose their own learning activities in the classroom.	11%	9%	12%	38%	12%	55%
10—Children learn best from their peers.	23%	7%	28%	25%	32%	36%
Number of Interviews Conducted	**184**	**182**	**183**	**8**	**182**	**11**

Note: The color scale is a gradient based on the proportion of agreement. Green = strong agreement; yellow = moderate agreement; orange = low agreement; red = poor agreement.

An overall "playfulness" score was further calculated for teachers and head teachers. This score reflects the total number of items on which the respondent selected the statement considered more supportive of LtP. As ***Table 77*** shows, out of a total of 10 items, both teachers and head teachers in Rwanda agreed with the more playful response, on average, roughly one-half of the time. Across all treatment groups, head teachers were marginally more supportive of LtP than were teachers.

Table 77. Mean Attitudes Scale Scores toward LtP—Teachers and Head Teachers

Respondent	Control-Non-SBMT	Treatment-Non-SBMT	Control-SBMT	Treatment-SBMT
Teachers	4.9	5.1	4.8	4.6
Head teachers	5.4	5.8	5.5	5.4

A second, simplified survey about attitudes was given to caregivers, who were also asked to select their preference between two statements. The survey results are presented in ***Table 78.*** The caregivers tended to favor more play-based practices in the classroom (Items 3 and 5), but their attitudes were less approving than those of the teacher, and fewer caregivers agreed with Statement 1 than were observed in other countries. The caregivers were also divided over the importance of creating a friendly classroom environment (Item 2), though the majority did agree that children learn best when having fun (Item 4). The area with the most limited support from caregivers was a willingness to relinquish a focus on discipline and order in the classroom—as shown by the relatively lower agreement with the statements regarding more play that appear in Item 6.

Table 78. Caregiver Attitudes toward LtP—Rwanda

Respondent	Control-Non-SBMT	Treatment-Non-SBMT	Control-SBMT	Treatment-SBMT
1- It is more important to teach creativity than academic subjects.	66%	62%	66%	61%

Table 78. Caregiver Attitudes toward LtP—Rwanda				
Respondent	**Control-Non-SBMT**	**Treatment-Non-SBMT**	**Control-SBMT**	**Treatment-SBMT**
2- Teachers should create a friendly classroom environment.	55%	49%	58%	70%
3- Teachers should use play to promote learning in the classroom.	94%	88%	86%	92%
4- Children learn best when they are having fun.	96%	92%	97%	88%
5- I would be happy if teachers use play to support learning in the classroom.	78%	75%	81%	73%
6- It is important for children to be free to talk.	33%	40%	31%	36%

What are the enabling factors associated with implementing LtP pedagogical methods in classrooms? What are the barriers associated with implementing LtP pedagogical methods in classrooms?

Support to LtP

The sections below present information on the existing environment in which the Scaling Learning Through Play program will be implemented, broken out by four treatment groups. These sections are organized by stakeholder. At baseline, questions to SBMs serving as trainers and meeting facilitators were limited to their attitudes toward play, so this group is not included in the enabling environment findings presented below (they will be incorporated at midline). Additional findings on the supports received by teachers, coaches, and DEOs are presented in ***Annex A*.**

Grade 1 Teachers

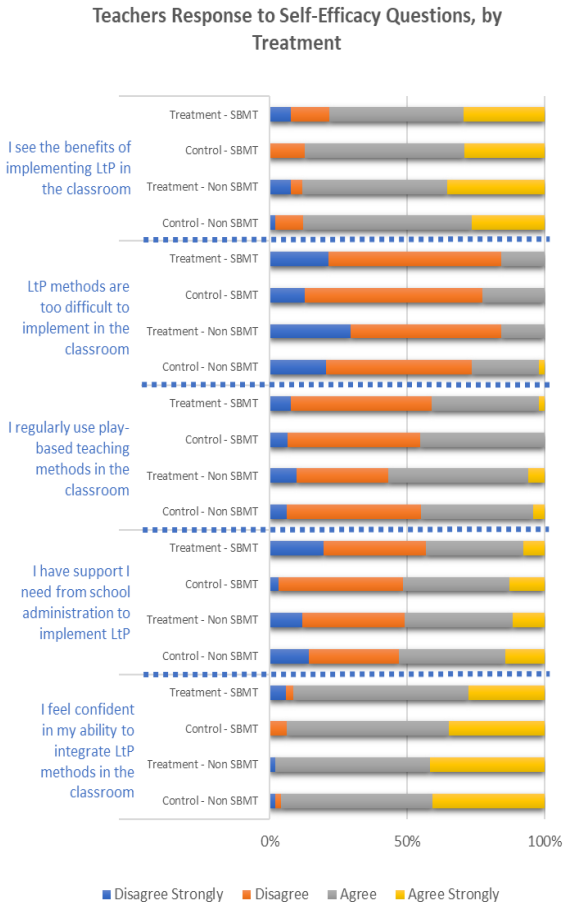
The most common type of support teachers reported receiving is training and mentoring. A slightly higher percentage of teachers in treatment non-SBMT schools reported receiving mentoring and more teachers in both treatment groups reported receiving coaching more frequently, but these differences were also slim **(*Table 79*).** Additional information on the frequency and content of these supports is presented in ***Annex A*.**

Table 79. Which of the Following Supports Do You Receive as a Teacher?				
	Control - Non-SBMT	**Treatment - Non-SBMT**	**Control - SBMT**	**Treatment - SBMT**
Training	82%	76%	71%	78%
Mentoring	78%	86%	71%	76%
Coaching (in-person)	33%	45%	32%	43%

Text messages or other remote advice or guidance	12%	14%	10%	12%
None of the above	0%	0%	6%	4%

Grade 1 teachers were asked questions aimed to measure their level of self-efficacy implementing LtP in the classroom **(*Figure 21*).** The majority of teachers agreed that they see the benefits of implementing LtP and disagreed that LtP methods are too difficult to implement in the classroom. Across all treatment groups, nearly all teachers agreed or strongly agreed that they are confident in their ability to integrate LtP in the classroom. Roughly one-half of teachers across groups said that they regularly use play-based teaching methods or have the support from school administration to do so.

Figure 21. Teachers' Self-Efficacy Using LtP, by Treatment Group



Caregivers of Grade 1 Children

To gain a fuller picture of children's exposure to play and play-based learning at home as well as in school, caregivers were asked about things they do to support their children at home. Additional findings on caregiver's involvement in school can be found in ***Annex A*.** As ***Table 80*** shows, when asked about activities they engaged in with their children in over the previous 3 days, over two-thirds of caregivers in all treatment groups said they helped children with homework and read or told stories. Roughly one-half of caregivers said they played games or cooked something together. There was some variation between treatment groups, which is likely a product of the broad and open-ended nature of the question.

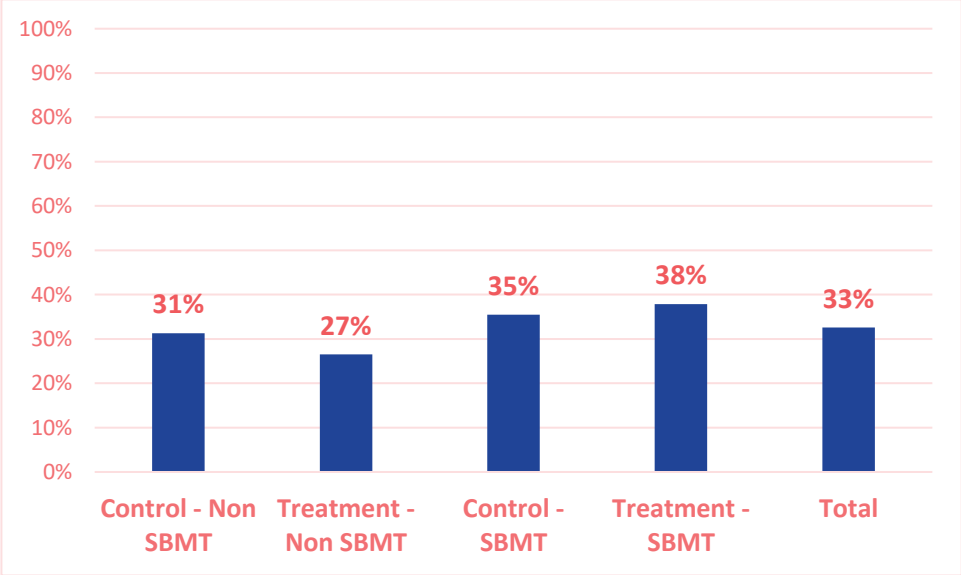
Table 80. Caregiver: In the Past 3 Days, Have You or Another Caregiver Engaged in Any of These Activities with Your Child?

	Control-Non-SBMT	Treatment-Non-SBMT	Control-SBMT	Treatment-SBMT
Helped with homework	89%	86%	91%	78%
Read storybooks	71%	69%	70%	72%
Told stories	68%	61%	72%	66%
Played a game inside the house	50%	48%	42%	51%
Played a game outside	57%	58%	41%	48%
Cooked something together	49%	44%	50%	46%

Table 80. Caregiver: In the Past 3 Days, Have You or Another Caregiver Engaged in Any of These Activities with Your Child?				
	Control-Non-SBMT	Treatment-Non-SBMT	Control-SBMT	Treatment-SBMT
Built something together	20%	29%	11%	22%
Made up a game together	25%	31%	19%	19%
Did a puzzle or board/card game	10%	14%	13%	14%
Other	5%	11%	11%	5%

Only 33% of caregivers replied that yes, they have play materials at home. **Figure 22** shows a breakdown of this question by treatment group:

Figure 22. Caregiver: Do You Have Any Play Materials at Home?



Among caregivers who reported having play materials at home, the types of materials varied and should be considered as descriptive findings not used for comparison at midline. Across all treatment groups, the highest proportion of caregivers said they had sports equipment at home, followed by dolls and homemade toys. Far fewer caregivers said they had books at home, which is surprising given the high percentage who responded they read stories with their child (***Table 81*).**

Table 81. Caregiver: What Types of Play Materials Do You Have at Home?				
	Control-Non-SBMT	Treatment-Non-SBMT	Control-SBMT	Treatment-SBMT
Balls or sports equipment	71%	52%	54%	73%
Dolls	32%	30%	13%	25%
Homemade toys	32%	19%	42%	23%
Craft supplies (paper scissors crayons paint coloring books)	16%	15%	4%	13%
Toys from the market	19%	11%	21%	13%
Blocks or other building toys	10%	0%	0%	10%

Table 81. Caregiver: What Types of Play Materials Do You Have at Home?				
	Control-Non-SBMT	Treatment-Non-SBMT	Control-SBMT	Treatment-SBMT
Other	7%	11%	4%	10%
Books	16%	11%	17%	3%

Head Teachers

Questions for head teachers at baseline focused on how they identify and support struggling teachers. As ***Table 82*** shows, most head teachers rely on student results, routine interactions with teachers, and time spent supporting the teacher to determine whether their teachers are struggling. Over one-half of head teachers said they also rely on coaching feedback or data, most likely referring to discussions with SBMs.

Table 82. Head Teacher: How Do You Know if a Teacher in Your School Is Not Performing as Expected?				
	Control-Non-SBMT	Treatment-Non-SBMT	Control-SBMT	Treatment-SBMT
Student results	76%	78%	76%	83%
Regular interaction/discussion with teachers	68%	78%	70%	77%
Coaching feedback	62%	59%	52%	73%
Supporting the teacher	72%	75%	76%	69%
Coaching data	54%	63%	49%	64%
Other (Specify)	18%	12%	18%	12%

The importance of mentors in Rwandan schools was further underscored when head teachers were asked what they do when a teacher is struggling. Over two-thirds of head teachers across treatment groups said they would ask another teacher to mentor the teacher or provide additional support themselves. Curiously, there was little difference in the percentage of head teachers from SBMT and non-SBMT schools who would engage a mentor or coach to support the teacher (***Table 83*****).**

Table 83. Head Teacher: What Action Do You Take if a Teacher Is Not Performing as Expected?				
	Control-Non-SBMT	Treatment-Non-SBMT	Control-SBMT	Treatment-SBMT
Ask another teacher to mentor the teacher	66%	75%	76%	83%
Provide additional support to the teacher yourself	78%	80%	70%	81%
Request additional coaching support for the teacher	64%	63%	58%	64%
Observe the teacher (more frequently)	64%	55%	58%	62%
Other	14%	10%	18%	15%

Coaches

Coach interviews were conducted with the Director of Studies at each school. At smaller schools, where there was no Director of Studies, the Deputy Head Teacher was interviewed. Additional findings on the supports coaches say they provide to support LtP can be found in **Annex A.**

Most of the coaches (88%) reported feeling confident or very confident in their ability to support teachers with playful and interactive or student-directed strategies in the classroom. Based on their experiences, the vast majority of coaches felt that the biggest challenge to teachers' increasing their use of play-based approaches was a lack of the appropriate materials or support (***Table 84***).

Table 84. Coaches: What Do You Think Are the Biggest Challenges to Increasing the Amount of Playful or Interactive Activities in the Classroom?

	Control-Non-SBMT	Treatment-Non-SBMT	Control-SBMT	Treatment-SBMT
Teachers do not have the appropriate materials/support	94%	84%	91%	78%
Teachers do not have the necessary knowledge/skills	58%	63%	56%	54%
Teachers do not support play in the classroom	8%	12%	6%	14%
It is difficult to find time for play	14%	14%	3%	14%
There is no incentive to increase play	8%	6%	13%	6%
Head teachers do not support play in the classroom	4%	10%	9%	0%

District Education Officials

DEO interviews were conducted with 11 DEOS and are presented in aggregate. Baseline interviews with DEOs focused on the support they provide to schools. Additional findings on DEOs' roles and communication provided at the district and sub-district levels can be found in ***Annex A***.

When asked what their offices' role is in supporting the Scaling Learning Through Play program, DEOs noted a myriad of responsibilities, though none mentioned overseeing mentoring or teacher meetings, or working with parents and communities (***Table 85***).

Table 85. DEOs: What Is Your (or Your Office's) Role in Supporting Scaling Learning Through Play?

Role	Percent
Communicating program expectations to sub-district and school-level staff	9%
Ensuring that teachers participate in program activities	9%
Providing coaching and/or school supervision	9%
Providing monitoring of program implementation	9%
Training sub-district or school-level staff	9%
Overseeing or providing funds for mentoring and/or teacher meetings	0%
Working with parents/communities on program components	0%
No role in supporting the program	0%

DEOs were also asked questions about access to play-based learning supports in their district. When asked "Does the curriculum used in your primary schools include play-based learning activities and teaching methods?" all 8 of the 11 DEOs said yes. As would be expected at baseline, only 2 DEOs also said that they have received training or orientation on "learning through play" or play-based learning. When asked "Do primary teachers in your

schools have materials for play-based learning in the classroom?" DEOs were largely in agreement (9 DEOs) that they do not.

What are the baseline levels of student literacy, numeracy, and SEL among Grade 1 students in treatment and control schools?

Assessment of SEL

Observed Student Behavior During Assessments

During the SEL assessment, six student behaviors were observed by the assessor (see ***Table 86***), providing a measure of student confidence and positive emotions. The total score was normally distributed around the mean of 8.2. ***Table 87*** shows that the mean score was similar across treatment groups (see **Annex A** for distributions of SEL scores).

Table 86. Student Behaviors

Behavior	Mean	Max
Does the child greet you confidently?	1.09	2
Does the child respond to questions confidently?	1.34	2
Does the child look you in the eye when listening to or responding to questions?	1.38	2
Actively attempts to engage interviewer	1.62	3
Shows pleasure in accomplishment and active task mastery	1.48	3
Child shows frequent positive emotions and behaviors	1.27	3
Total	8.18	15

Table 87. Observed Behavior Score by Treatment Group

Treatment Group	Average Score
Control-Non-SBMT	8.6
Control-SBMT	8.4
Treatment-Non-SBMT	7.9
Treatment-SBMT	7.8
Overall	8.2

Relationships

Table 88 shows that most children—around 75% to 90%—have a friend they can turn to for help in difficult situations or when learning and with whom they can share secrets. The distribution of the relationship score is skewed to the right, with 52% of children agreeing with all five questions. The mean was 4.2. ***Table 89*** shows that the mean percentage of children responding positively to relationship questions was similar across treatment groups, with an overall mean of 83.2%. In an additional question, children named up to 41 friends by name with a median number of 3 friends.

Table 88. Relationship Item Scores in Rwanda

Item	%
Do you talk to any of your friends when you feel sad? If Yes, whom do you talk to?	88.63
When you have an argument/fight with another child, do you go to any of your friends?	74.05
If you are working on something difficult, do you ask any of your friends for help?	90.53
Do you have a friend who helps you learn? If yes, who helps you learn?	84.06

Do you have a friend whom you tell your secrets to? If yes, who?	78.92
Do you talk to any of your friends when you feel sad? If Yes, whom do you talk to?	83.2

Table 89. Average Percent of Relationship Items

Treatment Group	Average Percent of "Yes" responses
Control-Non-SBMT	81.7
Control-SBMT	84.7
Treatment-Non-SBMT	84.5
Treatment-SBMT	82.6
Overall	83.2

Conflict Resolution

Children listened to five scenarios in which there was potential conflict with another student and were asked what they would do in that situation. For each scenario their response was placed into one of eight categories (see ***Table 90***). The most common responses across all scenarios were disengagement from the situation or appealing to authority (i.e., asking the teacher to intervene). ***Table 91*** shows that the average number of children giving each type of response was similar across treatment groups.

Table 90. Percentage of responses to Conflict Resolution Scenarios in Rwanda									
	Level 0		Level 1	Level 2		Level 3			
Scenario number	Verbal aggression	Physical aggression	Disengage	Command	Appeal to authority	Ask for reasons	Influence, give reasons	Help the other child	Missing/no response
1	1.1	16.9	28.3	9.8	36.0	1.4	3.0	0.9	2.6
2	0.4	1.5	38.0	3.5	45.1	1.7	6.5	1.2	2.1
3	0.3	1.7	41.0	3.8	36.4	1.8	7.4	3.0	4.6
4	0.4	4.6	32.8	6.1	47.9	2.4	2.6	0.4	2.8
5	1.8	10.9	32.7	4.2	42.6	1.6	2.2	0.7	3.4
Total	**0.8**	**7.2**	**35.2**	**5.6**	**42.4**	**1.8**	**4.4**	**1.3**	**1.2**

Table 91. Average Percent of Responses Conflict Resolution Scenarios by Treatment Group									
	Level 0		Level 1	Level 2		Level 3			
Treatment group	Verbal aggression	Physical aggression	Disengage	Command	Appeal to authority	Ask for reasons	Influence, give reasons	Help the other child	Missing/no response
Control-Non-SBMT	0.6	8.6	33.5	5.3	44.5	1.2	3.9	0.9	1.5
Control-SBMT	1.1	5.5	33.7	5.8	40.5	2.4	4.1	1.7	5.2
Treatment-Non-SBMT	0.9	6.8	36.4	5.1	40.8	1.5	4.0	1.5	2.9
Treatment-SBMT	0.7	7.0	34.3	5.8	40.3	2.1	5.2	1.1	3.4
Total	**0.8**	**7.2**	**35.2**	**5.6**	**42.4**	**1.8**	**4.4**	**1.3**	**1.2**

Each response was classified as Level 0, Level 1, Level 2, or Level 3 and received a corresponding score of 0, 1, 2, or 3. The distribution of the score, totaled across the five scenarios. The scores are normally distributed around the mean of 6.3.

For each scenario, students were asked why the other child in the scenario behaved the way they did. The explanation was coded according to whether the motive of the other student was seen as being negative toward the student being interviewed (e.g., "she doesn't like me"), neutral (e.g., "she was distracted"), or positive (e.g., "she was trying to help me"). ***Table 92*** shows the percentage of students with a negative attribute, referred to as 'hostile attribution bias.' ***Table 93*** shows that the average percentage was similar across treatment groups.

Table 92. Percentage Negative Attributions by Students

Scenario question	%
Why do you think the child took the toy away?	77.4
Why do you think the child wasn't helping to sweep the room?	82.4
Why do you think the child is building the tower/building a different way from you?	80.74
Why do you think your classmate hasn't given you a turn yet?	83.01
Why do you think your classmate stood in front of you and took your place?	84.2
Total	81.6

Table 93. Percentage Negative Attributions by Treatment group

Treatment Group	Average Percent of Negative Perception
Control-Non-SBMT	83.2
Control-SBMT	81.1
Treatment-Non-SBMT	82.4
Treatment-SBMT	79.5
Overall	81.8

Confidence

Students were read descriptions of two children—a child who shows confidence and a child who does not show confidence—and asked which one they identified with. ***Table 94*** shows the percentage of children identifying with the confident child across the seven questions. The average percentage of confident responses did not differ significantly across treatment groups ***(Table 95)***. The confidence scores were distributed well around the mean of 4.8.

Table 94. Percentage Students Identifying with Confident Person

Question	Average Percent of identifying with confident person
X often feels confident (not afraid) when going into a room full of people.	74.45
When children in the class have a question, they get X to ask the teacher.	60.81
X can lead other children to sing a song.	69.08
X starts conversations with adults he/she doesn't know well.	63.16
X likes to read stories out to the whole class.	67.34
X is trusted to look after younger children.	76.66
X shares his/her thoughts even if they are very different from others'.	69.5

Overall	68.7

Table 95. Percentage Students Identifying with Confident Person by Treatment Group

Treatment group	Average percentage identifying with confident person
Control-Non-SBMT	69.3
Control-SBMT	70.2
Treatment-Non-SBMT	66.5
Treatment-SBMT	69.4
Total	68.7

Creativity

Children were asked to imagine different uses for two everyday objects – a spoon and a pencil. On average, children in Rwanda named 3.4 different ways these objects could be used. ***Table 96*** shows students' average creativity scores (the number of different uses they name), by treatment group. The distribution of creativity scores skews left, with most children suggesting very few (between one and five) different uses for the two items.

Table 96. Average Creativity Score by Treatment Group

Treatment Group	Average
Control-Non SBMT	3.3
Control-SBMT	3.6
Treatment-Non SBMT	3.2
Treatment-SBMT	3.5
Overall	3.4

Student Literacy and Numeracy Outcomes

Shortened versions of a traditional EGRA and EGMA were administered to nearly 3,000 Grade 1 students across treatment groups in Rwanda. The EGRA consisted of four subtasks in Kinyarwanda, with an additional English reading passage and comprehension section. The EGMA consisted of six subtasks, all of which were administered in Kinyarwanda.

As previously stated, the purpose of these measures is simply to set a baseline against which program impact will be assessed at midline and endline. Therefore, it should not be surprising that performance was low across the majority of subtasks (as shown in ***Table 97***). This is to be expected given that students were assessed in Grade 1, at which point they are only just beginning to be introduced to these skills. For example, approximately three-quarters of students (across all groups) scored zero on the invented word and oral reading passage subtasks in Kinyarwanda, while oral reading passage in English and reading comprehension in both languages had upwards of 90% zero scores. The slightly lower performance in English is not surprising but somewhat concerning given the transition to English as a language of instruction. Performance was stronger in the letter sound identification task, which tests a lower order skill that is a starting point for the higher order EGRA skills. EGMA scores followed an even more conspicuous pattern of decreasing accuracy and fluency scores (with increasing proportions of zero scores) from the first to the fourth subtask.

While scores were relatively consistent across treatment groups, there were some statistically significant differences, which will be accounted for at midline and endline.

Table 97. EGRA and EGMA Results for Grade 1 Students in Rwanda, by Subtask and Treatment Group

		Control		Treatment	
Reading subtasks	**Measure**	**Control-non-SBMT**	**Control-SBMT**	**Treatment-non-SBMT**	**Treatment-SBMT**
		n= 1636			
Kinyarwanda Subtasks					
Letter Sounds	Fluency (correct letters per minute)	17.6 [±1.2]	19 [±1.4]	15.2 [±1.1]	18.9 [±1.1]
	Percentage correct of items attempted	46.9 [±2.3]	48.2 [±2.8]	40.3 [±2.3]	48.2 [±2.2]
	Percentage of students with zero scores	23 [±2.9]	22.4 [±3.6]	29.5 [±3.2]	22.4 [±2.8]
Invented Words	Fluency (correct letters per minute)	2.9 [±.5]	2.2 [±.5]	2.3 [±.4]	2.7 [±.4]
	Percentage correct of items attempted	13.7 [±1.9]	11.1 [±2.2]	10.8 [±1.8]	13.9 [±1.9]
	Percentage of students with zero scores	76.5 [±3.0]	78.3 [±3.5]	81.2 [±2.7]	73.6 [±3]
Oral Reading Passage	Reading Fluency (correct words per minute)	2.1 [±.4]	2 [±.4]	1.4 [±.3]	2 [±.3]
	Percentage correct of items attempted	10.6 [±1.6]	10.5 [±2.0]	7.6 [±1.4]	10.4 [±1.5]
	Percentage of students with zero scores	71.9 [±3.1]	71.1 [±3.9]	79.7 [±2.8]	72.9 [±3.0]
Reading Comprehension	Percentage correct	4.4 [±1.1]	3.7 [±1.2]	3 [±.9]	4.4 [±1.1]
	Percentage correct of items attempted	4.5 [±1.1]	3.7 [±1.2]	3.1 [±.9]	4.5 [±1.1]
	Percentage of students with 80% comprehension	1.5 [±.9]	1.1 [±.0]	1.2 [±.8]	2.2 [±1.0]
	Percentage of students with zero scores	90 [±2.1]	91.3 [±2.4]	92.7 [±1.8]	91 [±2.0]
English Subtasks					
Oral Reading Passage	Reading Fluency (correct words per minute)	0 [±.1]	.1 [±.1]	0 [±0]	.1 [±.1]

Table 97. EGRA and EGMA Results for Grade 1 Students in Rwanda, by Subtask and Treatment Group

		Control		Treatment	
Reading subtasks	**Measure**	**Control-non-SBMT**	**Control-SBMT**	**Treatment-non-SBMT**	**Treatment-SBMT**
		n= 1636			
	Percentage correct of items attempted	.2 [±.3]	.4 [±.5]	0 [±0]	.2 [±.2]
	Percentage of students with zero scores	99.6 [±.5]	99.2 [±.9]	100 [±.]	99.4 [±.6]
Reading Comprehension	Percentage correct	0 [±.1]	0 [±0]	0 [±0]	0 [±.1]
	Percentage correct of items attempted	0 [±.1]	0 [±0]	0 [±0]	0 [±.1]
	Percentage of students with 80% comprehension	0 [.]	0 [.]	0 [.]	0 [.]
	Percentage of students with zero scores	99.9 [±.4]	100 [±.]	100 [±.]	99.9 [±.4]
Math Subtasks					
Quantity Discrimination	Percentage correct	50.3 [±1.8]	47.9 [±2.2]	46.3 [±1.8]	51.2 [±1.8]
	Percentage correct of items attempted	53.6 [±1.7]	51.4 [±2.0]	49.7 [±1.7]	54.4 [±1.7]
	Percentage of students with zero scores	6.4 [±1.7]	6.2 [±2.1]	7.2 [±1.8]	5.7 [±1.6]
Missing Number	Percentage correct	12.8 [±.8]	11.5 [±.9]	10.5 [±.8]	13.3 [±.8]
	Percentage correct of items attempted	20 [±1.1]	18.5 [±1.3]	16.8 [±1.1]	20.6 [±1.1]
	Percentage of students with zero scores	33 [±3.3]	34.8 [±4.1]	40.8 [±3.4]	31.1 [±3.2]
Basic Addition	Fluency (correct addition problems per minute)	2.6 [±.3]	2 [±.3]	2.5 [±.3]	2.8 [±.3]
	Percentage correct of items attempted	20.8 [±2.1]	16.6 [±2.3]	20.1 [±2.1]	22.3 [±2.1]
	Percentage of students with zero scores	55 [±3.5]	60.8 [±4.2]	58.1 [±3.4]	53.1 [±3.4]

Table 97. EGRA and EGMA Results for Grade 1 Students in Rwanda, by Subtask and Treatment Group

		Control		Treatment	
Reading subtasks	**Measure**	**Control-non-SBMT**	**Control-SBMT**	**Treatment-non-SBMT**	**Treatment-SBMT**
		n= 1636			
Basic Subtraction	Fluency (correct subtraction problems per minute)	1.3 [±.2]	1.1 [±.2]	1.4 [±.2]	1.5 [±.2]
	Percentage correct of items attempted	10.7 [±1.5]	9.6 [±1.8]	11.3 [±1.7]	11.6 [±1.6]
	Percentage of students with zero scores	73.1 [±3.1]	75.1 [±3.7]	75 [±3.0]	72.6 [±3.0]
Addition Level 2	Percentage correct	3 [±.8]	1.8 [±.7]	2.6 [±.7]	2.7 [±.7]
	Percentage correct of items attempted	3 [±.8]	1.8 [±.7]	2.6 [±.7]	2.7 [±.7]
	Percentage of students with zero scores	91.4 [±2.0]	94 [±2.1]	91.8 [±1.9]	91.4 [±1.9]
Subtraction Level 2	Percentage correct	.9 [±.4]	.5 [±.3]	1 [±.4]	1.1 [±.4]
	Percentage correct of items attempted	.9 [±.4]	.5 [±.3]	1 [±.4]	1.1 [±.4]
	Percentage of students with zero scores	96.4 [±1.3]	98.1 [±1.2]	96 [±1.4]	96.3 [±1.3]

3. CONCLUSION

3.1 Measurement Considerations

This study incorporates the use of well-established measures (such as EGRA and EGMA) alongside constructs that have been less commonly measured in low- and middle-income countries. For example, we developed new play attitude scales, as well as a new SEL tool (which went through a rigorous development and piloting process, prior to administration). We are encouraged by the reliability and quality of these measures but there is always room for improvement. Accordingly, we are in the process of reviewing and refining (or dropping) items that did not function as intended. In order to maintain comparability across time points, we will carefully track all changes.

Although the intent of this study was to collect data on all programs prior to the start of any school-level implementation, this was not always possible due to COVID-related delays. Therefore, all impact estimates at midline and endline will control for any baseline differences. Furthermore, our student-level estimates of impact are strengthened by the student-level longitudinal design of this study.

Student literacy and numeracy performance was low across all programs and all treatment groups, due to the fact that we tested grade 1 students. The highly skewed distributions were expected at baseline. The same measures will be administered at midline and endline, in order to measure growth in performance. There is significant room for improvement across all subtasks and we do not have concerns about floor or ceiling effects as we progress through the grades.

The majority of SEL scores were normally distributed and show room for growth. This is encouraging for the tool as a measure of impact for students as they progress through grades 2 and 3.

Across countries, teachers expressed a high level of confidence and self-efficacy in their ability to implement the LtP approaches. However, most teachers haven't yet received LtP training and may not yet fully understand what these approaches entail. Therefore, it is possible exposure to the program could have unintended impacts on teachers' confidence and attitudes toward the use of learning through play, even as they increase their use of playful practices in the classroom.

At midline and endline, analyses will include interactions among indicators, in order to determine which aspects of learning through play are most strongly associated with changes in desired outcomes. For example, it will be interesting to see if shifts in teachers' and headteachers' attitudes towards a "less-disciplined" classroom environment might be concurrent with changes in instructional practice, including increases in group work—and if this leads to subsequent improvements in SEL competencies, literacy and numeracy.

3.2 Implications and Recommendations

Learning through Play in the Classroom

Teachers in all programs incorporated many basic best teaching practices in their lessons (e.g., student monitoring and support, material preparation, good instructional pace, making connections and sharing ideas, inclusionary practices). However, we recommend providing additional space and opportunities to incorporate more playful pedagogies in the classroom, as laid out below.

Student Engagement. Classrooms in all three programs showed evidence of strong levels of student engagement. This is a promising starting point for these programs. ***We recommend that programs now focus on improving the quality of engagement through more playful practices and socially interactive methods.***

Classroom Climate. For the large majority of observed lessons, the classroom climate was found to be 'ordinary' (meaning that there was neither a noticeably positive nor negative tone). While it's great to see extremely limited use of negative language or physical punishment, ***we recommend that programs seek to increase levels of positive climate over the course the implementation, as well as increased perseverance (or recognition of effort and encouragement of goal setting among students).***

Teaching Approach. Across all three programs, teachers spent the largest proportion of time explaining concepts, modeling activities and monitoring students' understanding, while students spent the largest proportion of time listening. These are all valuable and important components of teaching and learning but there was a noticeably limited use of more interactive or student-focused actions. ***We recommend that programs provide teachers with additional opportunities to incorporate more interactive approaches into their teaching, such as the use of open-ended questioning and turn-taking.***

Material Usage. Materials provide a great opportunity for programs to increase the use of playful practices in the classroom. Chalkboards and oral language were the most commonly used materials in two of the three programs, while one program had a larger proportion of book usage. Across the board, this left little time for the use of other written materials (such as letter/number cards or pocket charts), manipulatives and alternative materials (e.g., crafts, toys, illustrations, space, etc.). ***We recommend an increased use of these types of alternative materials (as well as book usage), in order to support improved access and use of print and engagement among students.***

Material Availability. In order for students to engage with materials, it is essential for those materials to be available in the classroom. Although some materials were available, **we recommend all programs to carefully examine the availability of materials and consider providing additional resources where possible (e.g., teaching aids, supplementary readers, workbooks, etc.). Increasing the availability of these materials in classrooms will improve students access to print and could be important for improving their performance on academic outcomes.**

Group Work. Although the majority of teachers in all three programs noted the importance of group work for student learning, less than 10% of the total class time was spent with students working in large or small groups. ***We recommend increased usage of group work and/or learning corners in order to improve social interaction and iteration among students.***

Student Agency. Arguably the area with the greatest room for improvement in these programs is student agency. Nearly all of class time across programs was dedicated to activities that were initiated and directed by teachers. Whether student-initiated or teacher-initiated and student-directed, ***we recommend providing students with increased levels of control over how they engage in an activity or, at a minimum, with improved opportunities for asking questions, expressing opinions and practicing skills. Increased student agency is at the foundation of improving learning through play in the classroom.***

LtP Enabling Environment

Learning Environment. Students in all programs reported incidences of physical and verbal aggression by their teachers in the past week. Roughly half (45%-55% across countries) of students said that teachers yell at them when they get the answer wrong. Interestingly, while students reported physical and verbal punishment, they also largely reported that they feel supported and listened to by teachers. Furthermore, verbal and physical aggression was rarely seen in classroom observations (which was likely impacted by the presence of an outside observer in the classroom). ***We recommend that programs further investigate the degree to which children feel safe and supported in classrooms and focus on ensuring that there are more positive and constructive interactions between teachers and students.***

Teacher Support. Head teachers in all programs often claimed that they provide additional support to struggling teachers themselves, although there was variation in who assists them with identifying a struggling teacher in the first place. ***We recommend that programs consider further investigating head teachers' capacity and availability to support struggling teachers in their schools, while identifying alternative avenues for support through other system actors (e.g., coaches, meeting facilitators, etc.).***

Training and Materials. Teacher interviews underscored the ongoing importance of training as a major support teachers receive – as well as teacher meetings. While teachers' understanding of how to use playful approaches in the classroom was a common barrier noted by trainers and coaches, the largest concern was with the availability of play-based materials (or lack thereof). ***We recommend that programs continue to focus both on training teachers on playful pedagogies, as well as the development/provision of materials (or through advocacy for approaches that require few or no additional materials).***

Ministry Stakeholders. As anticipated, DEOs largely see their role as communicating program activities and monitoring implementation in schools. **We recommend leveraging district and sub county officials to clearly communicate expectations around learning through play (including what learning looks like, in terms of student agency and collaboration). Support from district officials can often play a large role in ensuring that activities are implemented as intended in schools.**

Perceptions of LtP

Attitudes Toward Play. Generally speaking, caregivers, teachers and education stakeholders are in support of LtP in theory but are less in support of a more student-centered, "free" classroom environment. This may be due in part to a misconception that student agency and more student-centered approaches will lead to a less disciplined and more chaotic classroom. ***We recommend that programs continue their work of orienting key stakeholders on the benefits of learning through play in the classroom and understanding their concerns/objections in order to provide solutions for improving attitudes toward specific playful practices.***

Student Holistic Outcomes

Conflict Resolution. Across programs, there was little evidence of advanced conflict resolution skills among students. Only 5-10% of students responded to another child in a conflict situation by asking for reasons for their behavior or by trying to influence them. Perhaps this is unsurprising among Grade 1 students. ***We recommend increasing activities that involve collaboration and interaction with other students to enhance children's conflict resolution skills.***

Conflict Resolution. Interestingly, Rwandan students were much more likely (nearly twice as likely as those in Ghana) to attribute a hostile intention to another student in a conflict situation. This may reflect difference in cultural and experience of students across countries. It may also be the result of different interpretations by assessors across countries. ***We recommend that that the Scaling LTP team explores this issue further in Rwanda, while we plan to recalibrate and standardize coding procedures across teams at midline.***

Creativity. There were low creativity scores across all programs, with only an average of about 4 ideas generated in the alternate uses test. ***We recommend increasing play-based approaches (such as the use of games and symbolic play) in order to improve creativity rates.***

Literacy and Numeracy. EGRA and EGMA results represent baseline estimates for students in Grade 1 for all programs. Therefore, results were relatively low across all subtasks (as expected). This led to a large proportion of zero scores (i.e., students who were unable to identify a single item correctly), particularly for higher order skills. These same skills will be measured at each time point and low baseline estimates should not be seen as a point of concern for these programs.

ANNEX A – ADDITIONAL TABLES

A.1 Ghana Stakeholder Interviews

A.1.1 Teacher Interview - Ghana

	Control %	Indirect %	Direct %	GALOP %
Have you received in-service training?	65.2	61.7	87.7	80.4
When was the last time you received in-service teacher training?				
Within the past few months	43.5	31.9	59.6	45.7
Within the past year	6.5	17	24.6	19.6
Within the past 3 years	13	6.4	1.8	15.2
More than 3 years ago	2.2	4.3	1.8	0
I have never received in-service training	0	2.1	0	0
Have not received in-service training	34.8	38.3	12.3	19.6
What content from your last in-service training did you find most helpful?				
New instructional approaches or strategies	43.5	51.1	66.7	58.7
Lesson plan development	39.1	25.5	56.1	45.7
Conducting or using student assessments	21.7	12.8	26.3	17.4
Classroom management	37	38.3	52.6	41.3
Parental involvement	4.3	8.5	22.8	8.7
Creating materials for the classroom	28.3	14.9	28.1	19.6
Using materials in the classroom	23.9	12.8	33.3	23.9
Other	2.2	4.3	1.8	4.3
Do you feel that your last in-service training led to improvements in your teaching?	58.7	59.6	84.2	78.3
Do you feel that your last in-service training led to improvements in student learning in your classroom?	58.7	61.7	86	76.1
Do you feel that your last in-service training prepared you to implement more student-centered or 'playful' approaches in the classroom?	56.5	57.4	84.2	78.3
Which of the following teacher materials do you receive?				
Teachers' guide	54.3	55.3	57.9	63
Teaching aids (e.g., letter cards, charts, etc.)	34.8	31.9	43.9	45.7
Teacher's version of textbook	21.7	17	17.5	21.7
Lesson plans (or sample lessons)	52.2	36.2	54.4	54.3
Furniture (small tables for groups etc.)	10.9	8.5	1.8	15.2
Play-based materials:	10.9	10.6	45.6	32.6
None	19.6	17	5.3	15.2
Which of the following student materials do you receive?				
Textbooks	30.4	25.5	22.8	26.1
Supplementary readers	21.7	19.1	31.6	34.8
Student exercise books	17.4	10.6	33.3	39.1

	Control %	Indirect %	Direct %	GALOP %
Stationary (notebooks etc.)	2.2	2.1	21.1	10.9
Play-based materials (toys crafts etc.)	6.5	8.5	33.3	21.7
None	47.8	57.4	33.3	34.8
Do you receive support visits from a coach or mentor?	58.7	38.3	59.6	76.1
Who conducts these coaching/mentoring visits?				
Head Teacher/School Principal	26.1	12.8	14	41.3
District Education Officer/ District Education Director	6.5	8.5	10.5	8.7
Fellow Teacher/Mentor Teacher	4.3	4.3	1.8	6.5
SISO/ School Improvement Support Officer (formerly Circuit Supervisors)	15.2	12.8	22.8	17.4
Curriculum Lead/Teacher Champion	6.5	0	7	0
Other	0	0	3.5	2.2
No visits conducted	41.3	61.7	40.4	23.9
How often do you receive coaching/mentoring visits?				
More than once per month	37	10.6	29.8	30.4
Once per month	10.9	14.9	15.8	19.6
A few times per year	10.9	10.6	14	23.9
Once per year or less	0	2.1	0	2.2
Of the supports identified you receive in the coach/mentor visitor, which one is the most useful?				
Instructional approaches	23.9	14.9	12.3	21.7
Classroom management	8.7	4.3	12.3	15.2
Conducting/using student assessments	2.2	0	1.8	6.5
Creating student materials	4.3	4.3	0	0
Pair/group work for students	2.2	2.1	5.3	2.2
Increasing play in my classroom	6.5	6.4	15.8	15.2
Making my classroom more interactive/engaging for students	10.9	6.4	12.3	15.2
Does not receive supports	41.3	61.7	40.4	23.9
Has your teaching practice improved due to the coaching/mentoring you've received?	58.7	38.3	59.6	76.1
What about the coaching/mentoring made you change your teaching?				
The coach gave me the confidence to improve my teaching	39.1	25.5	31.6	41.3
The coach improved my skills as a teacher	45.7	27.7	35.1	37
The coach helped me increase 'play' in the classroom	21.7	21.3	38.6	39.1
The coach was very supportive/friendly	23.9	14.9	26.3	26.1
Coach observes my teaching and provides useful feedback	23.9	23.4	28.1	30.4
Coach models instruction	8.7	10.6	12.3	10.9
Coach and I discuss progress from last observation	10.9	12.8	12.3	6.5
Do teachers in your school participate in teacher meetings (such as Community of Practice or Community of Learning)?	60.9	76.6	89.5	84.8
How often do you attend [teacher COP/COL] meetings?				
Never	0	0	1.8	4.3
Once per year	0	10.6	7	4.3
A few times per year	32.6	29.8	38.6	32.6

	Control %	Indirect %	Direct %	GALOP %
Once per month	6.5	17	8.8	10.9
More than once per month	21.7	19.1	33.3	32.6
No meetings take place	39.1	23.4	10.5	15.2
What do you find most useful from these meetings?				
Feedback on how to improve	39.1	53.2	64.9	50
Discussions with other teachers	30.4	38.3	54.4	32.6
Having someone answer my questions	8.7	25.5	12.3	13
Time to practice teaching	15.2	14.9	15.8	6.5
Learning new information	37	53.2	64.9	41.3
Reminder/refresher of skills/information/approaches	21.7	23.4	22.8	23.9
Developing friendships/networks with other teachers	10.9	23.4	31.6	13
Other	0	0	1.8	0
Do you discuss how your own classroom instruction is going in these meetings?				
Never	0	6.4	8.8	4.3
Rarely	15.2	8.5	3.5	8.7
Sometimes	39.1	38.3	50.9	45.7
Most of the time	2.2	19.1	17.5	15.2
Always	4.3	4.3	7	6.5
No meetings take place	39.1	23.4	12.3	19.6
Do you discuss play-based strategies in these meetings?				
Never	6.5	6.4	5.3	10.9
Rarely	13	14.9	5.3	8.7
Sometimes	21.7	29.8	22.8	28.3
Most of the time	15.2	23.4	36.8	19.6
Always	4.3	2.1	17.5	13
No meetings take place	39.1	23.4	12.3	19.6

A.1.2 Coach Interview – Ghana

	Number of Coaches (out of 28)
Have you begun working as a coach/mentor for P3?	28
What materials, if any, does the [program] provide to help you work with the teachers as a coach?	
I haven't received any materials	6
Tablet	1
Observation tools	8
Teachers' guides	6
Student books	2
Coaching/Mentoring manuals	12
Videos	3
Other	9

A.1.3 Trainer Interview – Ghana

	Number of Trainers (out of 7)
How does the frequency of program trainings compare to other teacher trainings you've been involved in?	
About the same	1
Program trainings more frequent	6
How does the organization of program trainings compare to others?	
About the same	1
Program trainings more organized	6
How do the program training materials compare to others?	
Program trainings better/more materials	7
How does teacher participation in program trainings compare to others?	
About the same	1
Program trainings had more participation (e.g., more discussion practice group work; less lecture)	6
How do the allowances of program trainings compare to others?	
About the same	1
Program better allowances (transport etc.)	5
Program fewer allowances (transport etc.)	1
How do the teaching strategies from program trainings compare to others?	
More focus on playful interactive or student-directed activities in program trainings	7
For which of the following areas do you provide training to teachers?	
Instructional approaches	4
Classroom management	7
Conducting/using student assessments	5
Creating student materials	4
Pair/group work for students	6
Increasing play in the classroom	6
Making their classroom more interactive/engaging for students	7
Which of these areas do you feel is the most effective/useful for teachers?	
Classroom management	1
Pair/group work for students	1
Increasing play in the classroom	2
Making their classroom more interactive/engaging for students	3
Which training method do you use most in program trainings?	
Discussion	1
Modeling/demonstration	1
Small-group practice	5
Which method do you think is the most useful to teach a new instructional approach?	
Discussion	2
Modeling/demonstration	1

Small-group practice	4
Name 2-3 things that you would change to improve the training.	
Include more activities	2
Include more practices	4
Less lecture	1
Better allowances (transport etc.)	2
Increase use of technology	5
Include more discussion	2
More frequent trainings	3
Other	1

A.1.4 DEO Interview – Ghana

	Number of Officers (Out of 10)
Does the curriculum used in your primary schools include play-based learning activities and teaching methods?	10
Do primary teachers in your schools have materials for play-based learning in the classroom?	4
Have you received training or orientation on "learning through play" or play-based learning?	10
Who determined this role?	
Ministry leadership determined my [office's] role	2
The program determined my [office's] role	3
I determined my [office's] role	2
The ministry and program jointly agreed on the role I [my office] would play	3
How was this role communicated?	
I received an official letter about my role from regional/national ministry leadership	5
I was informed about my role during ministry meetings	1
I was informed about my role during meetings with the program	5
Have the responsibilities of your office changed as a result of P3?	2
Have the responsibilities of [sub-district] staff changed as a result of P3?	4

A.1.5 Caregiver Interview – Ghana

	Control %	Indirect %	Direct %	GALOP %
How many times in the last month have you attended a PTA meeting?				
None/0 times	45.1	51.5	40.9	41.3
Once	35.2	30.9	31.3	31.5
Twice	13.2	9.3	22.6	25
Three times	5.5	5.2	4.3	2.2
Four times	1.1	2.1	0	0
Five or more times	0	1	.9	0
How many times in the last month have you met with your child's teacher?				
None/0 times	30.8	44.3	50.4	39.1

	Control %	Indirect %	Direct %	GALOP %
Once	27.5	29.9	27	30.4
Twice	22	9.3	11.3	17.4
Three times	7.7	6.2	6.1	4.3
Four times	0	5.2	2.6	3.3
Five or more times	12.1	5.2	2.6	5.4
How many times in the last month have you volunteered in the classroom?				
None/0 times	81.3	95.9	93.9	87
Once	8.8	2.1	3.5	6.5
Twice	6.6	0	0	2.2
Three times	2.2	1	1.7	3.3
Four times	0	0	0	1.1
Five or more times	1.1	1	.9	0
How many times in the last month have you attended a school event (fair, sports event, cultural event, fundraiser, etc.)?				
None/0 times	84.6	92.8	81.7	81.5
Once	11	5.2	14.8	10.9
Twice	1.1	1	1.7	4.3
Three times	2.2	0	.9	1.1
Four times	0	1	0	0
Five or more times	1.1	0	.9	2.2
How many times in the last month have you talked to the head teacher?				
None/0 times	50.5	58.8	62.6	53.3
Once	28.6	22.7	17.4	18.5
Twice	8.8	9.3	12.2	10.9
Three times	4.4	3.1	6.1	7.6
Four times	2.2	6.2	0	1.1
Five or more times	5.5	0	1.7	8.7
How often do you speak to your child's teacher?				
Never	22	28.9	27.8	27.2
Daily	5.5	3.1	5.2	6.5
Weekly	28.6	15.5	26.1	23.9
Monthly	15.4	16.5	13	17.4
A few times a year	28.6	36.1	27.8	25

A.2 Kenya Stakeholder Interviews

A.2.1 Teacher Interview – Kenya

	Control %	Phase 1 %	Phase 2 %
Have you received in-service training?	61.2	62.5	44
When was the last time you received in-service teacher training?			
Within the past few months	28.6	50	38
Within the past year	14.3	10	4
Within the past 3 years	16.3	0	0
More than 3 years ago	2	2.5	2
I have never received in-service training	0	0	0
Have not received in-service training	38.8	37.5	56
Do you feel that your last in-service training led to improvements in your teaching?	59.2	60	44
Do you feel that your last in-service training led to improvements in student learning in your classroom?	59.2	60	44
Do you feel that your last in-service training prepared you to implement more student-centered or 'playful' approaches in the classroom?	49	60	40
Which of the following teacher materials do you receive?			
Teachers' guide	67.3	65	60
Teaching aids (e.g., letter cards, charts, etc.)	49	37.5	44
Teacher's version of textbook	53.1	45	46
Lesson plans (or sample lessons)	24.5	42.5	32
Furniture (small tables for groups etc.)	14.3	17.5	12
Play-based materials:	14.3	7.5	12
None	24.5	27.5	32
Which of the following teacher materials do you receive?			
Textbooks	67.3	57.5	56
Supplementary readers	42.9	22.5	18
Student exercise books	44.9	42.5	36
Stationary (notebooks etc.)	32.7	22.5	20
Play-based materials (toys crafts etc.)	10.2	7.5	8
None	26.5	32.5	34
Do you receive support visits from a coach or mentor?	40.8	52.5	40
Who conducts these coaching/mentoring visits?			
Head Teacher/School Principal	16.3	25	6
District Education Officer/ District Education Director	12.2	17.5	10
Fellow Teacher/Mentor Teacher	8.2	2.5	18
SISO/ School Improvement Support Officer (formerly Circuit Supervisors)	0	0	0
Curriculum Lead/Teacher Champion	0	0	0
Other	4.1	7.5	6
No visits conducted	59.2	47.5	60
How often do you receive coaching/mentoring visits?			
More than once per month	6.1	17.5	12
Once per month	14.3	17.5	8
A few times per year	18.4	17.5	18

	Control %	Phase 1 %	Phase 2 %
Once per year or less	2	0	0
Of the supports identified you receive in the coach/mentor visitor, which one is the most useful?			
Instructional approaches	12.2	15	6
Classroom management	4.1	15	2
Conducting/using student assessments	4.1	2.5	2
Creating student materials	14.3	7.5	10
Pair/group work for students	2	0	4
Increasing play in my classroom	2	2.5	0
Making my classroom more interactive/engaging for students	2	10	14
Does not receive supports	59.2	47.5	60
Has your teaching practice improved due to the coaching/mentoring you've received?	38.8	52.5	38
What about the coaching/mentoring made you change your teaching?			
The coach gave me the confidence to improve my teaching	22.4	35	28
The coach improved my skills as a teacher	24.5	50	38
The coach helped me increase 'play' in the classroom	10.2	27.5	16
The coach was very supportive/friendly	10.2	30	26
Coach observes my teaching and provides useful feedback	16.3	22.5	22
Coach models instruction	8.2	17.5	26
Coach and I discuss progress from last observation	12.2	15	18
Other	2	2.5	0
Do teachers in your school participate in teacher meetings (such as Community of Practice or Community of Learning)?	81.6	77.5	64
How often do you attend [teacher COP/COL] meetings?			
Never	4.1	5	2
Once per year	4.1	15	0
A few times per year	42.9	17.5	36
Once per month	18.4	30	16
More than once per month	12.2	10	10
No meetings take place	18.4	22.5	36
What do you find most useful from these meetings?			
Feedback on how to improve	46.9	55	26
Discussions with other teachers	49	45	32
Having someone answer my questions	28.6	30	12
Time to practice teaching	18.4	27.5	14
Learning new information	28.6	47.5	34
Reminder/refresher of skills/information/approaches	26.5	32.5	12
Developing friendships/networks with other teachers	10.2	22.5	12
Other	6.1	0	4
Do you discuss how your own classroom instruction is going in these meetings?			
Never	4.1	12.5	8
Rarely	2	2.5	4
Sometimes	34.7	12.5	20

	Control %	Phase 1 %	Phase 2 %
Most of the time	20.4	17.5	18
Always	16.3	27.5	12
No meetings take place	22.4	27.5	38
Do you discuss play-based strategies in these meetings?			
Never	30.6	20	28
Rarely	2	2.5	8
Sometimes	26.5	25	12
Most of the time	8.2	10	4
Always	10.2	15	10
No meetings take place	22.4	27.5	38

A.2.2 Trainer Interview – Kenya

	Number of Trainers (out of 14)
How does the frequency of program trainings compare to other teacher trainings you've been involved in?	
About the same	2
Program trainings more frequent	4
How does the organization of program trainings compare to others?	
About the same	0
Program trainings more organized	8
How do the program training materials compare to others?	
Program trainings better/more materials	8
How does teacher participation in program trainings compare to others?	
About the same	0
Program trainings had more participation (e.g., more discussion practice group work; less lecture)	8
How do the allowances of program trainings compare to others?	
About the same	3
Program better allowances (transport etc.)	0
Program fewer allowances (transport etc.)	6
How do the teaching strategies from program trainings compare to others?	
More focus on playful interactive or student-directed activities in program trainings	9
For which of the following areas do you provide training to teachers?	
Instructional approaches	5
Classroom management	3
Conducting/using student assessments	2
Creating student materials	5
Pair/group work for students	5
Increasing play in the classroom	7
Making their classroom more interactive/engaging for students	7
Which of these areas do you feel is the most effective/useful for teachers?	

Classroom management	0
Pair/group work for students	0
Increasing play in the classroom	1
Making their classroom more interactive/engaging for students	4
Which training method do you use most in program trainings?	
Discussion	1
Modeling/demonstration	2
Small-group practice	1
Which method do you think is the most useful to teach a new instructional approach?	
Discussion	0
Modeling/demonstration	6
Small-group practice	1
Name 2-3 things that you would change to improve the training?	
Include more activities	2
Include more practices	3
Less lecture	2
Better allowances (transport etc.)	4
Increase use of technology	2
Include more discussion	3
More frequent trainings	6
Other	1

A.2.3 DEO Interview – Kenya

	Number of Officers (out of 11)
Does the curriculum used in your primary schools include play-based learning activities and teaching methods?	9
Do primary teachers in your schools have materials for play-based learning in the classroom?	7
Have you received training or orientation on "learning through play" or play-based learning?	5
Who determined this role?	
Ministry leadership determined my [office's] role	0
The program determined my [office's] role	1
I determined my [office's] role	0
The ministry and program jointly agreed on the role I [my office] would play	2
How was this role communicated?	
I received an official letter about my role from regional/national ministry leadership	1
I was informed about my role during ministry meetings	0
I was informed about my role during meetings with the program	2
Have the responsibilities of your office changed as a result of the TuKu program?	0
Have the responsibilities of [sub-district] staff changed as a result of the TuKu program?	0

A.2.4 Caregiver Interview – Kenya

	Control %	Phase 1 %	Phase 2 %
How many times in the last month have you attended a PTA meeting?			
None/0 times	40.6	42.5	38.4
Once	29.2	28.7	26.3
Twice	20.8	18.7	17.2
Three times	7.3	3.8	9.1
Four times	2.1	6.3	3
Five or more times	0	0	6.1
How many times in the last month have you met with your child's teacher?			
None/0 times	12.5	10	9.1
Once	20.8	11.3	18.2
Twice	27.1	21.3	30.3
Three times	18.8	11.3	13.1
Four times	6.3	7.5	7.1
Five or more times	14.6	38.8	22.2
How many times in the last month have you volunteered in the classroom?			
None/0 times	75	52.5	49.5
Once	3.1	13.8	13.1
Twice	4.2	3.8	10.1
Three times	1	2.5	6.1
Four times	3.1	5	3
Five or more times	13.5	22.5	18.2
How many times in the last month have you attended a school event (fair, sports event, cultural event, fundraiser, etc.)?			
None/0 times	86.5	80	64.6
Once	5.2	15	18.2
Twice	7.3	3.8	9.1
Three times	1	0	3
Four times	0	1.2	1
Five or more times	0	0	4
How many times in the last month have you talked to the head teacher?			
None/0 times	37.5	31.3	32.3
Once	31.2	26.3	23.2
Twice	16.7	18.7	12.1
Three times	7.3	5	10.1
Four times	2.1	3.8	6.1
Five or more times	5.2	15	16.2
How often do you speak to your child's teacher?			
Never	3.1	1.2	3
Daily	18.8	26.3	18.2
Weekly	42.7	42.5	47.5
Monthly	28.1	27.5	19.2
A few times a year	7.3	2.5	12.1

A.3 Rwanda Stakeholder Interviews

A.3.1 Teacher Interview – Rwanda

	Control Non-SBMT (%)	Control SBMT (%)	Treatment Non-SBMT (%)	Treatment SBMT (%)
Have you received in-service training?	93.9	83.9	94.1	98
When was the last time you received in-service teacher training?				
Within the past few months	46.9	38.7	51	49
Within the past year	22.4	25.8	9.8	19.6
Within the past 3 years	22.4	19.4	23.5	27.5
More than 3 years ago	2	0	9.8	2
I have never received in-service training	0	0	0	0
Have not received in-service training	6.1	16.1	5.9	2
Do you feel that your last in-service training led to improvements in your teaching?	93.9	83.9	90.2	98
Do you feel that your last in-service training led to improvements in student learning in your classroom?	93.9	83.9	88.2	98
Do you feel that your last in-service training prepared you to implement more student-centered or 'playful' approaches in the classroom?	44.9	48.4	54.9	62.7
Which of the following teacher materials do you receive?				
Teachers' guide	87.8	93.5	96.1	94.1
Teaching aids (e.g., letter cards, charts, etc.)	83.7	80.6	88.2	84.3
Teacher's version of textbook	89.8	87.1	88.2	88.2
Lesson plans (or sample lessons)	89.8	83.9	90.2	86.3
Furniture (small tables for groups etc.)	16.3	6.5	15.7	17.6
Play-based materials:	4.1	12.9	15.7	13.7
None	0	3.2	0	0
Which of the following teacher materials do you receive?				
Textbooks	87.8	87.1	94.1	88.2
Supplementary readers	71.4	77.4	86.3	80.4
Student exercise books	71.4	77.4	76.5	68.6
Stationary (notebooks etc.)	61.2	67.7	56.9	62.7
Play-based materials (toys crafts etc.)	6.1	9.7	5.9	5.9
None	2	3.2	0	0
Do you receive support visits from a coach or mentor?	61.2	71	78.4	72.5
Who conducts these coaching/mentoring visits?				
Head Teacher/School Principal	44.9	48.4	62.7	49
District Education Officer/ District Education Director	2	0	0	0
Fellow Teacher/Mentor Teacher	14.3	22.6	11.8	19.6
SISO/ School Improvement Support Officer (formerly Circuit Supervisors)	0	0	0	0
Curriculum Lead/Teacher Champion	0	0	0	0
Other	0	0	3.9	3.9
No visits conducted	38.8	29	21.6	27.5

	Control Non-SBMT (%)	Control SBMT (%)	Treatment Non-SBMT (%)	Treatment SBMT (%)
How often do you receive coaching/mentoring visits?				
More than once per month	38.8	35.5	37.3	45.1
Once per month	40.8	51.6	51	45.1
A few times per year	10.2	9.7	3.9	7.8
Once per year or less	2	0	2	2
Of the supports identified you receive in the coach/mentor visitor, which one is the most useful?				
Instructional approaches	61.2	74.2	58.8	47.1
Classroom management	0	6.5	5.9	9.8
Conducting/using student assessments	8.2	3.2	7.8	9.8
Creating student materials	6.1	6.5	7.8	9.8
Pair/group work for students	6.1	0	3.9	3.9
Increasing play in my classroom	2	3.2	2	0
Making my classroom more interactive/engaging for students	10.2	0	11.8	17.6
Does not receive supports	0	0	0	0
Has your teaching practice improved due to the coaching/mentoring you've received?	93.9	93.5	98	98
What about the coaching/mentoring made you change your teaching?				
The coach gave me the confidence to improve my teaching	77.6	61.3	78.4	82.4
The coach improved my skills as a teacher	83.7	61.3	84.3	86.3
The coach helped me increase 'play' in the classroom	28.6	16.1	27.5	27.5
The coach was very supportive/friendly	59.2	32.3	49	41.2
Coach observes my teaching and provides useful feedback	71.4	45.2	70.6	66.7
Coach models instruction	49	54.8	62.7	51
Coach and I discuss progress from last observation	57.1	29	60.8	49
Other	8.2	12.9	3.9	7.8
Do teachers in your school participate in teacher meetings (such as Community of Practice or Community of Learning)?	93.9	96.8	100	98
How often do you attend [teacher COP/COL] meetings?				
Never	0	3.2	0	0
Once per year	0	3.2	2	0
A few times per year	6.1	0	5.9	11.8
Once per month	53.1	61.3	52.9	45.1
More than once per month	34.7	29	39.2	41.2
No meetings take place	6.1	3.2	0	2
What do you find most useful from these meetings?				
Feedback on how to improve	79.6	87.1	90.2	86.3
Discussions with other teachers	87.8	67.7	90.2	84.3
Having someone answer my questions	36.7	22.6	37.3	33.3
Time to practice teaching	53.1	41.9	60.8	52.9
Learning new information	71.4	54.8	74.5	66.7
Reminder/refresher of skills/information/approaches	53.1	41.9	52.9	43.1
Developing friendships/networks with other teachers	63.3	51.6	58.8	64.7

	Control Non-SBMT (%)	Control SBMT (%)	Treatment Non-SBMT (%)	Treatment SBMT (%)
Other	2	6.5	2	2
Do you discuss how your own classroom instruction is going in these meetings?				
Never	0	0	0	0
Rarely	6.1	6.5	5.9	3.9
Sometimes	20.4	12.9	21.6	23.5
Most of the time	44.9	51.6	49	49
Always	22.4	22.6	23.5	21.6
No meetings take place	6.1	6.5	0	2
Do you discuss play-based strategies in these meetings?				
Never	55.1	54.8	60.8	68.6
Rarely	18.4	9.7	19.6	0
Sometimes	16.3	12.9	7.8	15.7
Most of the time	4.1	12.9	9.8	11.8
Always	0	3.2	2	2
No meetings take place	6.1	6.5	0	2

A.3.2 Coach Interview – Rwanda

	Number of Coaches (out of 136*)
What materials, if any, does the program provide to help you work with the teachers as a coach?	
I haven't received any materials	30
Tablet	36
Observation tools	49
Teachers' guides	72
Student books	74
Coaching/Mentoring manuals	65
Videos	23
Other	17

**Sub-sample of respondents who said they've received supports from the program at baseline*

A.3.3 Trainer Interview – Rwanda

	Number of Trainers (out of 32*)
How does the frequency of program trainings compare to other teacher trainings you've been involved in?	
About the same	4
Program trainings more frequent	2
How does the organization of program trainings compare to others?	
About the same	6
Program trainings more organized	7
How do the program training materials compare to others?	

Program trainings better/more materials	7
How does teacher participation in program trainings compare to others?	
About the same	5
Program trainings had more participation (e.g., more discussion practice group work; less lecture)	8
How do the allowances of program trainings compare to others?	
About the same	4
Program better allowances (transport etc.)	5
Program fewer allowances (transport etc.)	8
How do the teaching strategies from program trainings compare to others?	
More focus on playful interactive or student-directed activities in program trainings	13
For which of the following areas do you provide training to teachers?	
Instructional approaches	21
Classroom management	20
Conducting/using student assessments	23
Creating student materials	19
Pair/group work for students	16
Increasing play in the classroom	13
Making their classroom more interactive/engaging for students	14
Which of these areas do you feel is the most effective/useful for teachers?	
Classroom management	1
Pair/group work for students	1
Increasing play in the classroom	3
Making their classroom more interactive/engaging for students	1
Which training method do you use most in program trainings?	
Discussion	18
Modeling/demonstration	2
Small-group practice	7
Which method do you think is the most useful to teach a new instructional approach?	
Discussion	10
Modeling/demonstration	7
Small-group practice	7
Name 2-3 things that you would change to improve the training?	
Include more activities	11
Include more practices	11
Less lecture	5
Better allowances (transport etc.)	19
Increase use of technology	15
Include more discussion	9
More frequent trainings	11
Other	1

**Sub-sample of respondents who said they've received supports from the program at baseline*

A.3.4 DEO Interview – Rwanda

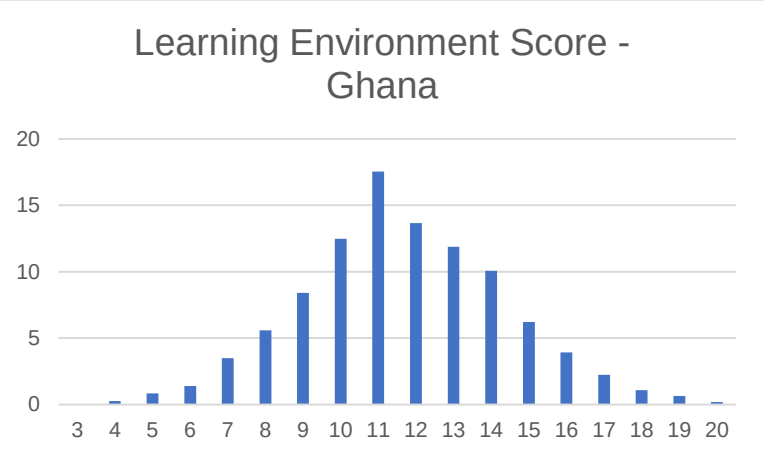
	Number of Officers (out of 11)
Does the curriculum used in your primary schools include play-based learning activities and teaching methods?	8
Do primary teachers in your schools have materials for play-based learning in the classroom?	2
Have you received training or orientation on "learning through play" or play-based learning?	2
Who determined this role?	
Ministry leadership determined my [office's] role	1
The program determined my [office's] role	0
I determined my [office's] role	0
The ministry and program jointly agreed on the role I [my office] would play	0
How was this role communicated?	
I received an official letter about my role from regional/national ministry leadership	1
I was informed about my role during ministry meetings	0
I was informed about my role during meetings with the program	0
Have the responsibilities of your office changed as a result of Scaling LtP?	0
Have the responsibilities of [sub-district] staff changed as a result of Scaling LtP?	0

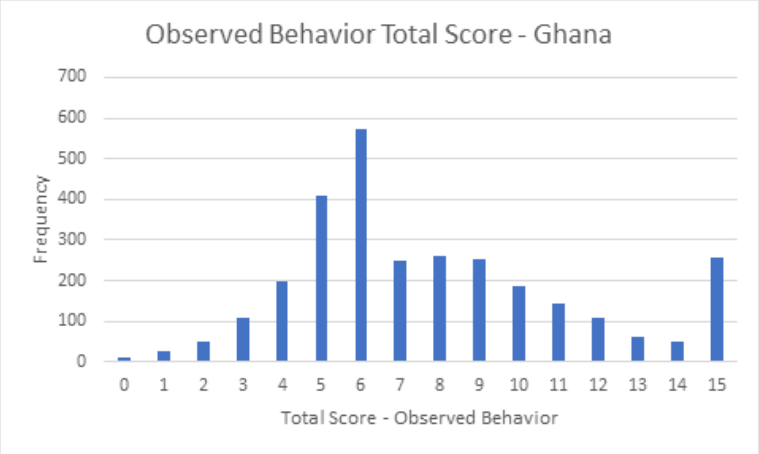
A.3.5 Caregiver Interview – Rwanda

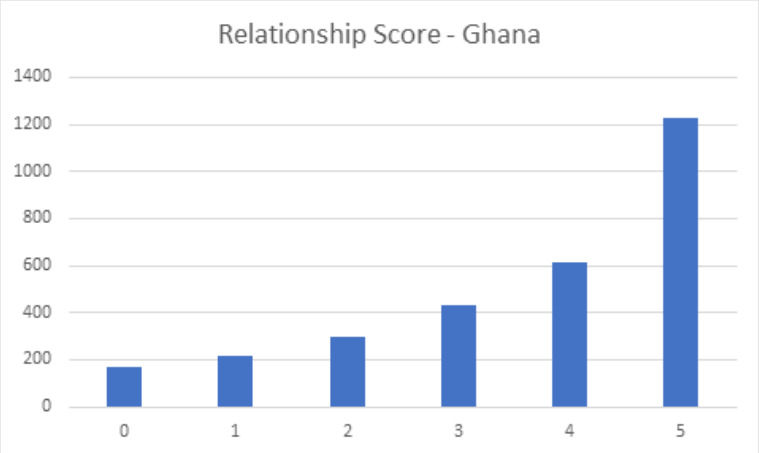
	Control Non-SBMT (%)	Control SBMT (%)	Treatment Non-SBMT (%)	Treatment SBMT (%)
How many times in the last month have you attended a PTA meeting?				
None/0 times	55.6	48.4	40.8	45.6
Once	26.3	30.6	32.7	33
Twice	11.1	12.9	17.3	11.7
Three times	1	6.5	3.1	3.9
Four times	5.1	0	1	2.9
Five or more times	1	1.6	5.1	2.9
How many times in the last month have you met with your child's teacher?				
None/0 times	18.2	21	22.4	16.5
Once	16.2	24.2	15.3	20.4
Twice	11.1	21	26.5	18.4
Three times	18.2	12.9	10.2	15.5
Four times	11.1	1.6	3.1	0
Five or more times	25.3	19.4	22.4	29.1
How many times in the last month have you volunteered in the classroom?				
None/0 times	62.6	62.9	50	57.3
Once	18.2	8.1	25.5	12.6
Twice	11.1	9.7	11.2	8.7
Three times	3	9.7	3.1	9.7

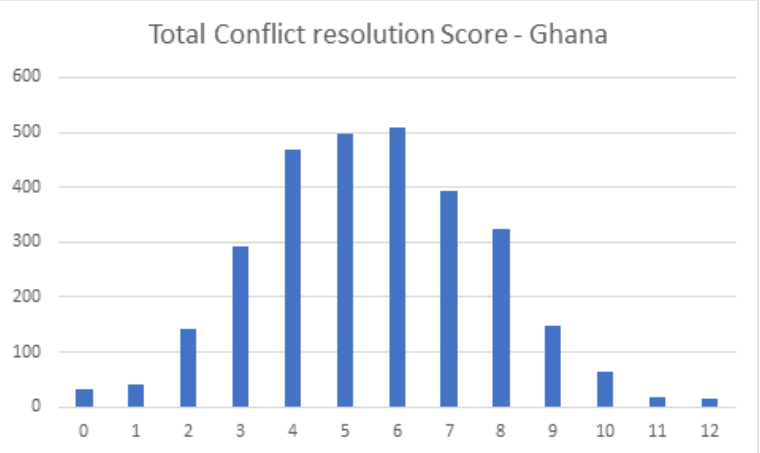
	Control Non-SBMT (%)	Control SBMT (%)	Treatment Non-SBMT (%)	Treatment SBMT (%)
Four times	0	4.8	2	2.9
Five or more times	5.1	4.8	8.2	8.7
How many times in the last month have you attended a school event (fair, sports event, cultural event, fundraiser, etc.)?				
None/0 times	80.8	79	75.5	76.7
Once	14.1	14.5	15.3	13.6
Twice	3	0	5.1	6.8
Three times	0	3.2	2	1.9
Four times	1	0	1	0
Five or more times	1	3.2	1	1
How many times in the last month have you talked to the head teacher?				
None/0 times	44.4	38.7	35.7	43.7
Once	13.1	19.4	23.5	18.4
Twice	11.1	9.7	16.3	13.6
Three times	11.1	9.7	8.2	6.8
Four times	5.1	9.7	4.1	3.9
Five or more times	15.2	12.9	12.2	13.6
How often do you speak to your child's teacher?				
Never	7.1	6.5	4.1	3.9
Daily	18.2	17.7	21.4	32
Weekly	43.4	48.4	41.8	26.2
Monthly	25.3	22.6	26.5	28.2
A few times a year	6.1	4.8	6.1	9.7

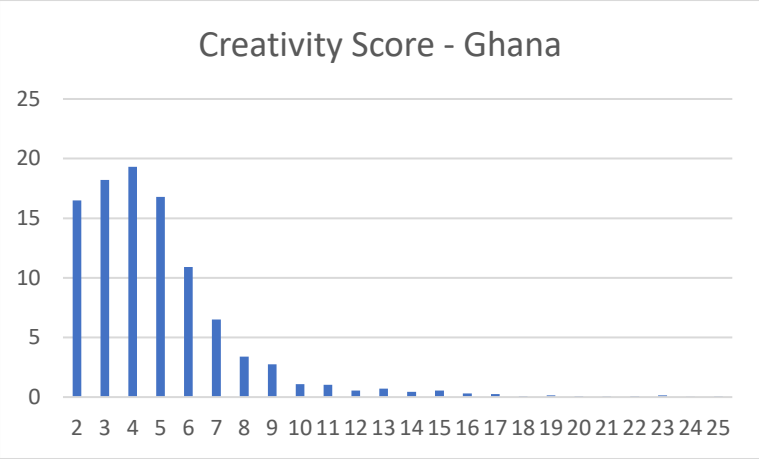
A.4 Ghana SEL Score Distributions



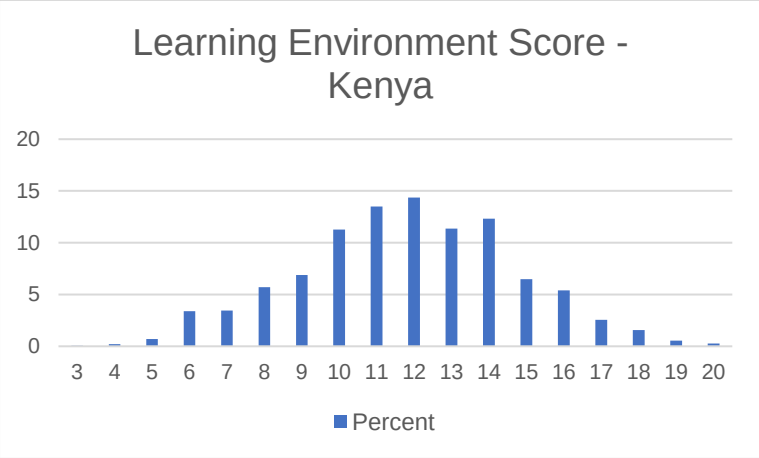


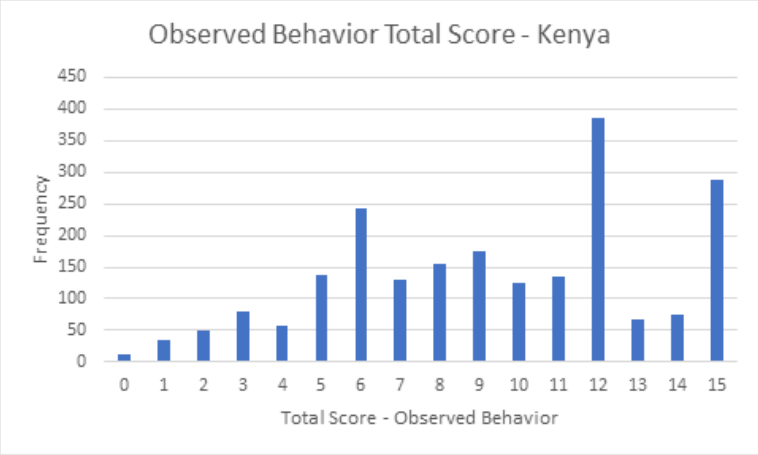


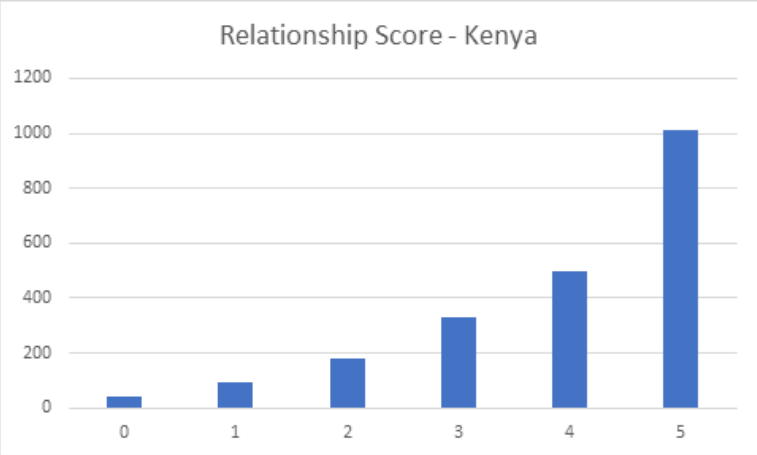


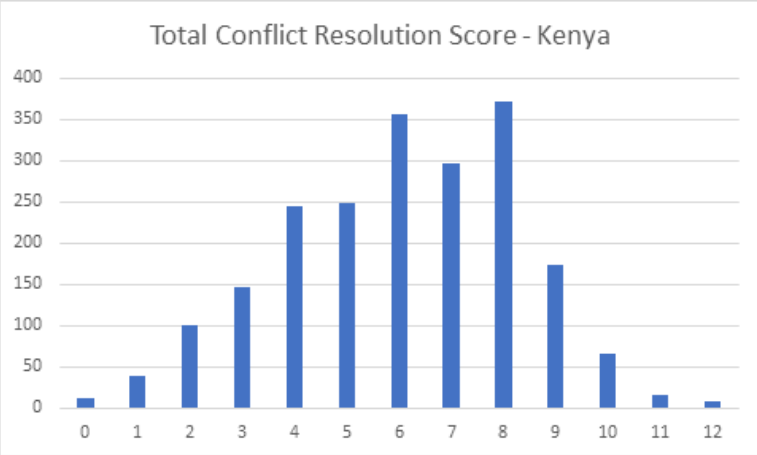


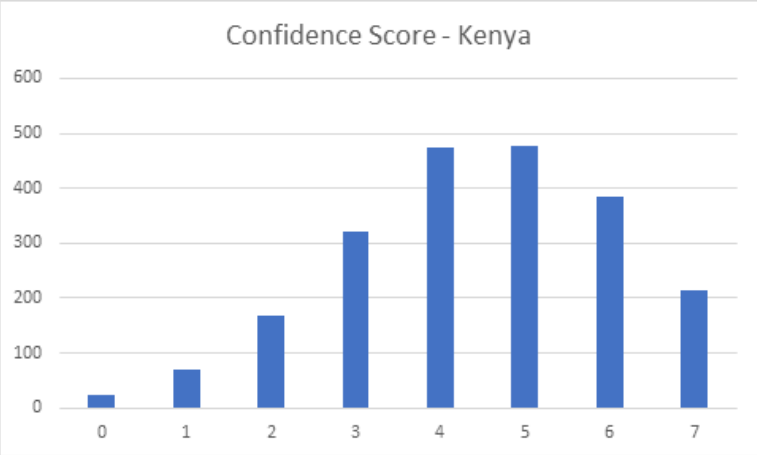
A.5 Kenya SEL Score Distributions

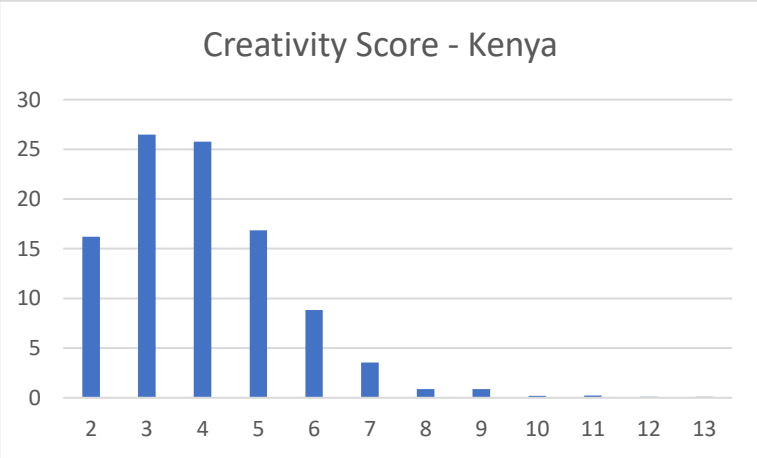




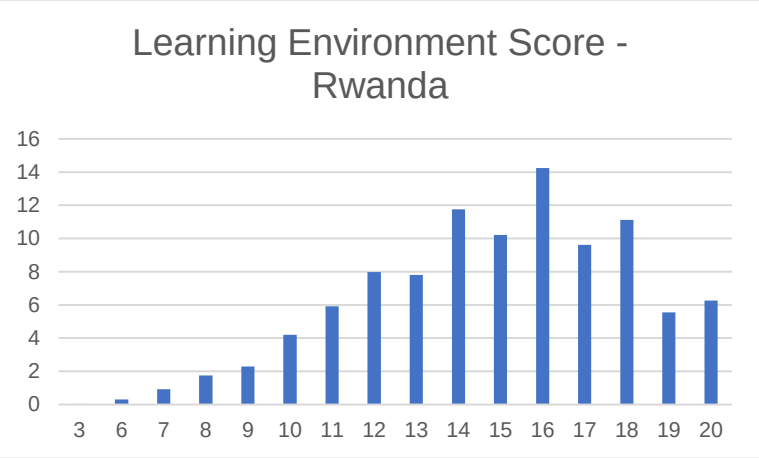

Relationship Score - Kenya
1200
1000
800
600
400
200
0
0 1 2 3 4 5

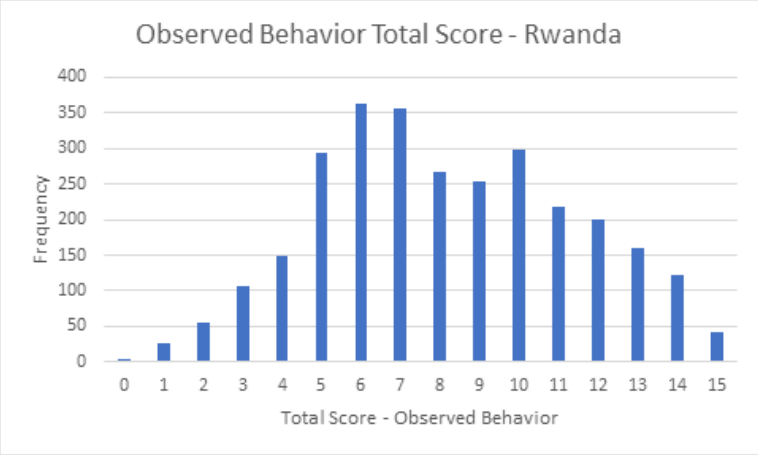

Total Conflict Resolution Score - Kenya
400
350
300
250
200
150
100
50
0
0 1 2 3 4 5 6 7 8 9 10 11 12

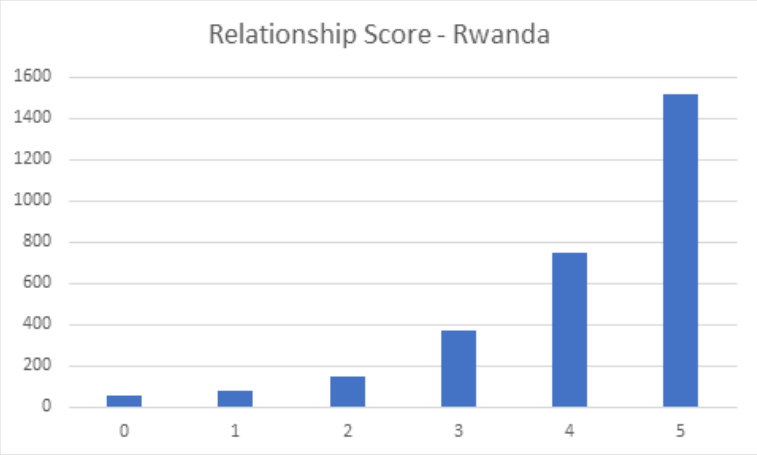

Confidence Score - Kenya
600
500
400
300
200
100
0
0 1 2 3 4 5 6 7

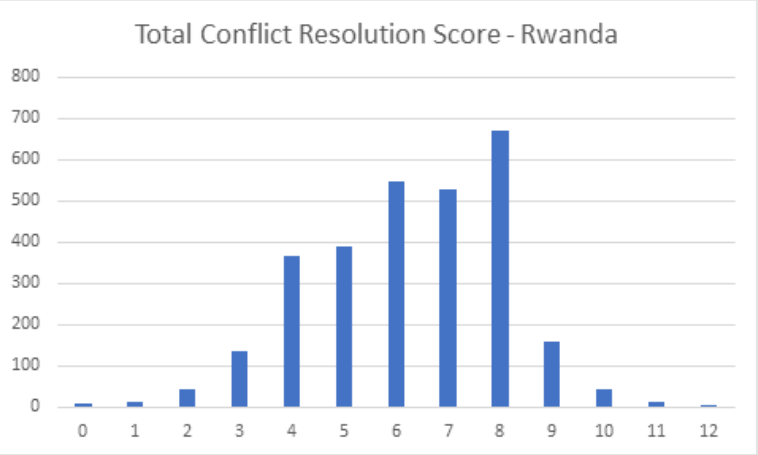

Creativity Score - Kenya
30
25
20
15
10
5
0
2 3 4 5 6 7 8 9 10 11 12 13

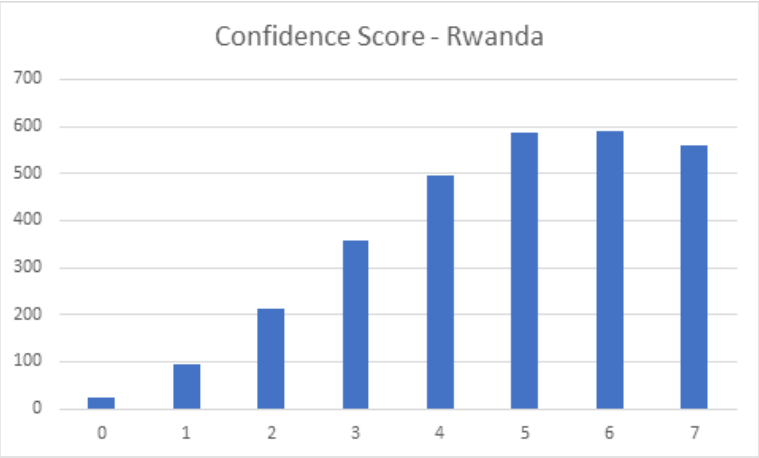
A.6 Rwanda SEL Score Distributions

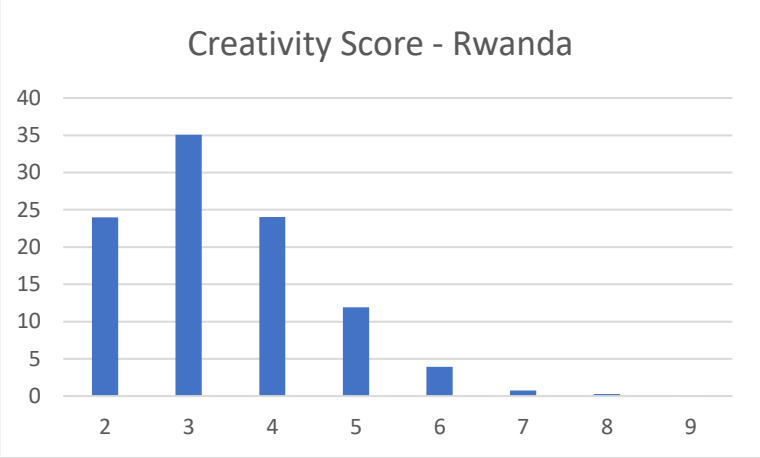









Confidence Score - Rwanda
700
600
500
400
300
200
100
0
0
1
2
3
4
5
6
7


Creativity Score - Rwanda
40
35
30
25
20
15
10
5
0
2
3
4
5
6
7
8
9

ANNEX B – RELIABILITY ANALYSIS

B1. EGRA – Cronbach's Alpha

Language	Alpha – Letters	Alpha – Invented Words	Alpha – Oral Reading	Alpha – Reading Comprehension
Asante-Twi	0.855933	0.961457	0.976513	0.883413
Dagbani	0.973305	1.004362	1.005102	0.99807
Ewe	0.929844	0.973146	1.005721	0.821422
Ga	0.809748	1.009676	N/A (all zero scores)	N/A (all zero scores)
Kiswahili	0.895121	0.84963	0.972172	0.659557
English (Kenya)	0.946639	0.8272	0.963929	0.830681
Kinyarwanda	0.951675	0.841433	0.981883	0.831131

B2. EGMA – Cronbach's Alpha

Country	Alpha – Quantity Discrimination	Alpha – Missing Numbers	Alpha – Basic Addition	Alpha – Basic Subtraction	Alpha – Addition Level 2	Alpha – Subtraction Level 2
Ghana	0.515894	0.254295	0.912108	0.910832	0.804785	0.72585
Kenya	0.241248	0.632099	0.182335	-	0.769446	0.712721
Rwanda	0.404892	0.701069	0.933122	0.939797	0.657228	-

B3. SEL

B3.1 Behavioral Observation- Alpha

0.86 - Ghana
0.92 - Kenya
0.85 - Rwanda

B3.2 Relationship Scale- Alpha

0.74 - Ghana (all items perform similarly)
0.62 - Kenya (all items perform similarly)
0.60 - Rwanda (all items perform similarly)

B3.3 Conflict Perception: Item-rest and Alpha

Ghana

	Item-rest	Alpha
Why do you think the child took the toy away?	0.45	0.62
Why do you think the child wasn't helping to sweep the room?	0.43	0.62
Why do you think the child is building the tower/building a different way from you?	0.39	0.64
Why do you think your classmate hasn't given you a turn yet?	0.37	0.65
Why do you think your classmate stood in front of you and took your place?	0.47	0.60
		0.68

Kenya

	Item-rest	Alpha
Why do you think the child took the toy away?	0.36	0.64
Why do you think the child wasn't helping to sweep the room?	0.39	0.61
Why do you think the child is building the tower/building a different way from you?	0.43	0.60

Why do you think your classmate hasn't given you a turn yet?	0.40	0.61
Why do you think your classmate stood in front of you and took your place?	0.36	0.62
		0.67

Rwanda

	Item-rest	Alpha
Why do you think the child took the toy away?	0.35	0.56
Why do you think the child wasn't helping to sweep the room?	0.39	0.54
Why do you think the child is building the tower/building a different way from you?	0.30	0.58
Why do you think your classmate hasn't given you a turn yet?	0.34	0.57
Why do you think your classmate stood in front of you and took your place?	0.40	0.53
		0.61

B3.4 Confidence- Item-Rest Correlation and Alpha

Ghana

	Item-Rest Correlation	Alpha
Richard/Emma often feels confident (not afraid) when going into a room…	0.38	0.71
When children in the class have a question, they get Richard/Emma to ask…	0.50	0.68
Richard/Emma can lead other children to sing a song…	0.47	0.69
Richard/Emma starts conversations with adults he/she doesn't know well…	0.39	0.71
Richard/Emma likes to read stories out to the whole class…	0.49	0.69
Richard/Emma is trusted to look after younger children…	0.35	0.72
Richard/Emma shares his/her thoughts even if they are very different…	0.50	0.69
		0.73

Kenya

	Item-Rest Correlation	Alpha
Richard/Emma often feels confident (not afraid) when going into a room…	0.22	0.41
When children in the class have a question, they get Richard/Emma to ask…	0.23	0.40
Richard/Emma can lead other children to sing a song…	0.25	0.39
Richard/Emma starts conversations with adults he/she doesn't know well…	0.17	0.44
Richard/Emma likes to read stories out to the whole class…	0.18	0.43
Richard/Emma is trusted to look after younger children…	0.20	0.42
Richard/Emma shares his/her thoughts even if they are very different…	0.21	0.41
	0.22	**0.45**

Rwanda

	Item-Rest Correlation	Alpha
Richard/Emma often feels confident (not afraid) when going into a room…	0.19	0.57
When children in the class have a question, they get Richard/Emma to ask…	0.31	0.53
Richard/Emma can lead other children to sing a song…	0.30	0.54
Richard/Emma starts conversations with adults he/she doesn't know well…	0.30	0.53
Richard/Emma likes to read stories out to the whole class…	0.33	0.52
Richard/Emma is trusted to look after younger children…	0.31	0.53
Richard/Emma shares his/her thoughts even if they are very different…	0.32	0.53
		0.57

B3.5 Learning Environment- Item-Rest Correlation and Alpha

Ghana

	Item-Rest Correlation	Alpha
Scored(+): Does your teacher praise you when you give the right answer?	0	0.63
Scored(+): Can you talk to at least one teacher in the school about things that..?	0	0.62
Scored(+): Last week, did children in your class ask the teacher questions..?	0	0.61
Scored(+): Last week, did some children in you class ask the teachers for help..?	0	0.61
Scored(-): Do your teachers encourage you to try out new things in class?	0	0.61
Scored(-): Do your teachers get angry or yell when you get the answer wrong?	0	0.63
Scored(+): Do your teachers help those who do not understand?	0	0.63
Scored(-): In the last week, did an adult in your school scream or yell..?	0	0.61
Scored(+): Can students share their thoughts in front of the class?	0	0.62
Scored(-): In the last week, did an adult in your school hit a child..?	0	0.65
		0.65

Kenya

	Item-Rest Correlation	Alpha
Scored(+): Does your teacher praise you when you give the right answer?	0.34	0.71
Scored(+): Can you talk to at least one teacher in the school about things that..?	0.28	0.72
Scored(+): Last week, did children in your class ask the teacher questions..?	0.44	0.69
Scored(+): Last week, did some children in you class ask the teachers for help..?	0.46	0.69
Scored(-): Do your teachers encourage you to try out new things in class?	0.47	0.69
Scored(-): Do your teachers get angry or yell when you get the answer wrong?	0.32	0.72
Scored(+): Do your teachers help those who do not understand?	0.38	0.70
Scored(-): In the last week, did an adult in your school scream or yell..?	0.43	0.70
Scored(+): Can students share their thoughts in front of the class?	0.44	0.70
Scored(-): In the last week, did an adult in your school hit a child..?	0.31	0.72
		0.73

Rwanda

	Item-Rest Correlation	Alpha
Scored(+): Does your teacher praise you when you give the right answer?	0.26	0.56
Scored(+): Can you talk to at least one teacher in the school about things that..?	0.30	0.55
Scored(+): Last week, did children in your class ask the teacher questions..?	0.35	0.54
Scored(+): Last week, did some children in you class ask the teachers for help..?	0.36	0.53
Scored(-): Do your teachers encourage you to try out new things in class?	0.39	0.53
Scored(-): Do your teachers get angry or yell when you get the answer wrong?	0.11	0.60
Scored(+): Do your teachers help those who do not understand?	0.31	0.55
Scored(-): In the last week, did an adult in your school scream or yell..?	0.12	0.59
Scored(+): Can students share their thoughts in front of the class?	0.29	0.55
Scored(-): In the last week, did an adult in your school hit a child..?	0.15	0.59
		0.59

B4. Classroom Observation – Factor Analysis

B4.1 Ghana – Factor Analysis

	Factor	Eigen Value	Ratio	Proportion
Integrated Pedagogies	Factor1	1.025	2.4049	1.0804
	Factor2	0.4262	-11.2939	0.4492
	Factor3	-0.0377	0.2319	-0.0398
	Factor4	-0.1628	0.5389	-0.1715
	Factor5	-0.302		-0.3183
Application	Factor1	1.3588	8.6082	1.1991
	Factor2	0.1578	-4.5712	0.1393
	Factor3	-0.0345	0.3707	-0.0305
	Factor4	-0.0932	0.3642	-0.0822
	Factor5	-0.2558		-0.2257
Responsiveness	Factor1	0.7641	14.2571	1.6169
	Factor2	0.0536	-0.5897	0.1134
	Factor3	-0.0909	0.3575	-0.1923
	Factor4	-0.2542		-0.538
Resources	Factor1	0.7047	2.2185	1.2034
	Factor2	0.3177	4.2779	0.5424
	Factor3	0.0743	-1.5508	0.1268
	Factor4	-0.0479	0.2544	-0.0818
	Factor5	-0.1882	0.6846	-0.3214
	Factor6	-0.2749		-0.4694
Efficiency	Factor1	0.2417	-1.5038	2.9848
	Factor2	-0.1607		-1.9848
Student-directed	Factor1	1.38	1.9661	0.9018
	Factor2	0.7019	9.8075	0.4587
	Factor3	0.0716	1.4735	0.0468
	Factor4	0.0486	-0.2818	0.0317
	Factor5	-0.1724	0.7242	-0.1126
	Factor6	-0.238	0.91	-0.1555
	Factor7	-0.2615		-0.1709

B4.2 Ghana – Factor Loading

	Factor 1	Factor 2	Factor 3	Factor 4	Uniqueness
Integrated Pedagogies	0.3643	0.3			0.7773
	0.4601	0.3308			0.6789
	0.5012	0.0691			0.744
	0.5336	-0.2905			0.6309
	0.3804	-0.371			0.7176
Application	0.3132	0.0683			0.8972
	0.7316	-0.1018			0.4544
	0.6809	-0.1714			0.507
	0.2943	0.2383			0.8566
	0.4185	0.238			0.7682
Responsiveness	0.4881	-0.0906			0.7535
	0.43	-0.1282			0.7987
	0.3958	0.1172			0.8296
	0.4292	0.1234			0.8005
Resources	0.2273	0.0148	0.1712		0.9188
	0.1421	0.034	0.1731		0.9487
	0.3893	-0.2941	-0.0369		0.7606
	0.4072	-0.2778	-0.026		0.7564
	0.3755	0.2584	-0.1133		0.7794
	0.4178	0.293	0.0096		0.7395

	Factor 1	Factor 2	Factor 3	Factor 4	Uniqueness
Efficiency	0.3476				0.8791
	0.3476				0.8791
Student-directed	0.2053	0.4023	0.1474	0.0334	0.7732
	0.44	0.4084	0.0477	0.0335	0.6362
	0.588	0.2209	-0.1226	-0.0401	0.5887
	0.4528	-0.0504	-0.1211	0.1212	0.763
	0.4809	-0.4291	0.0218	-0.0167	0.5839
	0.4646	-0.0479	0.0359	-0.1564	0.7561
	0.3825	-0.3681	0.1268	0.0727	0.6968

B4.3 Kenya – Factor Analysis

	Factor	Eigen Value	Ratio	Proportion
Integrated Pedagogies	Factor1	0.9836	2.0146	1.0213
	Factor2	0.4882	-22.5294	0.507
	Factor3	-0.0217	0.1089	-0.0225
	Factor4	-0.1989	0.6904	-0.2066
	Factor5	-0.2882		-0.2992
Application	Factor1	1.3199	6.0984	1.2678
	Factor2	0.2164	-2.2449	0.2079
	Factor3	-0.0964	0.5494	-0.0926
	Factor4	-0.1755	0.7858	-0.1686
	Factor5	-0.2233		-0.2145
Responsiveness	Factor1	0.5509	2.2723	1.0221
	Factor2	0.2424	2.3077	0.4498
	Factor3	0.1051	-0.2923	0.1949
	Factor4	-0.3594		-0.6669
Resources	Factor1	0.9795	1.4733	0.8738
	Factor2	0.6648	34.8191	0.5931
	Factor3	0.0191	-0.5972	0.017
	Factor4	-0.032	0.1289	-0.0285
	Factor5	-0.248	0.9444	-0.2212
	Factor6	-0.2625		-0.2342
Efficiency	Factor1	0.4286	-1.9575	2.0443
	Factor2	-0.2189		-1.0443
Student-directed	Factor1	1.5462	2.9089	0.905
	Factor2	0.5315	1.8122	0.3111
	Factor3	0.2933	223.1625	0.1717
	Factor4	0.0013	-0.0094	0.0008
	Factor5	-0.1398	0.63	-0.0818
	Factor6	-0.2218	0.7341	-0.1298
	Factor7	-0.3022		-0.1769

B4.4 Kenya – Factor Loading

	Factor 1	Factor 2	Factor 3	Factor 4	Uniqueness
Integrated Pedagogies	0.2575	0.3383			0.8192
	0.4203	0.3007			0.7329
	0.5173	0.2271			0.6808
	0.5003	-0.3606			0.6197
	0.472	-0.3191			0.6754
Application	0.2111	0.351			0.8322
	0.5835	0.1016			0.6492
	0.6026	0.072			0.6316
	0.413	-0.255			0.7644

	Factor 1	Factor 2	Factor 3	Factor 4	Uniqueness
	0.6333	-0.1128			0.5862
Responsiveness	0.1864	0.3264	0.1732		0.8287
	0.426	-0.1574	0.1817		0.7607
	0.4962	-0.1621	-0.1227		0.7125
	0.2974	0.2913	-0.1642		0.7997
Resources	0.0487	0.2495	0.0749		0.9298
	0.3094	-0.0874	-0.0989		0.8869
	0.5528	0.2909	-0.0111		0.6096
	0.455	0.4344	0.0063		0.6042
	0.4268	-0.3697	0.0582		0.6778
	0.4321	-0.43	0.0124		0.6283
Efficiency	0.4629				0.7857
	0.4629				0.7857
Student-directed	0.3272	0.246	-0.0637	-0.0232	0.8278
	0.5205	0.2315	-0.2408	-0.0007	0.6175
	0.5409	0.2703	-0.1147	0.017	0.6208
	0.4576	-0.3686	-0.0896	0.0107	0.6466
	0.5402	-0.0014	0.2999	-0.012	0.6181
	0.575	-0.3935	0.0261	-0.0046	0.5139
	0.209	0.2317	0.3456	0.0145	0.783

B4.5 Rwanda – Factor Analysis

	Factor	Eigen Value	Ratio	Proportion
Integrated Pedagogies	Factor1	0.9731	1.8024	0.9332
	Factor2	0.5399	22.0207	0.5177
	Factor3	0.0245	-0.1201	0.0235
	Factor4	-0.2041	0.7022	-0.1957
	Factor5	-0.2906		-0.2787
Application	Factor1	1.8685	13.147	1.1904
	Factor2	0.1421	-1.684	0.0905
	Factor3	-0.0844	0.6614	-0.0538
	Factor4	-0.1276	0.5571	-0.0813
	Factor5	-0.229		-0.1459
Responsiveness	Factor1	0.3322	4.2265	2.1528
	Factor2	0.0786	-1.7528	0.5094
	Factor3	-0.0448	0.2119	-0.2906
	Factor4	-0.2116		-1.3715
Resources	Factor1	0.6736	2.4346	1.2409
	Factor2	0.2767	9.8675	0.5097
	Factor3	0.028	-1.2474	0.0517
	Factor4	-0.0225	0.2367	-0.0414
	Factor5	-0.095	0.2986	-0.1749
	Factor6	-0.318		-0.5859
Efficiency	Factor1	0.4845	-2.1107	1.9003
	Factor2	-0.2296		-0.9003
Student-directed	Factor1	2.5966	5.19	0.8742
	Factor2	0.5003	1.6055	0.1684
	Factor3	0.3116	15.7105	0.1049
	Factor4	0.0198	-0.2052	0.0067
	Factor5	-0.0967	0.6297	-0.0325
	Factor6	-0.1535	0.7386	-0.0517
	Factor7	-0.2078		-0.07

B4.6 Rwanda – Factor Loading

	Factor 1	Factor 2	Factor 3	Factor 4	Uniqueness
Integrated Pedagogies	0.1079	-0.0182	0.1507		0.9653
	0.4111	0.3947	-0.028		0.6744
	0.5011	0.3619	0.0202		0.6176
	0.595	-0.2619	-0.025		0.5768
	0.4329	-0.4292	0.0001		0.6284
Application	0.4922	0.2688			0.6855
	0.6597	0.0328			0.5638
	0.7108	0.0656			0.4905
	0.4884	-0.2043			0.7197
	0.6688	-0.1507			0.53
Responsiveness	0.357	0.047			0.8703
	0.0868	0.2435			0.9332
	0.357	-0.1275			0.8563
	0.2641	0.0288			0.9294
Resources	0.0482	0.3356	0.0837		0.878
	0.3638	-0.0308	-0.1117		0.8542
	0.3869	0.2403	-0.0142		0.7923
	0.0966	0.2457	-0.0484		0.928
	0.4301	-0.1932	0.0643		0.7736
	0.4415	-0.0875	0.0433		0.7956
Efficiency	0.4922				0.7577
	0.4922				0.7577
Student-directed	0.1732	0.3212	0.2186	-0.0005	0.819
	0.2228	0.3302	0.1451	0.0439	0.8183
	0.4006	0.0479	0.3611	-0.0034	0.7068
	0.847	-0.2738	0.1126	-0.0537	0.1921
	0.7085	0.2287	-0.1791	-0.0828	0.4068
	0.6811	0.2541	-0.2599	0.0509	0.4014
	0.8205	-0.3066	-0.0077	0.0745	0.2272

ANNEX C – INSTRUMENTS

C.1 Attitudes Scale – Teachers

In the following set of questions, I will read you two statements. I want you to tell me which of the two statements you agree with most, or which statement you prefer.

For example, let me read two statements to you: Statement 1: I like tea. Statement 2: I like coffee. Which do you agree with? If they agree with both, say "Which do you agree with more." So, you see, perhaps you like both tea and coffee, but we want to know which statement you agree with most.

When answering the following questions, I want you to think of children in Grades 1 through Grade 3

Statement 1	***Statement 2***
Children learn best from the teacher.	***Children learn best from their peers.***
1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
It is important for the class to be orderly.	***It is important for children to be free to talk.***
1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
Children learn literacy/language best in small groups.	***Children learn literacy/language best as a whole class.***
1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
It is better if children tell the teacher when they don't understand.	***If children don't understand, they should try to find the answer independently before asking the teacher.***

1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
If children answer their own questions, they will be more likely to understand.	***If the teacher answers children's questions they will be more likely to get the correct answer.***
1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
Group work undermines the class's sense of togetherness.	***The class does not have to sit together to have a sense of togetherness.***
1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
Parents welcome more play-based approaches to learning in the classroom	***Parents prefer that I stick to traditional teaching approaches in the classroom***
1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
Children learn best by exploring new things on their own.	***Children learn best when new things are explained by the teacher.***
1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
Play is an important part of learning for children in primary school	***Play is only important for learning for preschool aged children***
1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
Children can choose their own learning activities in the classroom.	***Teachers are responsible for choosing the learning activities for his/her students.***

1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
It is more important that the teacher is engaged	***It is more important that the students are engaged***
1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	
A good teacher oversees the whole class at all times	***A good teacher is able to work with one group of students while other students work independently.***
1. Agree with 1 much more than 2 *2. Agree with 1 and 2 equally.* *3. Agree with 2 much more than 1*	

C.2 Attitudes Scale - Caregivers

In the following set of questions, I will read you two statements. I want you to tell me which of the two statements you agree with most, or which statement you prefer.

For example, let me read two statements to you: Statement 1: I like tea. Statement 2: I like coffee. Which do you agree with? If they agree with both, say "Which do you agree with more." So, you see, perhaps you like both tea and coffee, but we want to know which statement you agree with most.

When answering the following questions, I want you to think of children in Grades 1 through 3.

Statement 1	*Statement 2*
It is more important to teach creativity than academic subjects	***It is more important to teach academic subjects than creativity***
Teachers should create a friendly classroom environment	***Teachers should create a well-disciplined classroom environment***
Teachers should use play to promote learning in the classroom.	***Play should only happen outside of the classroom.***
Children learn best when they are having fun	***Children learn best when they are serious***
I would be happy if teachers focus on building my child's confidence and curiosity in the classroom	***I'd prefer teachers focus on academic skills in the classroom***
It is important for the class to be orderly.	***It is important for children to be free to talk.***

C.3 Timed Classroom Observation

Directions: Using a timing device, at 2-minute intervals capture a snapshot of what is happening at each level. The observer codes what is happening at the exact 2-minute interval, not what happened previously. Only one item per level, per interval, can be selected.

Lesson Minute	2	4	6	8	10	12	14	16	18	20	22	24	26
Level: Facilitator (teacher)													
Explain concept (X)													
Models (L)													
Affirms (F)													
Question Closed-Ended (QC)													
Question Open-Ended (QO)													
Monitors (R)													
Suggest (S)													
Analyze (Z)													
Management (M)													
Nothing (N)													
Level: Student													
Reading (R)													
Listening (L)													
Writing (W)													
Orally (O)													
Turn-taking (T)													
Physical (P)													
Manipulate (M)													
Getting (G)													
Not Engaged (N)													
Level: Source (State of Play)													
Following directions (F)													
Planning (P)													
Self-directed (S)													
Level: Materials													
Book (B)													
Chalkboard (CB)													
Written (W)													
Oral (O)													
Illustration (I)													
Manipulatives (M)													
Physical Space (P)													
Crafts (C)													
Repurposed (R)													
Stations (S)													
Toy (T)													
Nothing (N)													
Level: Focus (pick 2: organization and the nature of the social interaction)													
Whole Class (W)													
Large Group (L)													
Small Group (S)													
Individual (I)													
Corners (C)													
Competition (T)													
Cooperation (P)													
Conflict (F)													
Level: Engaged													
Most (M)													
Half (H)													
Few (F)													
Level: Climate (pick 2: tone and initiator)													
Perseverance (R)													
Positive Verbal (PV)													
Positive Physical (PP)													
Negative Verbal (NV)													
Negative Physical (NP)													
Ordinary (O)													
Teacher initiated (T)													
Child initiated (C)													
Not Applicable (N)													

Lesson Minute	28	30	32	34	36	38	40	42	44	46	48	50
Level: Facilitator (teacher) (Teacher)												
Explain concept (X)												
Models (L)												
Affirms (F)												
Question Closed-Ended (QC)												
Question Open-Ended (QO)												
Monitors (R)												
Suggest (S)												
Analyze (Z)												
Management (M)												
Nothing (N)												
Level: Student												
Reading (R)												
Listening (L)												
Writing (W)												
Orally (O)												
Turn-taking (T)												
Physical (P)												
Manipulate (M)												
Getting (G)												
Not Engaged (N)												
Level: Source (State of Play)												
Following directions (F)												
Planning (P)												
Self-directed (S)												
Level: Materials												
Book (B)												
Chalkboard (CB)												
Written (W)												
Oral (O)												
Illustration (I)												
Manipulatives (M)												
Physical Space (P)												
Crafts (C)												
Repurposed (R)												
Stations (S)												
Toy (T)												
Nothing (N)												
Level: Focus (pick 2: organization and the nature of the social interaction)												
Whole Class (W)												
Large Group (L)												
Small Group (S)												
Individual (I)												
Corners (C)												
Competition (T)												
Cooperation (P)												
Conflict (F)												
Level: Engaged												
Most (M)												
Half (H)												
Few (F)												
Level: Climate (pick 2: tone and initiator)												
Perseverance (R)												
Positive Verbal (PV)												
Positive Physical (PP)												
Negative Verbal (NV)												
Negative Physical (NP)												
Ordinary (O)												
Teacher initiated (T)												
Child initiated (C)												
Not Applicable (N)												

C.4 Post-Observation Checklist

Directions: This checklist is completed immediately after observing the lesson. Tick yes if you observed the behavior during the observation; tick no if you did not observe it. Some advanced observers may tick items as they are seen during the observation.

	Integrated Pedagogies		
1.	The teacher asked questions or raised topics that got students to respond to each other, not to her (i.e., active).	yes	no
2.	Student(s) persisted with the activity after the rest of the class has started to do something else (i.e., engage).	yes	no
3.	Student(s) started the task again or adjusted when they thought of a better idea/approach (i.e., iterative).	yes	no
4.	Student(s) had a chance to make connections between the new experiences/content to their prior knowledge or experiences (i.e., meaningful).	yes	no
5.	Students problem solved or had a chance to think critically or reflect about a topic.	yes	no
	Application: Students have generalized games.		
6.	Songs or energizers to start or divide activities.	yes	no
7.	Students played games with obvious rules.	yes	no
8.	Students had physical play.	yes	no
9.	Students used objects to represent something else.	yes	no
10.	Students had symbolic play.	yes	no
	Responsiveness: The teacher adapts to student behaviours.		
11.	The teacher gave students hints, suggestions, or redirection when they needed it.	yes	no
12.	The teacher circulated, leaned in, redirected, to support students as needed.	yes	no
13.	The teacher asked the class or a group for a solution to a student question or problem.	yes	no
14.	The teacher included students who did not volunteer to answer (i.e., Inclusion).	yes	no
	Resources: Materials that are needed for implementation are used.		
15.	Students used a resource that was dedicated for their use (e.g., book, workbook, slate, tablet).	yes	no
16.	Students used something to create something else (e.g., art supplies, items in nature, objects repurposed).	yes	no
17.	Students used a resource to solve a problem (e.g., manipulatives, math charts, dictionary).	yes	no
18.	Instruction used an item designed for interaction between teacher and students or between students (e.g., read aloud, pocket chart, counters, puzzle).	yes	no
19.	Students used physical space for learning activities (e.g., counting large steps for measurement, sit in a circle for sharing).	yes	no
20.	Students used stations/corner work/learning centers.	yes	no
	Efficiency: Instructional time is maximized.		
21.	The teacher had materials ready at the start of an activity or the class.	yes	no
22.	The pace of the instruction was lively.	yes	no
	Student-Directed (Ownership / Autonomy)		
23.	Students developed and shared ideas with either classmates or the teacher.	yes	no
24.	Students had a chance to select from choices. (state of play - exploring)	yes	no
25.	Students chose collaborators or roles.	yes	no
26.	Students made or changed rules for an activity.	yes	no
27.	Students chose how long to do an activity. (state of play - owning).	yes	no
28.	Students set goals or helped to plan how to use time.	yes	no
29.	Students chose when to move around.	yes	no

Material Inventory

Directions: Complete immediately after completing the binary checklist. Tick yes if teacher and student materials are available in the classroom; tick no if materials are not available in the classroom.

	Teacher Materials Teacher has [program] teacher materials IN THE CLASSROOM			
1a	Teachers' guide for language (reading/writing/listening/speaking)	yes	no	n/a
1b	Teachers' guide for mathematics	yes	no	n/a
1c	Teaching aids (e.g., pocket chart, letter cards, counters)	yes	no	n/a
	Student Materials Student [program] materials are available IN THE CLASSROOM			
2a	Textbooks for language (reading/writing/listening/speaking)	yes	no	n/a
2b	Textbooks for mathematics	yes	no	n/a
2c	Student workbooks for language (reading/writing/listening/speaking)	yes	no	n/a
2d	Student workbooks for mathematics	yes	no	n/a
2e	Supplementary readers	yes	no	n/a
2f	Student workbooks for mathematics	yes	no	n/a
	Student Textbook Ratio Student textbook ratio IN THE CLASSROOM			
3a	How many students are in the classroom? [Count students in the class]	|____|____|____|		
3b	How many students have textbooks for language in the classroom? [Count]	|____|____|____|		
3c	How many students have textbooks for mathematics in the classroom? [Count]	|____|____|____|		
	Library Materials Library materials in the school and classroom			
4a	Is there a school library?	yes	no	n/a
4b	Is there a classroom library?	yes	no	n/a
4c	How many students in the class have any books from the school/classroom library in their desk/bag? [Ask students to show the books]	|____|____|____|		

Observer Statement

Directions: Complete after the entering the timed tool codes, the binary checklist responses and the materials inventory. Respond to the five prompts:

Prompts
1. What was the learning content?
2. What did the teacher do?
3. How were resources or physical space used by the teacher or the students?
4. How would you describe the children's activity? (instructional or behaviour)
5. What is your opinion of this observation? Include tone.

ANNEX D – OVERVIEW OF PROGRAMS

Aga Khan Foundation in Kenya

The partnership will strengthen the capacity of 3,000 primary school teachers in implementing playful pedagogies in their everyday teaching and learning, thereby benefiting 120,000 children. It will reach out to 6,000 parents and community members to bolster acceptance for the use of learning through play methodologies in implementing the new curriculum. Four hundred school-based pedagogical leaders and resource teachers will be trained and engaged in supporting teachers' use of playful pedagogies at the school and classroom level.

Right to Play in Ghana

In partnership with five district governments and the Ministry of Education, the project will extend the use of Learning through Play teacher training package to **312 primary schools**. This will expand Learning through Play pedagogy coverage of all **451 government primary schools** in the five districts. It will build the capacity of **1,870 teachers and benefit 65,520 children in 312 primary schools.** It will also **reach approximately 14,400 parents and community members**, building their support for Learning through Play both in school and at home. Additionally, the project will **expand to 10,000 GALOP schools** through indirect support provided through master training for 6,000 schools and direct support at the district level for 4,000 schools.

UNICEF in Rwanda

This partnership will embed Learning through Play pedagogies into Rwanda's in-service teacher training system nationwide, refining school-based mentoring materials to better integrate Learning through Play pedagogies, and use these materials to **build the capacity of over 800 Mentor Trainers** and the **2,900 SBMs nationwide,** reaching approximately **29,000 teachers** and supporting integration of Learning through Play in their classrooms, resulting in approximately **1,450,000 primary school children** benefiting from the transformative power of learning through play.