

**COMMUNITY-BASED LAY-WORKER
FEASIBILITY CLUSTER RANDOMIZED
CONTROLLED TRIAL**

**STRONG FAMILIES, THRIVING
CHILDREN**

SUGIRA MURYANGO

Endline Data Collection Report

November 2019

Table of Contents

INTRODUCTION	4
Project Overview and Background	4
Project Objectives	4
ROLES AND RESPONSIBILITIES	5
Boston College	5
Laterite	5
METHODOLOGY	7
Design and Sampling	7
Figure 1. Number of Clusters Allocated to Treatment and Control	7
Figure 2. Number of Households Allocated to Treatment and Control	7
Research Instruments	7
Surveys at the Household	8
MDAT and Anthropometric Surveys	8
FIELD TEAM	9
Recruitment and Profile of Team Members	9
Field Supervisors	9
Field Coordinators	9
Enumerators	9
Field Team Structure	9
Surveys at the Household Field Team	9
Figure 3. Non-MDAT Field Team Structure	10
MDAT and Anthropometrics Field Team	10
Figure 4. MDAT and Anthropometrics Field Team Structure	10
Project Management	10
FIELD PLAN AND PROTOCOLS	12
Informed Consent	12
Strategy for Finding Households	12
Interview Protocols at Endline	13
Exclusion Criteria	13
Caregivers	13
Surveys Completed by Caregivers	14
Adverse Event and Risk of Harm Reporting	15
DATA COLLECTION TRAINING	16
Surveys at the Household	16
MDAT and Anthropometrics Surveys	16
SURVEY INSTRUMENT PILOTING	17
Surveys at the Household	17
Objectives	17
Implementation	17
Measure of Inter Rater Reliability (IRR)	17
Figure 5. Inter-rater reliability estimates for simultaneous observations of the Report on Child	17

MDAT	18
Objectives	18
Implementation	18
Measures of Rater Reliability (IRR)	18
DATA COLLECTION	19
Schedule	19
Completion Statistics	19
Figure 7. Number of Clusters by Treatment Status in Endline Dataset	19
Figure 8. Number of Households by Treatment Status in Endline Dataset	19
Figure 9. Surveys completed for six survey instruments by baseline district	20
Figure 10. Caregiver surveys completed at endline by baseline district	20
Attrition	20
Figure 11. Attrition in each wave of data collection	21
Figure 12. Respondent Attrition	21
Administration of Surveys	21
Surveys at the Household	21
MDAT and Anthropometrics Surveys	22
Survey Duration	22
Figure 13. Median duration for each survey	22
Participant Compensation	23
Challenges	23
DATA QUALITY MONITORING	24
Field Supervision	24
Real-time Completion Tracking	24
Routine Monitoring	24
Flagged Survey Responses	25
Figure 14. Flagged Responses Triggering the Risk of Harm Protocol	25
Figure 15. Additional Flagged Responses	26
Audio Audits	26
Inter Rater Reliability	27
Design	27
Results	28
Figure 16. Inter-rater reliability for components of the Report on Child	28
Figure 17. Inter-rater reliability for components of the Report on Child	28
DATA CLEANING	29
Duplicates Management	29
Master Datasets	29
Caregiver IDs	30
Figure 18. Identifying variables in each dataset	31

Appendices

Appendix 1. Risk of Harm Action Flowchart	32
Appendix 2. Entire household missing.....	33
Appendix 3. Partially Complete Households: Child, Caregiver, or Household Surveys missing	33
Appendix 4. Questions in the Audio Audit.....	36

INTRODUCTION

Project Overview and Background

This study is a cluster randomized controlled trial of the *Sugira Muryango* intervention delivered by community-based lay-workers trained and supervised by FXB Rwanda. The study was designed as a quantitative survey conducted in three districts of Rwanda (Nyanza, Ngoma, Rubavu) with 1,040 households and qualitative interviews with a subset of 40 caregivers. *Sugira Muryango* will be evaluated as a supplementary intervention for households participating in Vision Umurenge 2020 Programme (VUP) Public Works. Data is being collected at three time points:

- Baseline: immediately prior to the intervention (April-May 2018);
- Midline: immediately following the intervention (August-September 2018).
- Endline: one year following the end of the intervention (August-September 2019).

The project will evaluate the impact of the *Sugira Muryango* intervention for all households participating in the study as well as for VUP Classic Public Works and VUP Expanded Public Works separately.

Project Objectives

The main objective of this study is to determine whether the *Sugira Muryango* intervention, delivered by community-based lay-workers, is effective in improving early childhood development (ECD) outcomes and supporting vulnerable families. This will be measured through surveys with the primary caregiver and their intimate partner (when applicable), and through child development assessments and anthropometric measurements.

This study will be used to support the design of policy aimed at strengthening Rwanda's most vulnerable families.

The data collected will also be used by the Government of Rwanda and the World Bank to test the sustainability and scalability of *Sugira Muryango* when delivered by community-based lay-workers and linked to VUP Public Works programming.

ROLES AND RESPONSIBILITIES

Throughout the design process and the implementation of baseline, midline, and endline data collection, Laterite worked very closely with the team from Boston College to operationalize the study design.

Boston College

Boston College's responsibilities included but were not limited to:

Obtaining research permits: Boston College was responsible for obtaining approvals to conduct the study from Rwanda National Ethics Committee (RNEC) and the National Institute of Statistics of Rwanda (NISR).

Sampling: Boston College led the design of this cluster randomized control study, including developing the sampling strategy.

Designing research instruments: Boston College designed the research instruments and consent forms and completed initial translations. Prior to data collection, Boston College approved the final versions of the instruments.

Training: Boston College designed and led field team training on the Malawi Development Assessment Tool (MDAT).

Data collection: When Laterite identified potential cases of severe abuse, suicidality, or malnutrition, Boston College and FXB were responsible for assessing risk severity and making referrals for additional services or treatment.

Laterite

Laterite's responsibilities included:

Obtaining approvals from local authorities: After receiving approvals from the RNEC and NISR, Laterite was responsible for informing and obtaining clearance for conducting the study from local authorities at the district, sector, cell, and village levels.

Sampling: Laterite operationalized the sampling strategy by collecting updated VUP participant lists, conducting village listing and household rostering to identify households that were eligible to participate in the study, randomly selecting clusters of households based on Boston College's study design, and enrolling the selected households in the study.

Reviewing, coding, and testing research instruments: Laterite reviewed all survey instruments and made recommendations regarding translations, cultural appropriateness, and logic patterns. Once the research instruments were finalized, Laterite programmed them into SurveyCTO for electronic data collection. Enumerators and Laterite researchers tested multiple iterations of the surveys to ensure functionality and estimate the time required to administer each instrument.

Recruitment and training of field staff: Laterite recruited a team of data collectors from its roster of qualified personnel. Laterite developed and delivered training on all surveys completed at the

household and on taking and recording anthropometric measurements. Laterite also included in its training sessions an overview of the project, Laterite policies, and research ethics.

Data collection: Laterite led all data collection activities, including developing field plans, scheduling appointments, conducting interviews and assessments, managing enumerator teams, and addressing quality concerns while in the field.

Data cleaning, monitoring, and auditing: Throughout data collection, Laterite monitored the quality of data, audited audio recordings of a random subset of interviews, and worked with the team in the field to resolve duplicates or discrepancies. With this report, Laterite submitted a cleaned dataset to Boston College.

METHODOLOGY

Design and Sampling

Details regarding the methodology of this cluster randomized controlled trial are included in the Field Preparation Report dated June 2018.

Figure 1 and **Figure 2** show the target number of clusters and households to be enrolled in the trial. Because of the timing of cluster allocation to treatment assignment and cluster failure, there are an uneven number of Expanded Public Works clusters and households in the treatment and control arms of the trial and in the pooled analysis. Additional details on the process of cluster selection, treatment allocation, and the selection of additional clusters are described in detail in the Field Preparation Report.

Figure 1. Number of Clusters Allocated to Treatment and Control

	Control	Treatment	Total
Classic PW	75 clusters	75 clusters	150 clusters
Expanded PW	39 clusters	47 clusters	86 clusters
Pooled	98 clusters	100 clusters	198 clusters

Figure 2. Number of Households Allocated to Treatment and Control

	Control	Treatment	Total
Classic PW	375 households	375 households	750 households
Expanded PW	138 households	174 households	312 households
Pooled	513 households	549 households	1062 households

Research Instruments

Quantitative data collection utilized the survey instruments described below, which are broadly categorized according to the location where the surveys were administered. Boston College designed and translated the research instruments, which Laterite reviewed for appropriateness and effectiveness. Laterite coded all surveys for deployment using SurveyCTO.

Changes to the survey instrument between baseline and midline are detailed in the Midline Field Report, dated October 2018.

Between midline and endline data collection, survey edits were made based on feedback from the client, enumerators, and research team. Changes to the surveys were reviewed and discussed with the Boston College Team. Prior to endline data collection, Laterite shared the coded final research instruments in Excel format and enabled Boston College with electronic access to the surveys on SurveyCTO. The research team from Boston College reviewed and approved all final instruments at the start of data collection. Changes made between midline and endline data collection, and those made during endline data collection are compiled in the “FSI Survey Changes Endline” Excel file in the 13 - Endline Surveys folder on Box.

Surveys at the Household

The following surveys were administered at the respondent's domicile.

- **Report on the Child:** This survey was completed by the baseline primary caregiver, midline primary caregiver or the new endline caregiver for and with each eligible child in the household. It includes: questions for the caregiver regarding child feeding practices, food security, child health, and child discipline; the Ages and Stages Questionnaire (ASQ)-3; the Home Observation for Measurement of the Environment (HOME); and the Observation of Mother Child Interaction (OMCI).
- **Caregiver Report on Household:** This survey was completed by the primary caregiver in each study household or by his or her intimate partner or additional caregiver identified at midline and includes modules regarding the family composition, household assets, social protection, VUP participation, finances, and food security.
- **Caregiver Report on Self:** This survey was completed by each primary caregiver in the study household. If the primary caregiver had an intimate partner, the Caregiver Report on Self was completed by the intimate partner as well. This survey was also completed by one additional caregiver identified at each household (if they exist); in most cases a grandmother of the eligible child that had caregiving responsibilities of the child. Intimate partner and additional caregiver surveys were not considered mandatory for household completion, but every attempt was made to complete this survey with all caregivers when applicable. The survey includes modules regarding health, disability, parenting and co-parenting, the family unit, early childhood development knowledge, caregiver mental health, caregiver alcohol use, daily hardships, and intimate partner violence.

MDAT and Anthropometric Surveys

The following instruments were administered at a central location:

- **Malawi Development Assessment Tool (MDAT):** This observation-based assessment was completed by each eligible child in the household.
- **Child anthropometric measurements:** Measurements of weight, height or length depending on age, and mid-upper arm circumference (MUAC) were taken for each eligible child in the household.

FIELD TEAM

Recruitment and Profile of Team Members

Laterite maintains a pool of qualified personnel to act as field team members for quantitative and qualitative data collection projects. In order to qualify for field enumeration work, individuals must pass a rigorous testing process which evaluates logic, problem solving, and, most importantly, communication skills. The interview process includes an Excel test, a written test, and a situational face-to-face interview. Successful candidates are added to the Laterite roster. Laterite selected data collectors for this study from this roster of qualified candidates. The majority of the enumerators utilized during endline data collection were experienced members of the field team from baseline data collection. Below is a description of the profiles of the field staff who worked on this project:

Field Supervisors

Field supervisors are experienced members of the Laterite team who oversee data collection activities and manage both field coordinators and enumerators.

Field Coordinators

Field coordinators are ground-level supervisors who lead small teams of enumerators while conducting data collection themselves. They coordinate logistics for their teams and carefully monitor the team's daily data uploading.

Enumerators

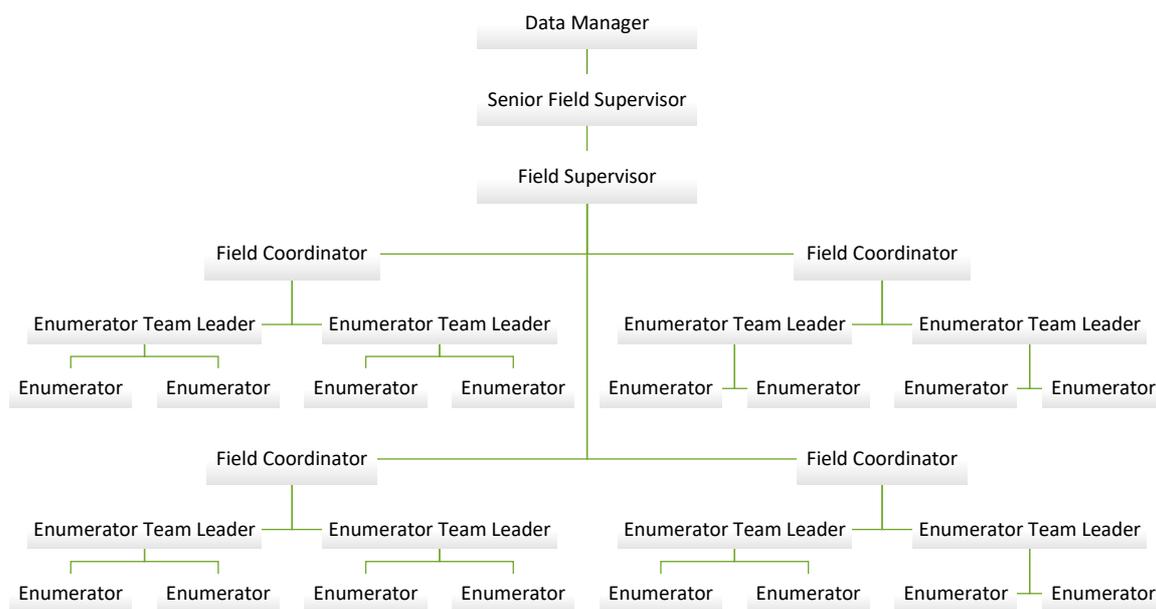
Enumerators are data collection specialists who conduct in-home and centralized interviews and assessments with adults and children in study households. Among groups of three enumerators, one was designated the field team leader to coordinate group logistics and liaise with the field coordinator.

Field Team Structure

Surveys at the Household Field Team

The household survey field team comprised 24 enumerators divided into eight sub-teams of three enumerators each. Each sub-team was led by an enumerator field team leader and field coordinators oversaw two sub-teams of three enumerators each. Laterite's full-time senior field supervisor in coordination with the field supervisor managed the team completing the surveys at the household, coordinated all data collection activities, and reported directly to the Data Manager. The structure of the household field team is outlined in **Figure 3**. Five of the 24 enumerators (20%) at Endline were new; the remaining 19 were also part of the baseline household enumerator team.

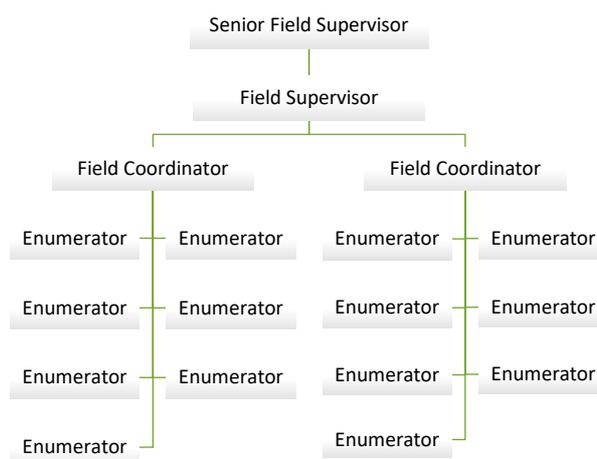
Figure 3. Non-MDAT Field Team Structure



MDAT and Anthropometrics Field Team

In each district, the MDAT and anthropometrics field team consisted of 16 enumerators divided into two field teams of seven enumerators and led by one field coordinator. The enumerators and field coordinators shared responsibilities of conducting MDAT assessments, taking anthropometric measurements, and conducting data entry. One of Laterite’s full-time senior field supervisors in coordination with the field supervisor managed all team activities. The structure of the MDAT field team is outlined in **Figure 4**. Three of the 16 enumerators at endline was new, the remaining 13 enumerators were also part of the baseline and midline MDAT team.

Figure 4. MDAT and Anthropometrics Field Team Structure



Project Management

Amani Ntakirutimana, Data Manager at Laterite Rwanda, managed all data collection activities and field staff for this project and was embedded with the field team for a significant portion of endline

data collection. Amani is one of the most experienced field researchers in Rwanda and specializes in large-scale data collection projects involving complex logistics.

Eric KAGABO led the non-MDAT team as the Senior Field Supervisor and was with the team in the field for the duration of the project. Eric was responsible for supervising the team in the field, leading daily debrief sessions with the team and report any identified issues to the data manager who will provide responses after consulting the research team at Laterite or Boston college. Juliet Kalimba led the MDAT and Anthropometrics team as the Senior Field Supervisor and was with the team in the field for the duration of the project. Juliet was in charge of overseeing MDAT activities in the field, answering questions from enumerators and draft field daily report.

FIELD PLAN AND PROTOCOLS

Informed Consent

Laterite enumerators sought informed consent from participants at least once during field preparation, baseline data collection, midline data collection, and endline data collection. Caregivers consented both themselves and their children.

The consent form, provided by Boston College, included the following information:

- Subject and purpose of the research;
- How participants were selected;
- That personal information collected would be used only for study purposes and would remain confidential;
- The type of questions to be asked and the estimated time to complete the surveys;
- Study contact details in case of questions or concerns; and
- That participation is voluntary and that participants may withdraw at any time.

Enumerators read aloud the consent form in Kinyarwanda given the low literacy level in the catchment area. The forms were signed or, when this was not possible, a thumbprint was given in lieu of a signature. Laterite and the participant each kept a signed copy of the consent form. Laterite stores consent forms at our main office in Kigali.

Strategy for Finding Households

Laterite attempted to contact all participants interviewed at baseline or midline by phone three weeks in advance of endline data collection. The aim of calls was to identify the current location of the household and confirm the availability of eligible children and their caregivers. In addition, Laterite collected information about any new primary caregivers or partners in the household.

Some households had no contact details from the previous rounds of data collection or were unable to be reached by phone. In these cases the Laterite team contacted the village leaders or CHWs to help reach the targeted household head.

Of the 1,049 households interviewed at baseline, 252 households were not able to be reached by phone. To contact these households, Laterite sent a team of 22 enumerators to contact these households and update the household information in person.

Before the deployment of the team for endline, all identified households were contacted again in the week before data collection and informed about the date, time, and venues of the household and MDAT/Anthropometrics interviews. These appointment calls were made by a team of enumerators was calling from the Laterite Kigali office.

Prior to endline data collection in each District, this information from appointment calls and household visits was compiled and shared with the field team. If a participant or household was not available on the day scheduled for their interview, they were added to a list of households for mop-up. Enumerators then followed up with these households and completed any outstanding surveys at the end of data collection in that District.

Interview Protocols at Endline

Exclusion Criteria

Participants were dropped from the study at endline under the following limited circumstances:

- If the child moved outside of the study districts (Rubavu, Ngoma or Nyanza), the household was considered no longer eligible and was not surveyed at endline.
- If a caregiver moved outside of the study districts (Rubavu, Ngoma, or Nyanza), they were not surveyed at endline.
- If a child had died, the household was considered no longer eligible and was not surveyed at endline.

Caregivers

Throughout the study, caregiver types are defined as follows:

- **Primary caregiver:** If available and participating in child-rearing, the biological mother of the eligible child. If the biological mother is not available, then the primary caregiver of the eligible child.
- **Partner of the primary caregiver (occasionally referred to as “secondary” caregiver):** The intimate partner of the primary caregiver.
- **Additional Caregiver:** No more than one additional person in the household who has a significant role in taking care of the child. A significant role is defined as someone who cares for the child the same amount of time as the primary caregiver, or more, or at least 3 days per week. In order of priority, assuming caregiving responsibilities are similar:
 - A grandparent of the eligible child.
 - An aunt or other relative of the child who takes care of the child.

Considerable effort was made to locate and survey all caregivers interviewed at baseline and midline. If a child had a new primary caregiver at endline or if their primary caregiver had a new partner at endline, they were added as new caregivers and interviewed during the endline stage. No new additional caregivers (G1, as defined below) were added during endline data collection. All new caregivers at endline were given a new data collection caregiver id. Below is a description of the seven types of caregivers, as defined at endline.

Baseline Caregivers

1. **Baseline Primary Caregiver (P1 or pcg):** A primary caregiver at baseline, who was NOT the primary caregiver at midline.
2. **Baseline Partner of the Primary Caregiver (P2 or scg):** An intimate partner of the primary caregiver from baseline, who was NOT the intimate partner at midline. *At endline, there are no caregivers of this type in the dataset.*

Additional Caregivers at Midline

3. Midline Primary Caregiver (P3 or acg1): The primary caregiver interviewed at midline – this may be the primary caregiver from baseline, or a new primary caregiver from midline. New midline primary caregivers were added using the following criteria:
 - a. A new primary caregiver who replaced P1 because P1 was no longer in the home at midline.
 - b. A biological mother who was not available at baseline but returned to the home.
 - c. A biological mother who was under 18 at baseline but was over 18 at midline.
4. Midline Partner of the Primary Caregiver (P4 or acg2): The partner of the primary caregiver interviewed at midline – this may be the partner from baseline, or a new partner from midline. New midline partners were added using the following criteria:
 - a. A new intimate partner of P1.
 - b. An intimate partner of P1 who was under 18 at baseline but is now 18.
 - c. An intimate partner of P3.
5. Additional Caregivers (G1 or acg3): The additional caregiver from midline. No new additional caregivers were added at endline.

Additional Caregivers at Endline

6. Endline Primary Caregiver (P6 or acg4): A new endline primary caregiver, defined as one of the following:
 - a. A new primary caregiver who replaced P3 because P3 is no longer in the home at endline.
 - b. A biological mother who was not available at baseline or midline but has returned to the home at endline.
 - c. A biological mother who was under 18 at midline but is now 18.
7. Endline Partner of the Primary Caregiver (P7 or acg5): A new endline partner, defined as one of the following:
 - a. A new intimate partner of P1 or P3.
 - b. An intimate partner of P1 or P3 who was under 18 at midline but is now 18.
 - c. An intimate partner of P6.

Surveys Completed by Caregivers

Each survey was completed by one or more respondent in the household. The following criteria were used to determine who should complete each type of survey:

- Caregiver Report on Self
 - Completed by: each caregiver identified in the household
 - Number completed: minimum 1; maximum 7 (theoretically, in practice the maximum number of caregivers interviewed per households is 5)
- Caregiver Report on Household¹

¹ As at midline, there are cases of two CGRH surveys per “household” at endline. We have created **house_id** which distinguishes between houses within the same household. We use “household” to refer to the family unit as it was surveyed at baseline, which now may be residing in one or more “house” and geographic location. See Data Cleaning section at end for more details.

- Completed by: Any caregiver surveyed in the household
 - The VUP recipient was prioritized for survey completion. If unavailable, the most consistent caregiver was prioritized.
 - If a child moved to a new household between waves of data collection, the CGRH was attempted for each household the child lived in.
- Number completed: minimum 1; maximum 3 (theoretically, in practice the maximum is 2 “houses” per household)
- Report on Child
 - Completed by: the baseline, midline or endline primary caregiver
 - If the baseline primary caregiver was available and still had caring responsibilities for the child, they were prioritized for survey completion at endline to ensure consistency in respondents over time.
 - Number completed: one per eligible child in the household
- MDAT & Anthropometrics
 - Completed by: any member of the household could bring the child to the MDAT & Anthropometrics assessments
 - Number completed: one per eligible child in the household

Adverse Event and Risk of Harm Reporting

An adverse event is any occurrence observed during research which suggests that a participant may be at risk for, or may have experienced physical, mental or sexual harm. This information may be directly reported by the participant or may be observed, e.g., bruising as a sign of physical abuse. Adverse events identified by the field team were managed according to a detailed risk of harm protocol, outlined in Appendix 1, developed by Boston College, which guided the field team through the appropriate response to different risks of harm.

Specific survey questions automatically triggered action according the risk of harm protocol. In these cases, enumerators were required to enter additional information at the end of the survey regarding participant risk and action taken in the field. They were further escalated to their field supervisors immediately and, in cases of a genuine risk to the participant or others surrounding them, were reported to Kalisa Godfroid a Senior Program Officer of the FSI/ECD Sugira Muryango Program. These flags noted in the survey were monitored in real time by the Senior Field supervisor, flagged in a Google Sheet that recorded actions taken and other comments, and were reviewed daily by the research team to ensure that the protocol was followed and escalated as appropriate. Cases of severe acute malnutrition encountered during the anthropometrics measurements and development delays encountered during the MDAT were also recorded in this Google Sheet and reported to Kalisa Godfroid. All risk of harm flags and their responses were communicated to Boston College using this Google Sheet.

DATA COLLECTION TRAINING

Surveys at the Household

Refresher training was held from August 5th to 9th, 2019. Refresher training for surveys completed at the household included an overview of the study's objectives and methodology, sampling and replacement strategies, research ethics, the field team's responsibilities, adverse event reporting and the risk of harm protocol, and a thorough review of the survey instruments. The review of survey instruments included presentations outlining:

- The Report on the Child survey, including:
 - All ASQ-3 surveys for 8- to 45-month-old children (ASQ-8-month to ASQ-42-month) with pictorial representations of milestones to ensure field team understanding;
 - The HOME survey;
 - The OMCI tool;
- The Caregiver Report on the Household survey; and
- The Caregiver Report on the Self survey.
- Change of caregivers
- Identification of additional caregivers and protocols

Field training was led by Amani Ntakirutimana, Data Manager at Laterite Rwanda, and presentations were adapted or created by Melissa Sutton, MD, Global Health Research Consultant at Laterite Rwanda. Kalisa Godfrod, Senior Program Officer of the FSI/ECD Sugira Muryango Program, attended the training for one day to train the enumerator team on the risk of harm and reporting protocols.

MDAT and Anthropometrics Surveys

Refresher training was held on August 5th - 9th, 2019 and was led by Senior Field Supervisor Juliet Kalimba. The training included a review of the questions and practices using the tool. On August 8, 2019, the teams attended one day of refresher training led by Jaya Chandna, a certified MDAT Trainer.

The training materials on the anthropometrics measurements were designed by Dr. Melissa Sutton, a medical doctor consulting for Laterite, prior to baseline in collaboration with the Laterite Senior Field Supervisor. The training included practical exercises on taking measurements as well as recording the information in the surveys.

SURVEY INSTRUMENT PILOTING

Surveys at the Household

Objectives

The endline pilot of the Surveys at the Household (Report on Child, Caregiver Report on Household, Caregiver Report on Self) was used as an opportunity for the enumerator team to re-familiarize themselves with the data collection tools, check for any errors in the coding of the research instruments, and to clarify any questions that were confusing to enumerators or respondents. Particular attention was given to questions that changed between midline and endline. In addition, on one of pilot additional data was collected to measure Inter-Rater Reliability (IRR).

Implementation

The pilot was held in Nyamirambo Sector of Kigali. The pilot took place over 2 days, Tuesday, August 13th and Wednesday August 14th 2019. We used the same pilot sample as baseline: we were able to find 44 of the original 48 households in the baseline pilot still living in Nyambirambo Sector. 25 households were interviewed on August 13th and the remaining 19 households were interviewed on August 14th.

On August 14th, the second day of pilot, we collected some additional data to measure IRR. For these 19 households, two enumerators simultaneously observed the Report on Child with one enumerator interviewing as normal and one interviewer silently observing and filling in the survey simultaneously.

Measure of Inter Rater Reliability (IRR)

Inter-rater reliability was evaluated with a kappa coefficient, κ , calculated using STATA version 15. We calculated IRR separately for the two observational modules of the Report on Child: the HOME Inventory and the OMCI. For the OMCI, the answer choices "very few" and "sometimes" are weighted as more similar than "never" and "often". The results are shown in **Figure 5** below. Overall agreement is high, as expected for simultaneous observations.

Figure 5. Inter-rater reliability estimates for simultaneous observations of the Report on Child

Battery	Percent Agreement	Kappa
HOME caregiver-child interaction (home_29 to home_49)	90.5%	0.81 (95% CI: 0.68-0.86)
HOME toys observation (home_13 to home_19)	99.1%	0.98 (95% CI: 0.79-1.00)
OMCI caregiver-child interaction (omci_1 to omci_19)	87.5%	0.77 (95% CI: 0.68-0.86)

MDAT

Objectives

The MDAT survey endline pilot was used as an opportunity for enumerators to re-familiarize themselves with the administration of the tool and to check for any errors in the coding of the research instruments. In addition, we collected test-retest data to measure IRR, thereby assessing enumerator consistency in implementing this complex pediatric developmental assessment.

Implementation

The MDAT pilot was held at the Centre Spirituel San Jose Carmelo in Kigali with a subset of the families interviewed during the baseline MDAT pilot. Families interviewed on the first day of the non-MDAT endline pilot were contacted to participate in the MDAT endline pilot. The MDAT survey endline pilot involved 27 children. To evaluate inter-rater reliability, children were assessed by a single enumerator and then reassessed by a different enumerator on the second day.

Measures of Rater Reliability (IRR)

Inter-rater reliability was evaluated with a kappa coefficient, κ , calculated using STATA version 15. Overall, inter-rater reliability was substantial with a $\kappa=0.62$ (s.e. 0.02) and an overall percentage agreement of **83%**. This is consistent with the IRR estimates from midline pilot.

DATA COLLECTION

Schedule

Endline data collection was carried out over six weeks from August 19, 2019 to September 30, 2019. Two weeks were spent in each district, and districts were surveyed sequentially starting in Rubavu, Nyanza then Ngoma. All data collection was completed in one district before data collection in the following district began.

The MDAT and anthropometric team followed behind the team completing the surveys at the household by two days, so that the MDAT survey was completed after the surveys at the household.

Completion Statistics

Data collection was conducted at endline in 1,029 of 1,049 households in the baseline sample. Details of the 20 households not surveyed at all at endline are included in Appendix 2. Because of the complexity of the endline “households”, we have not categorized households as “complete” or “not complete”. A breakdown of the number of clusters and households in the endline sample by treatment status and public works type is presented in Figure 7 and Figure 8, respectively.

Figure 7. Number of Clusters by Treatment Status in Endline Dataset

	Control	Treatment	Total
Classic PW	75 clusters	75 clusters	150 clusters
Expanded PW	38 clusters	47 clusters	85 clusters
Pooled	97 clusters	100 clusters	197 clusters

Figure 8. Number of Households by Treatment Status in Endline Dataset

	Control	Treatment	Total
Classic PW	362 households	362 households	724 households
Expanded PW	130 households	161 households	291 households
Pooled	492 households	523 households	1015 households

The final endline dataset includes 1,062 of the 1,085 children in the baseline sample. 22 of the 23 children not surveyed at endline were part of households not surveyed. In 1 household (hhid 114905 there is only a CGRH and CGRS with the partner of the primary caregiver). There are 36 households in the baseline sample with two eligible children. At midline, in two of the 36 households one child was not surveyed: in one household one child moved outside the districts in the study, in the other household one child was dropped because they were overage at baseline. At endline, in one of the households, one child was not surveyed because he/she moved to Kigali with the primary caregiver. Therefore, there are 35 households in the endline sample with two children surveyed. Details of the children not surveyed at endline are included in Appendix 3.

The final caregiver dataset includes 1,660 caregivers: 978 of the 1,049 primary caregivers in the midline sample, 410 of the 508 partners of primary caregivers in the midline sample, 22 new primary caregivers at endline, 28 new partners of primary caregivers at endline, and 215 additional caregivers. Additionally, we also interviewed 7 out of 16 caregivers that were the primary caregivers of the child

at baseline, but not at midline. Details of the caregivers not surveyed at midline are included in Appendix 3.

The numbers of each type of survey collected by baseline district and the number of each type of caregiver by baseline district are presented below in Figure 9 and Figure 10 , respectively.

Figure 9. Surveys completed for six survey instruments by baseline district

District	Number of households	Report on Household	Caregiver Report on Self	Report on Child	MDAT	Anthropometrics
Ngoma	343	344	539	346	353	353
Nyanza	344	348	553	350	350	354
Rubavu	342	345	568	353	352	353
Total	1,029	1,037	1,660	1,049	1,055	1,060

Figure 10. Caregiver surveys completed at endline by baseline district

District	Total number of caregivers	Caregiver Type						
		P1	P2	P3	P4	P6	P7	P5/G1
Ngoma	539	1	0	325	104	8	13	88
Nyanza	553	1	0	323	135	12	10	72
Rubavu	568	5	0	330	171	2	5	55
Total	1660	7	0	978	410	22	28	215

Note: see definition of caregiver types under: [Interview Protocols at Endline \ Caregivers](#)

P1 - Baseline (ONLY) Primary caregiver

P2 - Baseline (ONLY) partner of primary caregiver

P3 - Midline Primary Caregiver

P4 - Midline Partner of the primary caregiver

P6 - Endline Primary Caregiver

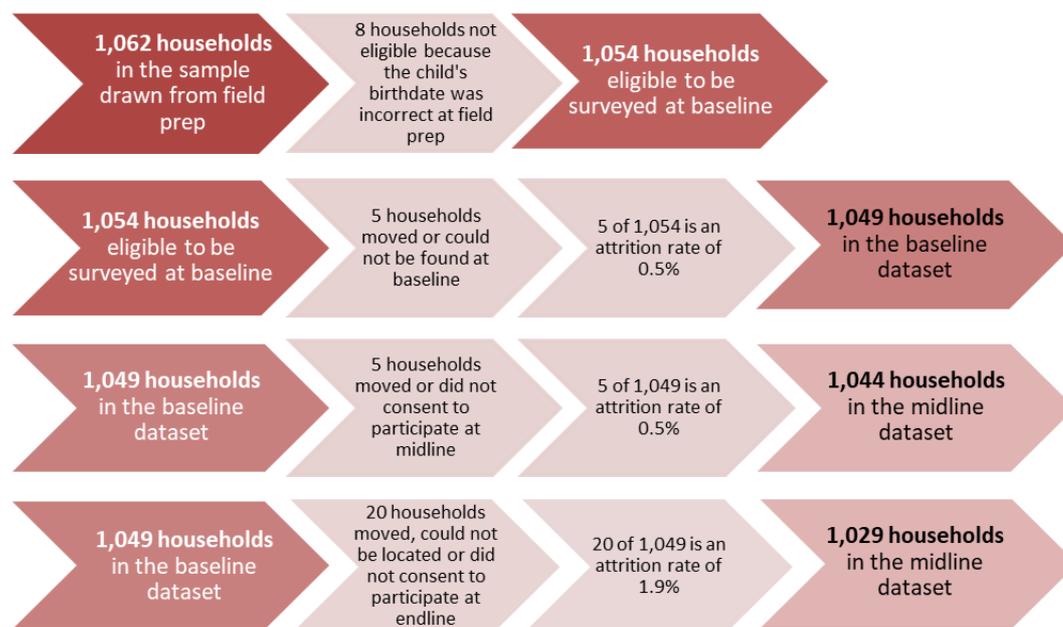
P7 - Endline partner of the primary caregiver

P5 - Additional Caregiver

Attrition

Between each wave of data collection, some households moved, declined to participate, or were otherwise not able to be surveyed. Figure 11 below outlines the attrition between the sample being drawn, baseline, midline and endline data collection. At endline, instead of considering “complete” households, we considered the number of households where we had *any* surveys complete.

Figure 11. Attrition in each wave of data collection



In addition to attrition of households, there is attrition of caregivers and children between waves of data collection. **Figure 12** outlines the attrition of caregivers and children between waves of data collection. The numbers used for baseline here are the number of caregivers and children in the 1,049 households in the baseline dataset, not the number surveyed at baseline. There was one primary caregiver, a number of partners of primary caregivers, and one child in the 1,049 households who were not able to be surveyed at baseline.

Figure 12. Respondent Attrition

Respondent Type	Baseline	Midline	Endline	Attrition * # (%)
Children	1,085	1,078	1,062	23 (2.1%)
Primary Caregivers (P3 & P1)	1,049	--	985	64 (6.1%)
Partners of Primary Caregivers (P4 & P2)	508	--	410	98 (19.3%)
Additional Caregivers (P5/G1)	--	269	215	54 (20.1%)
Children	1,085	1,078	1,062	23 (2.1%)

* Attrition is calculated between baseline and endline for children, primary caregivers, partners of primary caregivers; between midline and endline for additional caregivers.

NOTE: Attrition is calculated by looking at counts of caregiver type. However, some caregivers switch type between waves of data collection (e.g. some additional caregivers became new endline primary caregivers), and so these statistics should be interpreted with caution.

Administration of Surveys

Surveys at the Household

The Report on the Child, Caregiver Report on the Household, and Caregiver Report on Self were all performed in the home. The order in which the surveys were administered depended on whether the child was able to be assessed. If the child was present and awake, the Report on the Child was administered first, then the other surveys followed. If the child was asleep or not available, the

enumerator would start with the caregiver surveys instead. Due to the sensitive nature of the intimate partner violence questions in the Caregiver Report on Self, this survey was conducted only by enumerators of the same gender as the respondent and in as private a location as possible. In the case that a survey was not completed during the first visit—because there was a mismatch in the genders of the enumerator and respondent, for example—an appointment for mop up would be communicated to the household immediately and an enumerator would return at the earliest opportunity.

MDAT and Anthropometrics Surveys

Participants were scheduled to attend appointments for MDAT assessments and anthropometric measurements at centralized locations. In each cell Laterite in collaboration with community health workers selected a safe place—such as the cell office, a church, or school—to serve as the data collection site. Participants were notified in advance regarding the location and timing of these surveys.

On the day of the survey enumerators explained to caregivers that they would first play with the child then take measurements and reminded caregivers to do both sessions before leaving the venue. A dedicated enumerator took measurements for weight, height or length based on the age of the child, and MUAC and recorded observations on a paper form with a unique ID to identify the child. These measurements were entered into SurveyCTO at the earliest opportunity and always on the same day as data collection.

Survey Duration

Figure 13 below shows the median completion time for each type of survey at baseline, midline and endline. The anthropometrics measurements were first recorded on paper and then entered into the tablet later, so an accurate median time can't be calculated from the SurveyCTO output. Surveys at the household took less time at endline than midline or baseline, and respondents who had completed the surveys before were faster than their counterparts who were completing the survey for the first time. The MDAT took longer at endline, as the average survey time increases with child's age.

Figure 13. Median duration for each survey

Survey	Baseline Median Time (minutes)	Midline Median Time (minutes)	Endline Median Time (minutes)
Surveys at the Household			
Caregiver Report on Self	44	37	36
<i>P1: Baseline primary caregivers</i>	43	36	35
<i>P2: Baseline partners</i>	47	37	---
<i>P3: Midline primary caregivers</i>	---	42	35
<i>P4: Midline partners</i>	---	43	35
<i>P5/G1: Additional caregivers</i>	---	41	38
<i>P6: Endline primary caregivers</i>	---	---	43
<i>P7: Endline partners</i>	---	---	40

Caregiver Report on Household	30	20	20
Report on Child	53	49	49
MDAT Survey			
MDAT	47	52	61

Note: see definition of caregiver types under: [Interview Protocols at Endline](#) \ [Caregivers](#)

Participant Compensation

All study households received 5,000 RWF disbursed in two increments of 2,500 RWF at the completion of each of the household surveys and the MDAT, as compensation for their time and the cost of travelling to a centralized cell location for the MDAT and anthropometric assessment.

Challenges

The team encountered and addressed the following challenges during data collection:

- Errors in the participant list: There were a few households where caregiver names or other details were missing from the participant list because of missing or switched caregiver information in the field prep, baseline, and midline datasets. Enumerators were able to correct participant details in the surveys or add in missing details in the survey comments.
- The addition of new caregivers: Scheduling surveys was more difficult at endline as, despite making appointment calls to confirm caregiver details in advance, often the number of caregiver surveys required per household was not known until the household was visited. Where caregivers changed and additional caregiver surveys were required, a second mop-up visit to the household to complete the additional caregiver surveys was often required.
- Unavailability of a child in the household: There were cases where children went to visit other family or moved into a different household between appointment calls and data collection and so were not present at the household on the day planned for data collection. In these cases, enumerator teams followed up with other participants from the household, village leaders, CHWs and others to locate these children.
- Movement of participants: Between midline and endline, households and participants moved to both known and unknown locations. When the new location of the household was known, and inside the three research districts, the enumerator team interviewed participants at their new location and interviewed them there. No surveys were conducted if the new location of the household could not be identified. There are 36 households that have partially or completely moved from their midline location and could not be found at endline. In 13 of these households, the whole households had moved, and no surveys were done at endline.

DATA QUALITY MONITORING

Field Supervision

Throughout data collection, the field supervisor or senior field supervisor for the Surveys at the Household would choose one sub-team to accompany each day to confirm that interviews were conducted in the right households, protocols were followed, and ethical considerations were met. When enumerators raised issues during the day, the field coordinator reported them immediately to the field supervisor who communicated them to the Senior Field Supervisor then to the Data Manager for the final decision. The field supervisor would observe one or two surveys (with the exception of the Caregiver Report on Self since it included questions about particularly sensitive information) for each sub-team's member to confirm that questions were asked and answers recorded correctly. The field supervisor and senior field supervisor for the MDAT and anthropometric team roved among enumerators to check that assessments were conducted according to protocol and to offer guidance or corrections.

At the end of each day, the field supervisors updated the log of surveys completed and the Senior Field Supervisor updated the issue log sheet, highlighting any issues to be addressed in data cleaning, and compiled a daily field report for the Data Manager. The teams for both the surveys at the household and the MDAT and anthropometric assessments attended a daily debrief session led by the Senior Field Supervisor and Data Manager (if present). During the debrief, the team discussed issues from the day's surveys, and the Senior Field Supervisor or Data Manager provided solutions and communicated any changes in the survey. Before heading to the field each morning, the team met with the Senior Field Supervisor who updated the list of households with any replacements from the previous day, reviewed the schedule, and addressed quality concerns from the data collected in the previous days.

Real-time Completion Tracking

Laterite used SurveyCTO's feature of real-time publishing of form submission data into Google Sheets to track the progress of data collection compared to the targets in the field plan. A dashboard was created to track the completion of all surveys for each household. The Google Sheet dashboard was also used to track additional caregivers and households and flag duplicate surveys.

Routine Monitoring

Laterite used proprietary audit algorithms to review survey metadata to flag unusual submissions such as those with comparatively short or long durations or abnormal percentages of missing data. Concerning surveys were flagged to the Data Manager and Senior Field Supervisors for further investigation. In addition, Laterite routinely monitored the incoming data for survey duplication and outliers signaling potential measurement errors. Laterite reported survey completion statistics to Boston College in weekly Data Collection Reports.

Flagged Survey Responses

Several survey responses triggered the risk of harm protocol. These flags were monitored in real-time by field supervisors and senior field supervisors and twice weekly by the data monitoring team to ensure that the risk of harm protocol was followed appropriately. **Figure 14** shows the frequency of each flagged responses triggering the risk of harm protocol.

Laterite was asked by Boston College to flag several additional survey responses (see **Figure 15**). Responses were flagged and reported only if occurring more than 10 times in 3 months or, in the case of parq_4, if the response was, “almost always true.” Flagged responses were reported to Boston College in the weekly monitoring reports along with household information.

Figure 14. Flagged Responses Triggering the Risk of Harm Protocol

Variable	Question Text	Flagged Response	Instances Reported # (% of those asked)
Caregiver Report on Self			
hscl_20	Thoughts of ending your life (please tell us how many times it happened to you or stressed you last week including today)	A little / Quite a Bit / Extremely	62 (0.04%)
	Has any argument between you and your partner or most recent partner ever led you to do any of the following?		
cts_9	Used a knife or any other weapon to harm your partner	Yes	3 (0.21%)
cts_10	You hit your partner with something that could cause injuries to your partner	Yes	12 (0.85%)
cts_13	You used forceful means to be able to have sexual intercourses with your partner	Yes	17 (1.21%)
	Has your partner ever done any of the following things to you?		
cts_20	She/he used a knife or any other weapon	Yes	36 (2.55%)
cts_21	She/he hit you with an object that could cause injuries	Yes	83 (5.89%)
cts_24	She/he forced you have sexual intercourse when you didn't want	Yes	155 (10.99%)
Report on Child			
cd_12	Beat him/her up, that is hit him/her over and over as hard as one could (tell me if you or anyone else in your household has used this method in the past 30 days with [child name])	Yes	12 (1.15%)

Figure 15. Additional Flagged Responses

Variable	Question Text	Flagged Response	Instances Flagged # (% of those asked)
Caregiver Report on Self			
	Has any argument between you and your partner or most recent partner ever led you to do any of the following?		
cts_6	Shoving your partner	Yes	109 (7.7%)
cts_7	Grabbed your partner	Yes	59 (4.19%)
cts_8	Slapped or punched your partner	Yes	93 (6.6%)
cts_12	Kicked your partner	Yes	11 (0.78%)
	Has your partner ever done any of the following things to you?		
cts_17	She/he shoved you	Yes	207 (14.7%)
cts_18	She/he grabbed you	Yes	189 (13.41%)
cts_19	She/he hit you	Yes	247 (17.56%)
cts_22	She/he pushed you into a wall	Yes	145 (10.31%)
cts_23	She/he kicked you	Yes	105 (7.46%)
parq_4	I hit my child/children even when (s)he may not deserve it (Please pick how accurate you think these statements might be)	Almost always true	16 (0.96%)

Audio Audits

Audio from a sample of all surveys conducted at the household was automatically recorded using the in-built feature of SurveyCTO. The questions that triggered the audit were pre-selected and recordings were reviewed to assess whether (i) the interviews actually took place; (ii) enumerators were following proper interview procedures such as: explaining confidentiality and sticking to the script while asking sensitive questions; (iii) enumerators were asking questions with a respectful tone and without pushing the respondent or leading them towards a certain response; and (iv) explaining to the respondents that they could opt out of taking part (or answering questions in some parts of) in the survey. The questions audited are listed in Appendix 4. A random sample of 5% of all surveys were recorded for each question.

No audio audits were set for the MDAT and anthropometrics surveys since the proper administration of these assessments relied less on enumerator-participant dialogue. These surveys were instead administered under supervised conditions and enumerator behavior was closely monitored.

Throughout data collection, Kinyarwanda-speaking data auditors reviewed the audio recordings to confirm that there was dialogue between the enumerator and the respondents (i.e. that responses were not entered without actually asking the questions) and that enumerators were following proper interview procedures like explaining confidentiality and ensuring privacy for sensitive questions and asking questions in a respectful tone without pushing the respondent or leading them towards a certain response. Issues identified by the auditors were recorded and relayed to the data manager for proper follow-up and resolution with the field supervisors.

Inter Rater Reliability

Additional data was collected at endline to estimate inter-rater reliability for different components of administration of the survey instruments. This was done in addition to the existing data that has been collected for MDAT and Anthropometrics during each wave of pilot. Of particular interest was capturing information on the observational components of the Report on Child, which are most vulnerable to variation in administration.

Design

To assess reliability Laterite conducted three types of inter-rater reliability assessments:

1. **Simultaneous observation** by two enumerators for the Report on Child: the modules of the survey most of interest here were **the caregiver-child interactions observed in the OMCI and in the HOME** (home_20 to home_49). Though those two modules were most of interest, enumerators observed the full Report on Child as it was logistically easier to implement and may limit bias among enumerators as it will be less explicit exactly what was being measured in the simultaneous observation. Enumerators were instructed to observe independently (i.e. not communicate with each other about their responses during the assessment) We employed a variety of monitoring strategies we have used in other studies to try and flag any suspicious responses, for example by auditing surveys to see if enumerators communicate with each other during the paired assessments.
2. **Test/re-test reliability** for the **observational component on the home environment**, questions home_13 to home_19 where the enumerator observes the toys in the home. As this is the only part of the HOME or OMCI measuring something static, that is not expected to change substantially between assessments, this is the only module where test/re-test measurements provide meaningful information on variation between enumerators, as opposed to variation among respondent answers.
3. **Test/re-test reliability** for the **anthropometrics measurements**. These additional assessments were done both during pilot and during data collection. During data collection each type of assessment was implemented by District, as follows:
 - **Rubavu** - We started data collection in this district because of the risk of ebola spreading into this district: no additional assessments were done as the aim was to complete data collection as quickly as possible.
 - **Nyanza** - The data collection protocol was modified slightly to incorporate test/re-test reliability measurements. Every secondary caregiver interview included an additional module with the seven additional questions about the toys available in the home environment. This served as a second, independent assessment of this construct for the household.
 - **Ngoma** – The data collection protocol was modified slightly to incorporate simultaneous observations of the Report on Child for at minimum 10% of the households in the district. No test/re-test assessments for the home environment were done.

Results

Report on Child

Inter-rater reliability was evaluated with a kappa coefficient, κ , calculated using STATA version 15. We calculated IRR separately for the two observational modules of the Report on Child: the HOME Inventory and the OMCI. For the OMCI, the answer choices "very few" and "sometimes" are weighted as more similar than "never" and "often". The results are shown in **Figure 16** below. Overall agreement is high, as expected for simultaneous observations.

Figure 16. Inter-rater reliability for components of the Report on Child

Measure	Percent Agreement	Kappa
ASQ	98.17%	0.963 (95% CI: 0.994-0.932)
HOME: Simultaneous Observation	91.87%	0.835 (95% CI: 0.86-0.809)
HOME: test-retest reliability	86.00%	0.631 (95% CI: 0.67-0.592)

Anthropometrics

Inter-rater reliability was evaluated with inter-class correlation (ICC) calculated using STATA version 15. We calculated ICC separately for each anthropometrics measure. The results are shown in **Figure 17** below. Overall the results show the team has high consistency in taking anthropometric measurements: height and weight had near perfect agreement, while MUAC had very high agreement.

Figure 17. Inter-rater reliability for components of the Report on Child

Measure	ICC
Weight	0.999 (95% CI: 0.999 - 0.999)
Height	0.999 (95% CI: 0.998 - 0.999)
MUAC	0.977 (95% CI: 0.958 - 0.988)

DATA CLEANING

Laterite cleaned data iteratively throughout data collection to identify issues as they arose. At the end of endline data collection, a master data cleaning file was compiled that included all deduplication and cleaning done throughout data collection, as well as the merging of datasets, de-identification of the dataset, coding of missing values, and other steps to prepare the data for Boston College. All data cleaning was completed in STATA version 15.

Duplicates Management

Duplicates identified during data quality monitoring by unique identifiers were managed as follows:

- True duplicates (all variables identical): one copy was deleted using the *duplicates drop* command.
- Partial duplicates (not all variables identical): both copies were flagged using the *duplicates tag* command, removed from the dataset, logged, and investigated by the Data Manager. Following investigation, the Master Corrections .csv file and/or the STATA .do file (depending on the issue identified) were updated and the duplicates were cleaned with the subsequent round of data.

Duplicates reports were generated biweekly during the first week of data collection then with decreasing frequency as field errors decreased and data quality improved.

Master Datasets

The data cleaning process involved reshaping the raw dataset from SurveyCTO, containing all form submissions, and transforming it into cleaned, de-duplicated, de-identified datasets. Reshaping and merging the results of the five surveys, Laterite created three datasets: Caregiver, Household, and Child. The folder **14-Endline Data Cleaning** contains all the necessary files (raw data .csv files, .dta files, auxiliary files & .do files) used to produce the final datasets. Details about the inputs and structure of the files is outlined Readme.docx file included with the data sets.

The first step of the data cleaning process converted the raw data into the labelled Stata format dataset by running the import .do files that are output from SurveyCTO to transform the raw .csv files into STATA .dta datasets. This step was repeated for each version of each of the five surveys: Anthropometrics, Caregiver Report on the Household, Caregiver Report on Self, MDAT, and Report on Child. One of the inputs to this .do file is the survey-specific corrections file, which is used to make changes to values in the dataset by referring to the submission key and variable name. The corrections files are updated manually by the data monitoring and cleaning team at Laterite to incorporate changes suggested by the data collection team in the field or address values recorded erroneously.

At midline, for variables like name that might be used to verify identity, corrections were not made in the corrections file, but instead added to a separate “corrections_midline_surveys” document and merged into the dataset as a new variable (e.g. “name_corrected”). At endline, there were very few corrections to survey information, and therefore these were added to the main corrections csvs.

Data cleaning is done in one .do file for each dataset. The data cleaning process involved appending survey versions, resolving duplicates, pulling relevant information from the field prep data, comparing caregiver information across field prep, baseline, midline, and endline data collection, merging in the translations of enumerator comments, labelling variables, and re-coding skipped or missing values according to the Boston College team's guidance. The child data cleaning files also reshape the MDAT data. The final outputs include at least one survey for **1,029 households**, including:

- Caregiver dataset with **1,660** observations from the Caregiver Report on Self
- Household dataset with **1,037** observations ("houses") from the Caregiver Report on Household, representing **1,020** baseline households.
- Child dataset includes **1,062** observations, including **1,049** complete Report on Child surveys, **1,055** complete MDAT assessments, and **1,060** complete anthropometric measures.

In the cleaned endline datasets, we differentiate between a **household**, the primary sampling unit, and a **house**, which is each separate physical residence. By definition there was only one house per household at baseline. At midline and endline, because of changes in the family structure over time, there may be more than one separate residence, or house, per household.

Observations from the three datasets can be linked using the household ID variable (hhid), a unique identifier for each household. Children can be further linked to the house where they live using house ID (house_id), a unique identifier for each residence within a household. Currently house_id is not in the Caregiver dataset and cannot be used to link caregivers to their house. Household ID (hhid) should be used to link surveys between time points; house ID should not be used to link between time points. **Figure 18** shows the identifying variables in each dataset.

Caregiver IDs

A caregiver who changed caregiver type between waves of data collection (e.g. was a G1 at midline but the primary caregiver at endline) kept the same cg_id so that they can be tracked between waves of data collection, but was given a new caregiver type in the endline dataset.

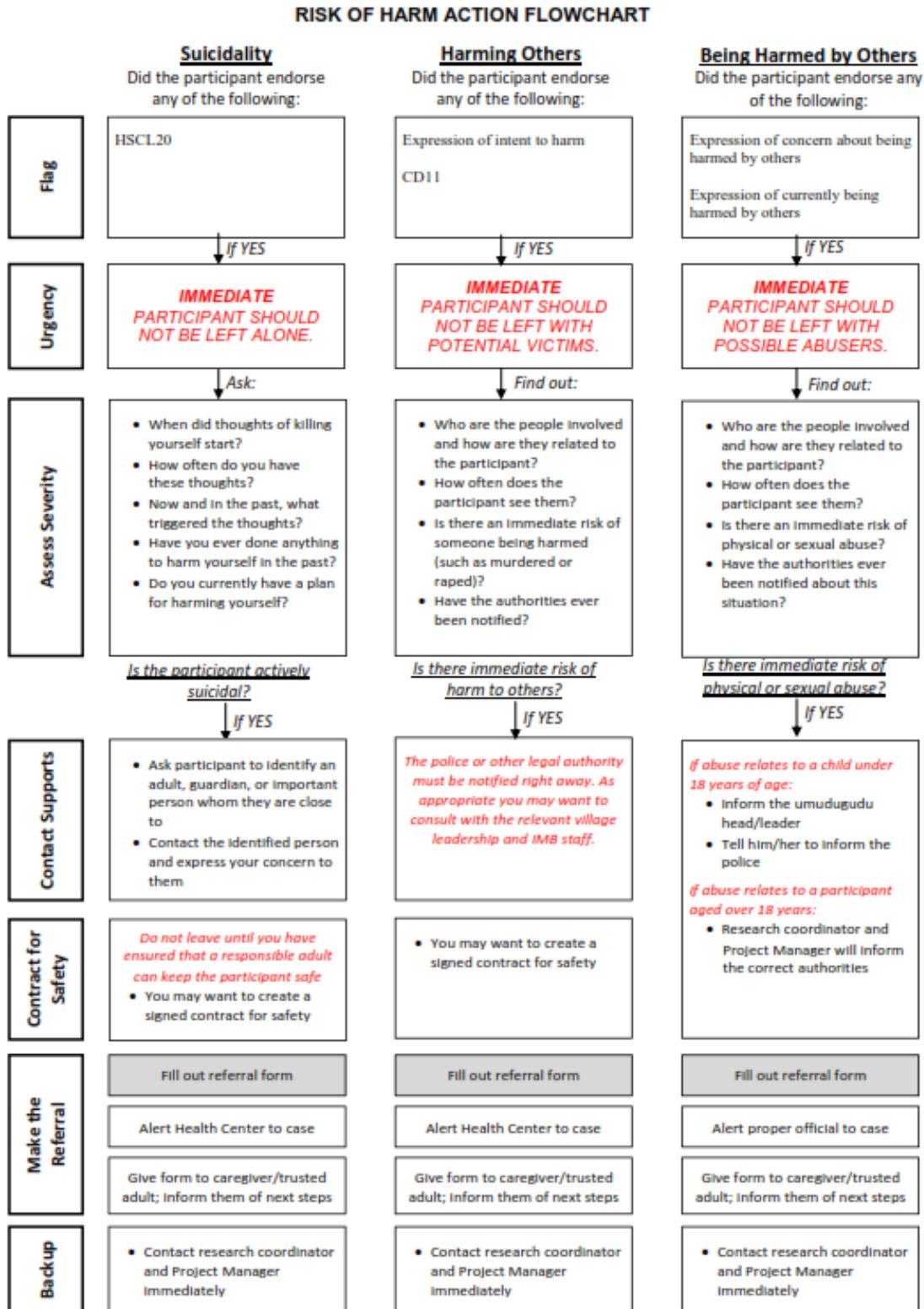
Caregiver IDs were created using the following naming conventions, for the role the respondent had in the household when they were first interviewed. The first number in the ID indicates the caregiver type, and the second number indicates the wave of data collection they were first interviewed at.

- **Primary caregivers:**
 - **P1:** 200000 - 219999
 - **P3:** 220000 - 239999
 - **P6:** 240000 - 259999
- **Alternate Caregivers:**
 - **G1:** 620000 – 639999
- **Partners of primary caregivers:**
 - **P2:** 300000 - 319999
 - **P4:** 320000 - 339999
 - **P7:** 340000 - 359999

Figure 18. Identifying variables in each dataset

Household Dataset	Caregiver Dataset	Child Dataset
<p>Unique ID: house_id hhid links houses in the same household hhid is consistent across all three timepoints of data collection child_in_hh flags which household the child resides in ...</p>	<p>Unique ID: cg_id hhid links caregivers in the same household cg_id is consistent across all three timepoints of data collection. caregiver (caregiver type) is not consistent across time points, but reflective of the position of the individual in the household at each time point.</p>	<p>Unique ID: child_id house_id specifies which house in a household the child lives in hhid links children in the same household child_id is consistent across all three timepoints of data collection</p>

Appendix 1. Risk of Harm Action Flowchart



Sugira Muryango Endline Field Report

Appendix 2. Entire household missing

hhid	Child 1 ID	Child 2 ID	P3 ID	P4 ID	G1 ID	P1 ID	P6 ID	P7 ID	Comments
100768	404498		200761	300406					Child died
101000	404781		201004		625927				Household/ child moved outside Area
101948	402631		201945	300877					Caregiver(s) not available /cannot be found during visits
102267	403130		202261	301052					Household/ child moved outside Area
104231	403806		204234						Child died
104253	401759		204250		638427				Household/ child moved outside Area
104622	400019		204625						Child died
105584	405586		205585		627522				Caregiver(s) not available /cannot be found during visits
107564	407565		207566	303541					Household/ child moved outside Area
109227	409226		209222		625263				Household/ child moved outside Area (Uganda)
109269	409260		209265						Household/ child moved outside Area
110309	410306		210308		628451				Household/ child moved outside Area
113126	413127		213128	306327					Caregiver(s) not available /cannot be found during visits
113485	413487		213482	306547					Caregiver not available or could not be found during visit
113615	413618		213612		626086				Household could not be located
114181	414189		214183	306919					Household/ child moved outside Area
114682	414685		214689						Household Refused to Participate
115069	415065		215061	307394					Mother and child moved to Uganda
115435	415438		215433						Child died
115443	415446		215440	307597					Child died

Appendix 3. Partially Complete Households: Child, Caregiver, or Household Surveys missing

hhid	Child IDs	Caregiver IDs	Child 1			Child 2			CGRS							CGRH	Comments	
			Ant	MDAT	ROC	Ant	MDAT	ROC	P1	P2	P3	P4	G1	P6	P7			
100981	402851	200988	1	1	0				0	0	0	0	0	0	0	0	0	Caregiver(s) not available /cannot be found during visits
101854	403811	201853	1	1	0				0	0	0	0	0	0	0	0	1	The primary caregiver who is the child's uncle wasn't available while the team was in Nyanza
102549	402224	202540 628977	1	1	0				0	0	0	0	1	0	0	1	1	Caregiver not available or could not be found during visit

Sugira Muryango Endline Field Report

hhid	Child IDs	Caregiver IDs	Child 1			Child 2			CGRS							CGRH	Comments
			Ant	MDAT	ROC	Ant	MDAT	ROC	P1	P2	P3	P4	G1	P6	P7		
103079	402690	203074 638488	1	1	0				0	0	0	0	1	0	0	1	Caregiver not available or could not be found during visit
105513	405510	205512 302696	1	1	0				0	0	0	0	0	0	0	0	Household/ child moved outside Area
105716	405712	205710 302798	1	1	0				0	0	0	0	0	0	0	0	This caregivers separated, the mother moved to outside the district with her child
107111	407110	207117 639846	1	1	0				0	0	0	0	0	0	0	0	Caregiver moved outside Area
107230	407236	207232	1	1	0				0	0	0	0	0	0	0	0	Household/ child moved outside Area
109402	409408	229397 620996 209401	1	1	0				0	0	0	0	0	0	0	0	Household/ child moved outside Area
109918	409911	209919 304456	1	1	0				0	0	0	0	0	0	0	0	Household/ child moved outside Area
110736	410734	210739 304882	1	1	0				0	0	0	1	0	0	0	0	Caregiver not available or could not be found during visit (primary caregiver in hospital)
114905	414908	214907 307305	0	0	0				0	0	0	1	0	0	0	1	Caregiver moved outside Area
100505	400478	200507 638084	1	1	1				0	0	0	0	1	0	0	1	Caregiver not available or could not be found during visit
100526	404175	237857 200529 359193	1	1	1				0	0	0	0	0	0	1	1	Caregiver not available or could not be found during visit
100539	404517	200532 633140 357202	1	1	1				0	0	0	0	1	0	1	1	Caregiver not available or could not be found during visit
100759	401051	200757 635503	1	1	1				0	0	0	0	1	0	0	1	Caregiver not available or could not be found during visit
101308	400949	201307 636391	1	1	1				0	0	0	0	1	0	0	1	Caregiver not available or could not be found during visit
101887	404073	201889 634965	1	1	1				1	0	0	0	0	0	0	1	Caregiver not available or could not be found during visit
103318	403319	203311 301669	1	1	1				0	0	0	1	0	0	0	1	Caregiver not available or could not be found during visit
104081	401265	204086 620493	1	1	1				0	0	0	0	1	0	0	1	Caregiver moved outside Area
104407	400795 500018	204402 621177	1	1	1	1	1	1	0	0	0	0	1	0	0	1	Caregiver moved outside Area
106132	406135	206136 626242 341334	1	1	1				0	0	0	0	0	0	1	1	Caregiver moved outside Area
106141	406143	206148 626055	1	1	1				0	0	0	0	1	0	0	1	Caregiver not available or could not be found during visit
106476	406477	206472 639791	1	1	1				0	0	0	0	1	0	0	1	Caregiver not available or could not be found during visit
106634	406630	206632 637966 358960	1	1	1				0	0	0	0	0	0	1	1	Caregiver not available or could not be found during visit
107483	407484	207488 303507 620253	1	1	1				0	0	0	1	1	0	0	1	Caregiver not available or could not be found during visit
107656	407650	207651 303593	1	1	1				0	0	0	1	0	0	0	1	Caregiver not available or could not be found during visit
107884	407881	207887 629239	1	1	1				0	0	0	0	1	0	0	1	Caregiver moved outside Area
109649	409645	209648 304328	1	1	1				0	0	0	1	0	0	0	1	Caregiver not available or could not be found during visit

Sugira Muryango Endline Field Report

hhid	Child IDs	Caregiver IDs	Child 1			Child 2			CGRS							CGRH	Comments
			Ant	MDAT	ROC	Ant	MDAT	ROC	P1	P2	P3	P4	G1	P6	P7		
111246	411249	231098 636594 211242 355098	1	1	1				1	0	0	0	1	0	1	1	Caregiver not available or could not be found during visit
111465	411460 500459	211466 305336	1	1	1	1	1	1	0	0	0	1	0	0	0	1	Caregiver not available or could not be found during visit
112005	412003	212008 307953	1	1	1				0	0	0	0	0	0	0	1	Caregiver has a mental disability and was not able to be surveyed
112325	412328	212320 305780	1	1	1				0	0	0	1	0	0	0	1	Caregiver not available or could not be found during visit
112897	412896	212894 306166	1	1	1				0	0	0	0	0	0	0	1	Caregiver not available or could not be found during visit
113448	413440	213449 306526	1	1	1				0	0	0	1	0	0	0	1	Caregiver not available or could not be found during visit
114952	414956	214955 628641	1	1	1				0	0	0	0	1	0	0	1	Caregiver not available or could not be found during visit
115138	415136	215134 307451	1	1	1				0	0	0	1	0	0	0	1	Caregiver not available or could not be found during visit
115585	415581	215582	1	1	1				0	0	0	0	0	0	0	1	Caregiver has a mental disability and was not able to be surveyed
101241	402838	201246 300605	1	0	1				0	0	1	0	0	0	0	1	
101335	402126	201332 637268 244591	1	0	1				0	0	0	0	0	1	0	1	Caregiver not available or could not be found during visit
101397	401023	201393	1	0	1				0	0	1	0	0	0	0	1	
101475	403000	201479	1	0	1				0	0	1	0	0	0	0	1	
109216	409218 500301	209211 624085	0	0	0	1	1	1	0	0	1	0	1	0	0	1	Child dropped since overage at baseline
109234	409239 500317	209233 631413	1	1	1	0	0	0	0	0	1	0	0	0	0	1	G1 moved outside area, ROC/CGRS missing
109848	409849 500349	209841 307869 627579	1	1	1	1	1	0	0	0	1	1	1	0	0	1	
113365	413364	213369 629408	0	0	1				0	0	1	0	1	0	0	1	
114203	414206	214208 306935	1	0	0				0	0	1	1	0	0	0	1	MDAT and ROC missing because child was sick
115599	415597	215596 307717	0	0	1				0	0	1	1	0	0	0	1	The MDAT and Anthro are missing because the child was sick

Appendix 4. Questions in the Audio Audit

Variable	Question Text	Why Audited
Caregiver Report on Household		
vup_1 to vup_3	vup_1; What type of public works job have you participated in? vup_2; About how many months have you/your household worked in Public Works (VUP) in the past 3 months? vup_3; Are you or is someone in your household currently working in public works (VUP)? (By currently, we mean within the past week)	Ensure enumerator stuck to the script; that it was clear to the respondent that the questions were <i>about the past week</i> ; and that the enumerator was not rushing or being rude to the respondent
fc_1 to fc_3	fc_1; In total, how many people live in your household? fc_2; In total, how many children 17 years old or younger live in your household? fc_3; How many children under 5 are in the household?	Ensure that the enumerator asked the questions in a clear and easy to understand manner; that the question script wasn't altered; and that the enumerator was courteous and not rushing
ha_ws12 to ha_ws14	ha_ws12; Do you have any soap or detergent or ash/mud/sand in your household for washing hands? ha_ws13; Do you normally wash your hands after you have been to the toilet? has_ws14; The last time your children passed stools, what was done to dispose of the stools?	Find out if the respondent was at ease; ensure the enumerator was reading the questions as scripted; the interview was not rushed and the enumerator was courteous
Caregiver Report on Self		
alc_inst to alc_2	alc_inst; INSTRUCTIONS: Due to the fact that alcohol carries some health risks and can interfere with some medications, it is important that we ask you questions about your alcohol consumption. Your answers will be kept confidential, that's why we ask you to be honest. Try to answer questions about alcohol beverages. Ask for clarification whenever needed: alc_1; How many times do you drink? alc_2; How many drinks containing alcohol do you have on a typical day when you are drinking?	Ensure reminder of confidentiality was read as written; that the respondent acknowledged having heard and understood the confidentiality reminder; Check that the enumerator was clear, courteous, at ease, not rushed and not harsh.
cts_inst to cts_note1	cts_inst; INSTRUCTIONS: Now I am going to ask you questions about your partner (boyfriend/girlfriend/husband/wife). cts_1; Have you ever been married or had a partner?	Check that the enumerator ensured no one else was present before asking these questions; Ensure that the enumerator clearly explained to the

	<p>cts_2; Are you married now or have a partner?</p> <p>cts_3; Do you currently live with your partner?</p> <p>cts_note1; INSTRUCTIONS: These questions might remind you of difficult events or circumstances. At any point if you do not wish to answer these questions let us know and we can skip a particular question or move to the next section.</p> <p>Has any argument between you and your partner or most recent partner ever led you to do any of the following?</p>	<p>respondent that they could choose to opt out of answering this set of questions; that the enumerator received acknowledgement from the respondent of having understood the instructions; and that the enumerator was clear, courteous, at ease, and not harsh.</p>
Report on Child		
<p>chh_momhiv to chh_hivb</p>	<p>chh_momhiv; Let me remind you that all of your answers are confidential, and that the information you provide is very important for the survey. Did you/the mother know if you were HIV+ when you were pregnant with \${_calc_child_name}?</p> <p>chh_momhivb; If HIV+, did you/the mother receive treatment to prevent transmission of HIV to your baby?</p> <p>chh_hiv; Has \${_calc_child_name} ever been tested to see if he/she has HIV?</p> <p>chh_hivb; Let me remind you that all of your answers are confidential, and that the information you provide is very important for the survey. Could you please tell me what was the result of your child's last test for HIV?</p>	<p>Ensure reminder of confidentiality was read as written; Check that the enumerator was clear, courteous, not rushed and not harsh; and that the respondent was at ease.</p>
<p>_check_child_bday to note_confirm_age</p>	<p>_check_child_bday_confirm; Was [child name] born on [child birthday]?</p> <p>child_bday_doc; What document was used to verify the date of birth?</p> <p><i>[non-spoken questions omitted here]</i></p> <p>According to our records, [child name] is [age in months] months old today. Can you confirm that this is correct?</p>	<p>Check that birthday corrections were being done correctly, and to identify any issues in confirming birthday documentation; Check that the enumerator read the question as written, and that they were clear, at ease, and not harsh or forceful.</p>
<p>cd_inst to cd_13</p>	<p>Gross motor objects available (e.g., ball, rope, ring, stone).</p>	<p>Check that the enumerator was clear, at east, not rushed, and that the enumerator looked for the objects mentioned or the participant brought them to the enumerator.</p>