



MOZAMBIQUE McGovern-Dole  
International Food for Education and  
Child Nutrition Project  
Educating Children Together Phase 3  
(ECT3)

Baseline Evaluation

1/14/2022

# Educating Children Together Phase 3 (ECT3) Baseline Evaluation Report

Program: McGovern-Dole International Food for Education and Child Nutrition

Agreement Number: FFE-656-2015-011-00

Funding Year: Fiscal Year 2021

Project Duration: FY2020 – FY2025

Implemented by: WORLD VISION

Evaluation Authored by: INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE  
(IFPRI)

FELICIANO CHIMBUTANE, CARLOS LAUCHANDE, CATALINA HERRERA-  
ALMANZA, NAUREEN KARACHIWALLA, JESSICA LEIGHT

DISCLAIMER: This publication was produced at the request of the United States Department of Agriculture. It was prepared by an independent third-party evaluation firm. The author's views expressed in this publication do not necessarily reflect the views of the United States Department of Agriculture or the United States Government.

## List of Acronyms

CESC	Centro de Aprendizagem e Capacitação da Sociedade Civil
CI	Confidence interval
CNBS	Comité Nacional de Bioética em Saúde
COVID-19	Coronavirus disease
CSB	Corn soy blend
ECT	Educating Children Together
EGRA	Early Grade Reading Assessment
IFPRI	International Food Policy Research Institute
INDE	Instituto Nacional de Educação / Institute for the Development of Education
LB	Literacy Boost
MINEDH	Ministério de Educação e Desenvolvimento Humano / Ministry of Education and Human Development
McGovern-Dole	McGovern-Dole International Food for Education and Child Nutrition Program
NGO	Non-Governmental Organization
RCT	Randomized controlled trial
SD	Standard deviation
SE	Standard error
SDEJT	District Services of Education, Youth and Technology
TARL	Teaching at the Right Level
UL	Unlock Literacy
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
USDA	United States Department of Agriculture
WV	World Vision
ZIP	Zona de Influência Pedagógica / Pedagogical Influence Zone

# Table of Contents

<b>Executive Summary</b> .....	6
<b>1 Introduction</b> .....	8
<b>2 Evaluation Design</b> .....	12
<b>2.1 Randomized controlled trial: quantitative methods</b> .....	13
<b>2.2 Baseline survey</b> .....	14
<b>2.3 Quality assurance and safety</b> .....	19
<b>2.4 Data Management and Security</b> .....	19
<b>2.5 COVID-19 Specific Precautions</b> .....	20
<b>3 Strategy for analysis</b> .....	21
<b>3.1 Summary statistics</b> .....	21
<b>3.2 Baseline balance</b> .....	21
<b>3.3 Determinants of reading</b> .....	22
<b>3.4 Strengths and limitations</b> .....	23
<b>4 Results</b> .....	24
<b>4.1 Summary Statistics</b> .....	24
<b>4.2 Summary of Performance Indicators</b> .....	35
<b>4.3 Baseline Balance</b> .....	37
<b>4.3 Determinants of Reading</b> .....	46
<b>5 Conclusion</b> .....	50
<b>References</b> .....	52

## List of Tables

Table 1. Number of Observations.....	11
Table 2. Sample Sizes .....	15
Table 3. Cronbach's alpha reliability coefficient.....	18
Table 4. Summary statistics for grade 4 students .....	25
Table 5. Summary Statistics of Grade 4 EGRA Scores .....	28
Table 6. Summary Statistics of Teachers.....	30
Table 7. Summary Statistics of Deputy School Directors.....	33
Table 8. Performance Indicators .....	35
Table 9. Baseline balance, student characteristics .....	39
Table 10. Baseline balance, teacher characteristics .....	43
Table 11. Baseline balance, deputy school director characteristics .....	45
Table 12. Correlation between performance in reading fluency and social, economic and linguistic variables.....	46
Table 13. Correlation between reading fluency and teachers' characteristics and school directors' profiles.....	47
Table 14. Estimation of the variance components in the empty model .....	48
Table 15. Estimates for two-level models-reading.....	49

## List of Figures

Figure 1. Study Design .....	13
Figure 2. EGRA Scores by Gender .....	27

## Executive Summary

The objective of this report is to present an overview of the baseline survey conducted for the USDA funded Educating Children Together – Phase 3 (ECT3) impact evaluation by IFPRI. World Vision has been implementing the program in two previous phases (ECT1 & ECT2) since 2013 delivering a comprehensive school meals program to all schools in two districts in Nampula province (Muecate, Nacaroa). As part of the current phase three (ECT3) of the project, World Vision is also implementing additional programming in all the target schools. This includes an early grade literacy program called Unlock Literacy (UL) that entails teacher training in early grade literacy and community participation. The teacher training comprises training of all teachers in the target schools on good pedagogical techniques for early-grade reading. The community participation element entails community meetings and training sessions for parents in supporting literacy, and reading camps in which a literacy volunteer in the community meets with children weekly outside of school (usually on the weekends) to create fun reading exercises and motivate children to read.

IFPRI is conducting the ECT3 impact evaluation using a randomized controlled trial designed to analyze the effects of Unlock Literacy implemented in conjunction with school meals compared to the provision of school meals, health and nutrition (all schools will receive the school meals program). The sample includes 160 rural public primary schools in two districts (Nacaroa and Muecate) in Nampula province. Randomization is conducted at the level of the school. All 160 schools will receive school meals. Fifty of those schools are assigned to each of two treatment arms; in one treatment arm, schools will receive meals and Unlock Literacy teacher training programming, and in the second treatment arm, schools will receive meals, UL teacher training and community action programming. The remaining 60 schools will be assigned to the control arm and will receive school meals only. The evaluation design aims to answer the following questions:

- What is the effect of literacy training on students' literacy and other academic outcomes?
- Does enhanced community action further improve children's literacy outcomes?
- Do the effects of the program differ by baseline characteristics, including child gender, household socioeconomic status, and a baseline measure of school quality?
- What is the correlation between the home environment as it pertains to literacy (parental engagement, book ownership, parental support) and child literacy outcomes?

The baseline survey for this impact evaluation was conducted between June and August 2021 by ELIM Serviços working under the supervision of IFPRI; the full sample of surveyed respondents included 1,522 students, 160 teachers, 160 deputy school directors, and 157 school cooks. The key topics included in the surveys included school characteristics and resources, pedagogical methods, participation in prior interventions including ECT2, exposure to COVID-19 related shocks, household characteristics and literacy practices (reported by students), and literacy assessment (as measured by the administration of the EGRA).

In this report, we present a range of summary statistics to characterize the sample and better understand the context for the evaluation. We present evidence that baseline characteristics are suggestive of meaningful challenges linked to low attendance and poor nutrition in the sample, as well as low baseline performance on the EGRA exam. In general, the majority of baseline characteristics are balanced across treatment arms as would be expected given the randomized design. This will allow us to identify the effects of ECT3 programming by comparing covariates across treatment arms

at endline program performance evaluation in 2024 following the impact evaluation midline data collection in two years (in 2023).

# I Introduction

Human capital investments in early years have important implications for long-term resilience and poverty; better nutrition and higher levels of learning improve access to and productivity in work and thus incomes, with impacts on future generations as well (Akresh et al., 2018; Maluccio et al., 2009). However, relatively little is known about how complementarities between nutrition and learning can be exploited to enhance human capital outcomes. In Mozambique, deficits in both health and education are large and persistent: anemia prevalence was 49.5 percent in rural areas among children 0–5 in 2018 (SETSAN, 2018). Mozambique had one of the lowest educational attainment rates in the world, characterized by primary completion below 40% (Mambo et al., 2019), with less than one-third progressing to secondary school (UNESCO Institute of Statistics). According to the last grade three national reading assessment (INDE/MINEDH, 2017), Nampula is one of the provinces that performed below the national average, especially in more remote and less developed districts.

Extensive empirical evidence relates child health and nutrition to the school performance of children. In particular, early childhood stunting (Mendez and Adair, 1999) and anemia (Soemantri et al., 1985) have been associated with poor performance on tests and cognitive assessments. Acknowledging this association, many governments have established school feeding programs with the intent of improving school performance through reduced undernutrition. However, despite the popularity of these programs in both developed and developing settings, while there is abundant evidence on their positive impacts on nutrition and enrolment (Alderman and Bundy, 2012; Drake et al., 2017), there is a relative dearth of evidence regarding their effectiveness on learning outcomes (Aurino et al., 2020). Moreover, these studies suggest that while school meals do raise learning outcomes, they may not be sufficient when complementary educational inputs such as teachers, pedagogy, or infrastructure are lacking or are of poor quality. Consequently, the potential for the addition of an early child literacy program to a school meals program may be substantial.

In fact, within many schools in developing countries, teachers still use a call and answer style of teaching in which instruction takes the form of a lecture with little student interaction (Glewwe and Muralidharan, 2016). This method offers little scope to differentiate instruction to account for the large heterogeneity in preparation levels often observed in early grade classrooms. It is likely that such a pedagogy leaves many children behind, and recent empirical reviews have shown that “teaching at the right level” (TARL) pedagogical interventions generally have positive impacts on test scores in both reading and math. However, there is often some difficulty in getting teachers to adopt different teaching methods (Glewwe and Muralidharan, 2016; Muralidharan, 2017; Beg et al., 2019). The weight of the evidence on teacher training implies that programs can be effective but also shows that choices in program design can have large effects on program efficacy. Similarly, to foster more inclusive education, UNESCO (2020) recommends the adoption of Universal Design Learning-UDL at the government level. UDL is a practice aimed at the inclusion of all learners regardless of the barriers they can face arising due to inequality, gender discrimination, poverty, and disability stigma, among other factors. Nevertheless, a recent review in low-and-middle-income countries indicates that for UDL to be a successful strategy, a better understanding of the local teaching practices is necessary to overcome teacher’s resistance to its implementation (McKenzie et al., 2021).

In a recent review, Popova et al. (2021) conducted a systematic review of 39 separate teacher training programs. Their review shows that teacher training programs generally display gains in child test scores, but that these gains are substantially larger if there are incentives associated with the training,



the training is focused on a specific subject, the training takes place in a face-to-face setting, or is focused on lesson enhancement. Longer trainings and those accompanied by follow-up and coaching tend to be even more effective. More recently, other community-level programs have been added to TARL-type of interventions. Usually, these community-level interventions aim to increase involvement of parents and communities in literacy promotion, for example, community meetings, reinvigoration of school councils, increasing parental involvement, and reading camps. However, the current evidence does not indicate that these community-level interventions are necessarily successful in improving children's learning outcomes (Barrera-Osorio et al., 2020).

The International Food Assistance Division (IFAD) of the United States Department of Agriculture (USDA) funds school meals in many developing countries, including Mozambique, where meals and literacy programming are delivered by a number of implementing organizations, including World Vision. IFPRI is partnering with World Vision to evaluate the impacts of the 5 year project **Educating Children Together - Phase 3 (ECT3)** on nutrition and literacy outcomes among primary school-age children in Mozambique. ECT3 is an integrated project funded at \$25.3m and involves school meals, enhanced community action intervention combined with a literacy model called Unlock Literacy (UL) in the target districts of Nampula Province in Northern Mozambique.

Educating Children Together (ECT) has been in operation since 2013 and included three phases. The first, ECT1, provided school meals to students in Nacaroa and Muecate, and in phase 2, ECT2, a literacy program was added to the school meals. WV partnered with Save the Children to implement a program called Literacy Boost (LB), which was operational between 2016 and 2020. ECT3 will run from October 2019 to September 2024. LB had some similar components to UL, but World Vision has since taken over the provision of the program and updated and improved components of LB. UL began in the summer of 2021 after schools reopened and it is this third phase of ECT, ECT3, that we discuss in this report.

World Vision delivers a comprehensive school meals program to all schools in the two districts in Nampula province, Muecate and Nacaroa. The school meals consist mainly of corn soy blend plus (CSB+) porridge with micronutrients added and are provided to all teachers and students at the school, every school day. Farmer groups in the community also assist by providing other foods such as fruits and vegetables. World Vision will also implement UL, which entails two components: 1) teacher training in early grade literacy and 2) community action. The teacher training comprises training of all teachers in the schools on good pedagogical techniques for early-grade reading. Teachers are taught the five phases of reading: letter knowledge, sounding out words, reading fluency, vocabulary, and comprehension. Teachers learn to create a print-rich environment in their classrooms and ensure that children remain motivated while learning to read. They are also provided with materials including books and classroom aids. These materials are in the local language, using locally relevant exercises, and are targeted at the appropriate grade level. The community action element entails community meetings and training sessions for parents in supporting literacy, and reading camps in which a teenage literacy volunteer in the community meets with children weekly outside of school (usually on the weekends) to create fun reading exercises and motivate children to read. The camp leaders are provided with training as well as learning materials to use with the students. In addition to the World Vision UL interventions, there will be additional community action component added on that which entails support for reading camp leaders. One teacher in each school will be appointed to serve as a focal point between the reading camp and the school, coordinating with school staff and the

school council. The teacher will help the reading camp leader to recruit students and ensure that they stay in and regularly attend the reading camp.<sup>1</sup>

The Phase 3 (ECT3) impact evaluation will be led by IFPRI. World Vision has been implementing the program in two previous phases (ECT1 & ECT2) since 2013 delivering a comprehensive school meals program to all schools in two districts in Nampula province (Muecate, Nacaroa). As part of the current phase three (ECT3) of the program, World Vision is also implementing additional programming in all the target schools, namely, Unlock Literacy. IFPRI will evaluate ECT3 by conducting a randomized controlled trial (RCT) designed to analyze the effects of Unlock Literacy and an enhanced community-level intervention implemented in conjunction with school meals. The enhanced community-level intervention will involve reading camps that will be supported by an appointed teacher in the school who will help the reading camp leader establish and maintain the reading camps, and be a liaison between the camps and the school management committee. The sample includes 160 rural public primary schools in two districts (Nacaroa and Muecate) in Nampula province. Randomization will be conducted at the level of the school. Fifty schools will be assigned to each of two treatment arms; in one treatment arm, schools will receive meals and Unlock Literacy teacher programming, and in the second treatment arm, schools will receive meals, UL teacher and community programming, and teacher support to reading camps.

This report documents the main results of the impact evaluation baseline data collection. Data was collected by this research team and comprised school-based surveys conducted by ELIM Serviços, in partnership with IFPRI. At this baseline, data was collected from grade 4 students as well as school staff. The target cohort for this evaluation was previously grade 3 students. However, given that baseline data was collected between June and August 2021, following a nearly year-long interruption in schooling, the baseline survey was conducted with grade four students. Grade four students graduated from grade two but received only minimal instruction in grade 3 in the 2020 school year due to COVID-related closures. Like students entering grade 3, they only had 2 years of literacy instruction and accordingly, their literacy level should proxy for grade 3 literacy in future cohorts. A roster of grade 4 students was created to list all students enrolled in the class (whether present that day or not), and 10 students were randomly chosen to be administered the Early Grade Reading Assessment (EGRA) in Portuguese,<sup>2</sup> as well as a short interview to collect data on demographic outcomes and nutrition-related information. We also interviewed the grade 4 teacher, the deputy school director (equivalent to the school principal), and the school cook, who prepares the school meals.

An IFPRI impact evaluation endline survey will be conducted after two years of implementation, in 2023. That survey data will primarily entail a survey of the new cohort of students enrolled in grade

---

<sup>1</sup> Alongside these interventions, a partner organization called Centro de Aprendizagem e Capacitação da Sociedade Civil (CESC) provides community mobilization around education and other topics. They encourage participation in the local and national government processes and learn about their rights and responsibilities as citizens of a democracy.

<sup>2</sup> All schools under WV ECT3 follow the monolingual program in Portuguese. In this program, Portuguese is the language of learning and teaching right from grade 1 onwards. It is also taught as a subject from grade 1. So, they are introduced to early grade literacy (reading and writing) in Portuguese from grade 1. This is different from the language policy in schools or classes that follow the bilingual program in Portuguese and in a local language. In these cases, a local language (the children's mother tongue) is used as the language of teaching and learning from grade 1 to 3 and then Portuguese takes over at grade 4.

3, though data on dropout among the original 2020 grade 4 cohort will also be collected. A final performance evaluation of the entire project, including literacy, health, and nutrition activities, is also expected to be done in 2024 before the award end date of September 30, 2024.

The impact evaluation baseline sample included 1,522 grade 4 students. In addition to students, 160 teachers, 160 headteachers, and 157 school cooks were interviewed at baseline (see Table 1). Our student sample is comprised of students who, on average, are 12 years old, and half of them are girls. These students have low educational and nutritional levels. We observed in our baseline survey that students' EGRA performance was generally poor; average scores were less than 50% overall. This might reflect the detrimental effects of COVID-19; for instance, 57% of the students reported doing nothing to continue learning during the lockdown period. Learning loss during school closures at least partly explain the low learning levels observed. Despite the re-opening of schools, the government introduced a shift system to help reduce crowding in classrooms so an average student attended roughly 3.6 of 5 days in a week at baseline. Moreover, teachers reported an overall decrease in class size. This reduction in school enrollment might be associated with children needing to work to support their families during the pandemic; 31.4% of students have worked in the last year to earn money. Furthermore, students' nutrition rates are low: 41.2% percent of students reported being hungry during the school day, and 63% of students said that their family had to forgo some meals due to COVID-19.

**Table 1. Number of Observations**

	Female	Male	Total
Grade 4 Student	801	719	1520
Grade Teacher	120	40	160
School Director	21	139	160
School Cook	83	76	159
Total	1025	974	1999

Regarding the teachers' characteristics, 75% of teachers in the sample are female and have completed high school education and, on average, have eight years of teaching experience. Although the official language of teaching and learning in the schools surveyed in Portuguese, very few teachers use this language as their primary language of instruction (3.8%). This is unsurprising as few grade 4 students report speaking Portuguese at home (3%), and teachers may resort to instructing in the local language as a coping strategy. Despite this use of the local language for instruction, almost all teachers reported that reading materials are in Portuguese. This presents challenges since instructional materials need to be translated from Portuguese to the primary language of instruction for students to understand (in this context, Emakhuwa). In our sample, 73% of the teachers had access to the grade 4 reading curriculum, and 70% received reading materials. Almost a quarter of teachers reported receiving these materials from World Vision or USDA.

Most school directors in our sample are male (87%) and have completed high school (90%). Most directors have teaching experience, with 95% reporting still teaching at their school and having 14.2

years of average teaching experience. Directors were also surveyed on the status of school facilities. All schools reported having latrines, and 94% of latrines were gender specific. However, only 3.1% of directors reported that their school had access to electricity. All directors mentioned that the school had a school council, and parents participated in almost all of them. Regarding the availability of resources, most directors (95%) reported that the school has reading materials that grade 4 students can use. Indeed, 97% of the directors reported that the school has participated in the Literacy Boost program in the past.

We compare the means of the baseline characteristics and test for any statistically significant differences across treatment arms. Since respondents were randomly assigned to treatment arms, there should be no significant differences in means between treatment arms. However, when testing differences between a large number of variables, some will be statistically significant by chance. We thus also estimate the average difference across all variables between each of the three groups as well as the difference in variables across all groups and variables. Sampling error can result in differences in means for some variables by chance in samples of practical size. While there are some differences in variables across treatments, the joint covariate tests indicate good balance overall and that imbalance should not be of great concern for the analysis for the impact evaluation. Moreover, we plan to include such covariates in our models when analyzing the effects of the intervention to ensure that results are consistent.

Finally, we analyze the determinants of students' reading fluency using the baseline data. In particular, by estimating hierarchical linear models, we identify the predictors of students' performance related to students' background, teacher characteristics, deputy school director profile and school management. The results show that a student's reading fluency is statistically significantly correlated with his/her background variables, such as having books at home; however, variables related to teacher characteristics and the deputy school director profile are not statistically significant. One hypothesis could be that the factors related to educational disruption around the school closures due to the pandemic dominate here, engulfing any effect of an individual teacher or deputy school director.

The next section will discuss the design of the study followed by the description of the analysis strategy that will be used. The results will follow in section four and we will conclude with a discussion and recommendations in section five.

## **2 Evaluation Design**

The impact evaluation of ECT3 uses a combination of quantitative and qualitative methods. Analysis of the baseline data is what we present in this report. Specifically, we will outline the methods used for the study below and then the methodology for data analysis in this report. The analysis will be split into three parts: summary statistics of the indicators of interest across the districts of Nacarua and Muecate, tests for balance across the experimental arms developed for the study, and an analysis of predictors of student reading ability.

## 2.1 Randomized controlled trial: quantitative methods

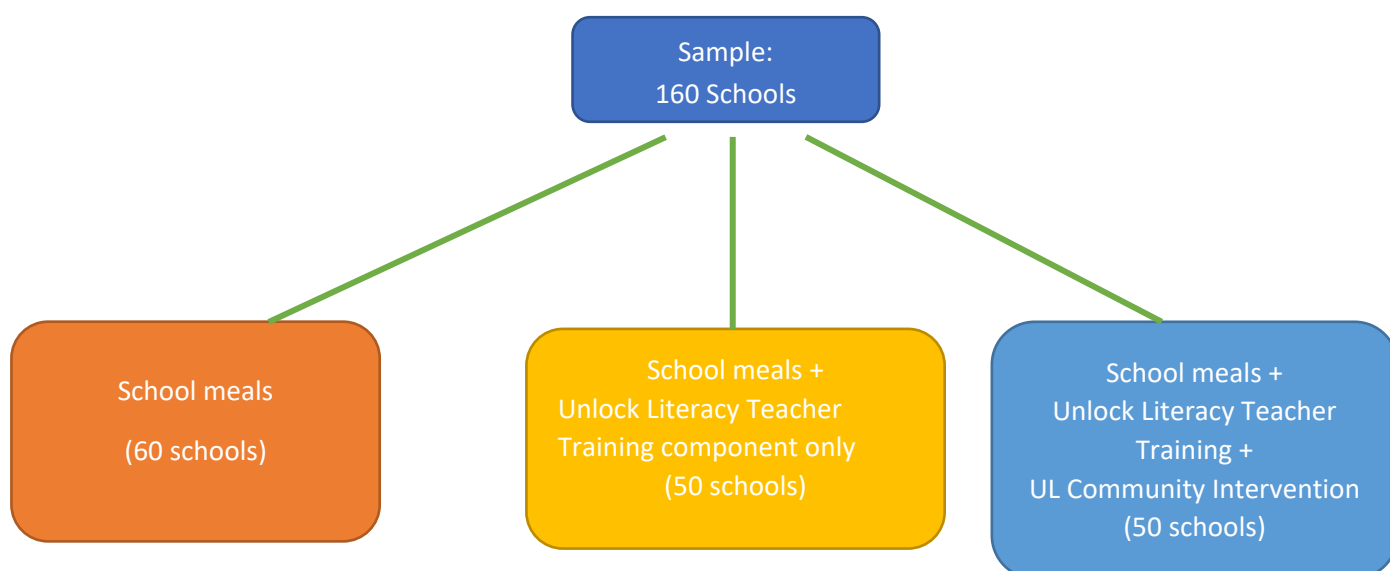
We begin with describing the overall approach to the evaluation. This impact evaluation is a randomized control trial designed to analyze the effects of Unlock Literacy and community action, implemented in conjunction with school meals, in a sample of 160 rural, public, primary schools in two districts (Nacaroa and Muecate) in Nampula province. Schools were randomly assigned to three different groups. In one treatment arm, 50 schools will receive school meals and Unlock Literacy teacher programming. In a second treatment arm, 50 schools will receive school meals and full UL programming (including teacher training and enhanced community reading camps). The remaining 60 schools will be assigned to the control arm and will receive school meals only.

The randomization of schools was conducted using administrative data provided by World Vision. For each school, data was available on the number of students enrolled in the school and the number of teachers in the school. We used these variables to create four groups of schools (strata):

1. Above median school enrollment and above median number of teachers
2. Below median school enrollment and above median number of teachers
3. Above median school enrollment and below median number of teachers
4. Below median school enrollment and below median number of teachers

The randomization was conducted so that each of these four groups had approximately the same number of schools in each experimental arm. This process was conducted to improve balance in school characteristics between the groups; to ensure that some similar types of schools were not grouped together by chance but that each type of school had an equal probability of being randomized into a particular experimental arm. Below, we show the study design graphically.

**Figure 1. Study Design**



The research questions of interest can be described as follows:

1. What are the baseline levels of learning and nutrition outcomes?
2. What is the effect of literacy training on students' literacy and other academic outcomes?

3. Does community action further improve children's outcomes?
4. Do the effects of the program differ by baseline characteristics, including child gender, household socioeconomic status, and a baseline measures of school quality (student academic outcomes and school facilities)?
5. What is the correlation between the home environment as it pertains to literacy (parental engagement, book ownership, parental support) and child literacy outcomes?

Given this design we will be able to measure the impact of the combination of Unlock Literacy and school meals compared to school meals only, and the impact of the combination of UL and school meals compared to both school meals only and school meals, the current UL program, and an enhanced community program. Note that we will not be able to estimate impacts of the school meals program only since all groups will receive them.

## **2.2 Baseline survey**

### **Overview**

Baseline data collection for this survey was conducted between June and August 2021 in the 160 target schools in Nacaroa and Muecete districts. The target cohort for this impact evaluation was grade three students, and a repeated cross-section design will be used. However, given that baseline data was collected following a nearly year-long interruption in schooling due to COVID-19, the baseline survey was conducted with grade four students. Grade four students graduated from grade two, but received only minimal instruction in grade three in the 2020 school year due to COVID-related closures; accordingly, their level of literacy should be an appropriate proxy for grade three literacy in future cohorts.

Data collection was carried out by a team of 20 interviewers from the survey firm ELIM Serviços. These enumerators were divided into 4 teams with 4 enumerators and one supervisor each. Each team collected data in 2 schools per day on average.

Enumerators were trained on the questionnaires in a 7-day training event, which included one day in the field in non-study sites to practice interviewing. During the training, all enumerators were trained on how to administer each questionnaire, the logistics of the school visits, and on research and data collection ethics. They were trained in how to carry out the interview, including line-by-line explanation and interpretation of the questionnaire, the flow and skip-patterns, definitions, and explanations of how to handle unusual cases and when to contact the supervisor for assistance. Supervisors participated in the enumerator training but also received additional training related to their supervisory role. This included detailed and special training on how to obtain informed consent for child respondents – including a detailed protocol for obtaining parental approval. Moreover, we emphasized that when interviewing respondents, enumerators should emphasize that the respondent does not need to answer any particular question should they not want to. World Vision training on child protection was also provided.

Data was collected on tablets using CSPro. The target sample included one fourth grade teacher, one deputy school director, and one school cook in each of the 160 schools, as well as ten grade four students. The key survey instruments included the following.

1. Grade 4 roster and attendance (1 page): generated a roster of all fourth-grade students and recorded information about their attendance pre- and post-pandemic and their dropout status.
2. Deputy school director questionnaire (9 pages): gathered data on characteristics of the school, head teacher, teachers, and students; school meals and take-home rations; overall enrollment; teacher attendance; and mental health (depression, anxiety, fear of COVID-19, and trust).
3. Fourth grade teacher questionnaire (6 pages): gathered data on the characteristics of the fourth-grade teacher and his/her class; teaching methods; and mental health (depression, anxiety, fear of COVID-19, and trust).
4. Fourth grade child questionnaire (2 pages): gathered data on characteristics of the child, their family, and interactions with the fourth-grade teacher.
5. EGRA exam (7 pages): a one-on-one early grade literacy exam that was administered by enumerators on tablets to fourth grade students.
6. School cook questionnaire (1 page): gathered data on characteristics of the cook, how often school meals are provided, how often they are provided, and measures to ensure the quality and safety of the food.
7. Qualitative Instrument: (19 pages): gathered data through focus groups on students, teachers, principals and farmers' perceptions of the World Vision programming and the impacts of COVID-19 related to their lives, schools, and communities. Additionally, semi-structured interviews were conducted with World Vision employees and Ministry of Education staff.

The sample of schools in the quantitative baseline data collection includes 77 schools in Nacarua and 83 schools in Muecate. The sample of individuals surveyed is summarized in Table 2 below, for students, teachers, deputy directors and cooks.

**Table 2. Sample Sizes**

	Muecate		Nacarua		Total
	Female	Male	Female	Male	
Grade 4 Student	409	391	392	328	1520
Grade 4 Teacher	58	25	62	15	160
School Director	12	71	9	68	160
School Cook	44	39	39	37	159
Total	523	526	502	448	1999

Note that surveys with school cooks could not be conducted in three schools in which the school cook was not present during the survey period. In addition, in Muecate the target number of grade four students was slightly under the target number of 770 due to the presence of some schools in which the required number of grade four students providing parental consent forms were not present on the survey date. This was largely due to the system of alternating days for students that was put in place upon the re-opening of schools to ensure that adequate physical distancing could be maintained.

Qualitative data collection at baseline included four focus groups with grade four students (two with boys and two with girls). These four focus groups were conducted at two schools, one in each district, and each included six students. The numbers of deputy directors, teachers, farmers, and parents participating in focus groups in each district is also summarized below.

Nacarao:

Quantity of focus groups	Population category	Male	Female
2	Students	6	6
1	Deputy School Directors	4	1
1	Teachers	3	2
1	Farmers	5	4

Muecate:

Quantity of focus groups	Population category	Male	Female
2	Students	6	6
1	Deputy School Directors	2	3
1	Teachers	1	4
1	Farmers	3	2
1	Parents	3	2

In addition, semi-structured interviews were conducted with three World Vision Employees (one Project Coordinator, one Education Specialist and one M&E Manager) as well as three government officials (two education officials and one health official).

We will now provide some additional details on data instruments and data collection procedures.

## Early Grade Reading Assessment (EGRA)

The Early Grade Reading Assessment (EGRA) was initially developed by RTI international to measure children's progress towards reading mastery and has since been updated and revised to fit local contexts. In our study, we use the EGRA to measure differences in the effectiveness of the ECT3 school meals and pedagogical techniques interventions. The EGRA is a 15-minute assessment that measures a student's aptitude in 5 crucial reading subdomains. The EGRA offers multiple instruments based on purpose, but in the version of the EGRA that we use, we test student's oral vocabulary, listening comprehension, letter recognition, familiar word reading, and reading comprehension. The different sections of the instrument are described below.

**Oral vocabulary:** This subtask measures children's oral receptive and production language skills of individual words and phrases related to body parts, common objects, and spatial relationships. The sub-task included two sub-parts: in the first subpart, eight prompts that required children to perform an action (e.g., "show me your knee") were used to determine their level of understanding of basic Portuguese oral vocabulary; in the second subpart, children were requested to follow six instructions given orally (e.g., "place the pencil next to the paper") and perform the required actions. The score is the number of prompts and instructions correctly performed (maximum score of 14).



**Concepts about printed material:** This subtask measures children's emergent literacy skills by asking them to demonstrate their readiness to handle and read print material. The children were asked 10 questions assessing recognition of the front and back covers of a book, the direction in which to read, the title of a story in a book, the page numbers in which a given story is located, etc. The score is the number of correct answers given.

**Letter name identification and reading:** This subtask is used to measure whether children can identify and read letters presented in random order, both lowercase and uppercase. In this subtask, children were presented with a chart of 100 letters and asked to read as many of these letters as they could within one minute. The score is the number of letters correctly identified and read in one minute.

**Familiar words reading:** This subtask assesses the ability of children to decode printed words and read them correctly. The task reflects both the accuracy and fluency of reading, which are fundamental skills for developing the ability to read and comprehend what is read. The children were presented with a card containing 30 words common to their daily life, including their school life, and asked to read as many words as possible in one minute. The score is the number of words correctly read in one minute.

**Listening comprehension:** This subtask is used to measure whether children have basic knowledge of the language in question and whether they can process what they hear in that language. In this subtask, the enumerators read aloud a short text comprising 56 words for the children and then asked them four questions to check their comprehension. The children were not given a copy of the text to refer to when answering the questions. The score is the number of correct answers given by the children to the four questions asked.

**Oral reading fluency:** This subtask assesses the speed, accuracy, and expressivity at which children read texts. The task reflects the ability to translate letters into sounds, recognize familiar words, decode unfamiliar words, and make sense of the text's meaning. The children were given a card with a narrative text of 120 words and asked to read as many words as possible in one minute. The metric of oral reading fluency was the number of correct words per minute (cwpm) read by the student.

**Reading comprehension:** This subtask assesses the ability of the children to extract and construct meaning out of the texts they read. Studies have shown that oral reading fluency is a predictor of reading comprehension (e.g., Daane et al., 2005; Abadzi, 2011), hence the relationship between EGRA oral reading fluency and reading comprehension subtasks. After reading the narrative text used as the stimulus for the oral reading fluency subtask, the children were asked up to five questions based on how much text they had read. For example, the enumerator only asked the first question to those children who managed to read at least nine words of the text given. The score is the proportion of comprehension questions correctly answered.

As a tool to assess early reading skills, EGRA is often administered to early grades (1-3). However, grade 4 students were tested for the ECT3 impact evaluation baseline. It should be noted that grade 2 and 4 students took the test for ECT2 program baseline study, while grade 3 students took the test for the program midline evaluation. As mentioned before, the decision to administer the test to 4th graders in the program endline evaluation was because these students did not attend grade 3 classes in 2020 due to school closures in the context of COVID-19. For various reasons, for some grades that do not undergo national examinations (e.g. grades 1, 2, 3, 4 and 5), the government automatically promoted all students to the next grade. As a result, the students in grade 4, when the 2021 school year started, had only been exposed to literacy classes for two years (grade 1 and 2). So, in practice, they can be compared to grade 3 students in a normal situation. WV requested that the students tested had had at least two years of exposure to literacy development in the school context.

### Reliability of the EGRA Test

Reliability refers to how dependably or consistently a test measures a characteristic. Reliability in item response theory (IRT) measures the extent to which the measure is independent (free) from groups (samples) as well as from the test items; in other words, if we apply many versions of the test for the same group, they must get the same score and same ranking (Lord, 1980). The reliability of a test is indicated by the *reliability coefficient* and is denoted by " $\alpha$ ". It is expressed as a number ranging between 0 and 1.00, with  $\alpha = 0$  indicating no reliability, and  $\alpha = 1.00$  indicating perfect reliability. In other words, the higher the  $\alpha$  coefficient, the more likely the items measure the same underlying concept. A minimum  $\alpha$  coefficient between 0.65 and 0.8 (or higher in many cases) are considered indicative of a reliable assessment;  $\alpha$  coefficients that are less than 0.5 are usually unacceptable.

Table 3 shows the reliability coefficients for each subtask and the entire EGRA. In all subtasks except reading comprehension, in the intervention group, the reliability is within the range of the recommended values. The  $\alpha$  reliability varies from 0.64 to 0.98, except for reading comprehension, which is below 0.5 (but it is quite close at 0.45). Although the  $\alpha$  coefficient for the entire EGRA test is not high (0.58), in three out of 5 subtests the reliability is very high. Additionally, reading fluency subtest, the strong predictive measure of reading comprehension (Snow, 2002), has the highest  $\alpha$  coefficient (0.93). One can argue that the measurement is independent (free) from groups (samples) as well as from the test items.

**Table 3. Cronbach's alpha reliability coefficient**

Subtask	Intervention
Letter name identification and reading (100 items)	92.5
Familiar words reading (30 items)	88.7
Listening comprehension (4 items)	0.64
Oral reading fluency (161 items)	0.93
Reading comprehension (4 item)	0.45
Full EGRA	0.58

## **2.3 Quality assurance and safety**

Careful quality assurance protocols were used to ensure fidelity to high-quality data collection principles. Supervisors monitored enumerators' work and directly observed interviews as appropriate. Regular data checks were conducted by ELIM and by IFPRI to identify any anomalies in the data. In addition, there was careful adherence to consent procedures to ensure that all households have the opportunity to provide informed consent for the participation of their children. As earlier mentioned, these procedures were specified in submissions to the IFPRI Institutional Review Board as well as the Comité Nacional Bioética em Saúde in Mozambique and approved by both committees.

Consent was sought from the parents of students as follows. Before starting fieldwork, a verification exercise took place whereby the survey firm visited schools to introduce themselves and collect information on school and class timings to organize fieldwork better. At that time, consent forms were dropped off with the school director, who sent the forms home with students. Students then returned the signed forms (if the parent consented) before the school visit. For adults who were surveyed, written consent was sought on the day of the survey and were provided with a consent statement that included: 1) objectives of the study; 2) study procedures; 3) risks and benefits of participating in the study; 4) strategies used by researchers to minimize risks; 5) costs/compensation associated with participating in the study; 6) the duration of the interviews, 7) the voluntary nature of the study and the participant's right to refuse to answer questions or leave the study; 8) that all information would be confidential, that nobody would be able to identify any particular individual's responses, and that their data would be kept securely; and 9) contact information for study staff. The consent statement was also retained by the supervisor, and a copy was provided to the respondent to take home. Their consent was recorded in the CSPro software, whereby a box was checked if the respondent consents to participate. Respondents had the opportunity to ask any questions and raise any concerns, and the enumerators were prepared to address any issues.

## **2.4 Data Management and Security**

The enumerators interviewed the respondents individually. All data was collected on tablets and provided to the project manager and research team each night via a password-protected Dropbox folder. Interviews were not recorded as consent was not sought for recordings, but rather enumerators typed responses into the tablet. The project manager conducted quality checks of the data collected. The data was kept archived with the main IT specialist of the survey firm. Only the project manager and the Director of the company had computer access to these files.

Only authorized individuals have access to the dataset with identifiers, which are secured through a combination of restricted dissemination of information and storage in a password protected file. Original data containing the identities of the respondents is not shared with any other institutions apart from IFPRI. Team members involved with this project are made aware of this provision.

Names and other easily recognizable identifiers were entered with IDs in a separate electronic file from all other data. This electronic file containing names is held separately from all other data files and is kept only by managers of the project at IFPRI. This information is only retained so that it can be used easily to help find respondents for the next round of data collection. Study identifiers (school and individual IDs) are included in each data file so that data from the several instruments collected

within a school may be linked together and with future survey rounds. However, these are not meaningful to casual observers without access to the original study logs. All data files are always maintained under password protection. Public use data will include no identifiers.

Protection of personally identifiable information is only accessible to the project managers and the research assistant engaged in data cleaning and is stored in a secure electronic environment. No PII has been shared with external stakeholders, policymakers, or with World Vision, and these provisions will hold for the duration of the project.

## **2.5 COVID-19 Specific Precautions**

Several procedures were followed to protect both enumerators and respondents during training and enumeration of the surveys. During training, enumerators were mandated to wear masks, they were trained in a large space to allow for physical distancing (6 feet apart), and their temperature was checked every day. If anyone was not feeling well, they were asked to stay home.

During enumeration, there was strong supervision regarding compliance to protocols by supervisors and the field manager. Enumerators were instructed to terminate the interview in case of non-compliance. All respondents were provided with masks and enumerators also wore masks. All enumerators were tested daily with a no-touch thermometer (thermo-gun) to determine whether they had a fever. Those showing signs of illness were not permitted to work and returned home until they were well. All enumerators were guaranteed paid sick leave.

All enumerators used Personal Protective Equipment (PPE). They were provided with masks, that they changed daily, as well as gloves and hand sanitizer. They cleaned their hands with sanitizer before and after each interview. Enumerators were also instructed not to shake hands with anyone and to use an alternative form of greeting instead.

Precautions being made for COVID-19 were explained to respondents, as well as potential risks. Enumerators emphasized that participation is voluntary if there is any discomfort.

All interviews took place outside when possible or in a well-ventilated space. To create the class lists, the students were moved outside the classroom and the enumerator remained 6 feet away from the teacher and students. For the one-on-one interviews, respondents were interviewed outside, so that a distance of 6 feet could be maintained.

Testing students on the EGRA requires that the enumerator and student be in relatively close proximity because the enumerator must be able to see what students are pointing to on the cards. Students were provided with separate cards from which they read letters, words, and texts; the cards were laminated, and the enumerator followed the student reading while entering information on the tablet. In addition, clear plastic screens were set up on a table to separate the enumerator and the student, while enabling them both to see the tablet and cards clearly. Tablets and laminated cards were sanitized between each interview and students also sanitized their hands after the interview.

### 3 Strategy for analysis

As mentioned above, there are three types of analysis presented in this report: summary statistics on key variables for the two districts, balance tests of the same indicators, and determinants of student learning reading development. In this sub-section, we describe the analysis used.

#### 3.1 Summary statistics

We first identified key variables of the study including the indicators to be reported per World Vision ToRs and other variables that we determined were important to the study. We report the overall mean and standard deviation of each variable combining the two districts, and then report the mean and standard deviation of each variable for each district separately. Finally, we test whether there are statistically significant differences in characteristics of the two districts. Statistically significant differences are identified by the following regression:

$$Variable = \alpha + \beta D + \varepsilon$$

D is a dummy variable equal to one if the district is Muecate. The coefficient  $\beta$  represents the mean of the variable of Muecate district. The coefficient alpha represents the mean of district Nacarua (reported as the constant term  $\alpha$  in the regression).  $\varepsilon$  represents the error term and corrects for any correlations within schools and thus the error term is clustered at the level of the school. If  $\beta$  is statistically significant, it indicates that there is a statistically significant difference between the two districts.

#### 3.2 Baseline balance

We use the same variables identified for the district analysis in the test of baseline balance. The balance tests are used to determine how similar the schools are across the three experimental groups. Ideally, the three groups should look approximately equal with few differences between them to ensure that different types of schools are equally represented in the three groups. However, there are also statistical methods that can be used in the final analysis at impact evaluation endline in 2023 that would correct for any imbalances. Here, statistically significant differences are identified using the following regression:

$$Variable = \alpha + \gamma T1 + \theta T2 + \varepsilon$$

T1 is a dummy variable equal to 1 if the school was randomized into the school meals + UL teacher training group and T2 is a dummy variable equal to 1 if the school was randomized into the school meals + UL teacher training + enhanced community action group.  $\alpha$  represents the school meals only group. The coefficients represent the means of each variable for the three groups of schools and if the coefficients are statistically significant, it indicates that that particular group is different from the school meals only group. We also conduct a test for differences between T1 and T2, representing the

difference between the school meals + UL teacher training group and the school meals + teacher training + enhanced community action group (that  $\gamma$  and  $\theta$  are jointly equal to zero).

### 3.3 Determinants of reading

Informed by the conceptual framework of School Effectiveness Research (SER), (Creemers, 2002), a multilevel regression model was used to identify the predictors of students' performance related to the following inputs: *student background, teacher characteristics, deputy school director profile and school management*. The importance of inputs on learner's achievement is well documented in the literature. For instance, using multi-level analysis, Lee, Zuze and Ross (2005) have shown significant relationships in learning of factors such as school resources, teacher quality, student socio economic status (SES), access to books at home, frequency of use of the instructional language at home and attending preschool.

The analysis in this section followed two stages. At the first stage, bivariate analysis was used to explore the relationship between Reading Fluency, one of the key EGRA subtests, and potential predictors related to students' background, teachers' characteristics, the deputy school directors' profile, and school management. The choice of reading fluency as the key variable in this analysis is based on the findings that it is both a strong predictive measure of comprehension and a consequence of it. As Snow and the RAND Reading Study Group (2002:13, as quoted in RTI (2015:25)) put this, "Some aspects of fluent, expressive reading may depend on a thorough understanding of a text. However, some components of fluency — quick and efficient recognition of words and at least some aspects of syntactic parsing [sentence structure processing] — appear to be prerequisites for comprehension." The predictive power of reading fluency on comprehension (see also Daane et al., 2005; Abadzi, 2011) underscores the importance of these abilities as prerequisites for students' academic success. That is, the ability to acquire new knowledge and skills in and outside education contexts depends largely on their ability to read fluently and extract meaning from texts.

On students' background the following variables were considered: gender, age, access to books at home, language used at home, household asset index<sup>3</sup> and number of meals consumed per day. Under teachers' characteristics the variables explored were: age, gender, experience, education level, use of participative methods of teaching, and participation in ZIP (Zona de Influência Pedagógica) meetings. For the deputy school director's profile and management, the following variables were included in the analysis: gender, age, years of experience, level of education, frequency of teachers' class observation and support, participation of school in reading camps, frequency of school council meetings, and an index of school infrastructure (availability of toilets, clean drinking water, and electricity).

In the second stage, variables with statistical significance in the bivariate analysis were included in a Hierarchical Linear Model (HLM) in order to test the strength and direction of prediction since there are different potential factors at the student and school level that can account for variation in student achievement.

HLM is a type of regression analysis for multilevel data, where the dependent variable is at the lowest level. This study deals with data at school and student level, that is, students are nested within schools. Three arguments justify the choice of multilevel analysis instead of an Ordinary Least Square (OLS)

---

<sup>3</sup> The index was estimated using principal component analysis. It includes possession of a radio, mobile phone, and bicycle.

regression of disaggregated data: (i) standard errors and tests based on OLS regression are questionable because the assumption of independent residuals is invalid; (ii) there is a need to disentangle variability at the various levels in this study to compare the variation between and within schools; (iii) with students nested within classes, and classes nested within schools, these grouping effects imply that learners are no longer independent and their responses are correlated, and hence there is a loss of independence of the error term among observations (Raudenbush and Bryk, 2002). This loss of independence in the error term constitutes a serious violation of a key assumption underlying a large body of parametric statistical procedures, but it is properly accounted for using HLM. The equation below represents the model used.

$$Reading_{ij} = \beta_{0ij} + \sum \beta_{1ij}[\text{student-level factors}] + \sum \beta_{2ij}[\text{school-level factors}] + \sum \beta_{3ij}[\text{cross level interaction}]$$

where  $Reading_{ij}$  is the reading score for student  $i$ , nested in group (school)  $j$ ,  $\beta_{1i}$  is the regression coefficient of a variable, and  $\beta_{0ij} = \alpha_{00} + U_{0j} + R_{ij}$  where  $\alpha_{00}$  represents the overall mean of reading fluency,  $U_{0j}$  is a random effect at the group level (school level), and  $R_{ij}$  random effect at the individual level (student level).

### 3.4 Strengths and limitations

This impact evaluation is designed as a randomized controlled trial, and this design has a number of strengths. It allows us to rigorously identify the causal effects of the interventions of interest, given that the randomization process ensures that schools in the control and treatment arms are similar on average in observable and unobservable characteristics. Conducting data collection at baseline and endline will also allow us to verify that schools in the multiple experimental arms are in fact similar on observables, and to adjust our analysis at endline for any differences observed at baseline.

The RCT design also entails some weaknesses. We are not necessarily able to identify the channels for the treatment effects we observe, though the evaluation design entails the collection of additional data designed to identify channels. RCTs are also not necessarily generalizable to other contexts outside of the context in which the evaluation is conducted, though this can be a general challenge in many evaluation designs.

Finally, this impact evaluation will not identify the impact of the school meals themselves since they are being provided in all schools in the sample. The impact evaluation will identify the impacts of the components of Unlock Literacy given that school meals are provided in the school. This again reduces generalizability, but because school meals are already prominent in many contexts and are becoming more so, we do not see this as a substantial hinderance.

## 4 Results

In this section we will present the results of the baseline data collection. We will first present summary statistics that will enable us to gain an understanding of the context in which these schools operate, looking at students, teachers, and deputy school directors in turn. We will also consider differences in characteristics across the two districts. Second, we will present results on the degree of balance between the three experimental groups of schools in the study.

### 5.1 Summary Statistics

#### Grade 4 Students

Table 4 shows demographic characteristics of grade 4 students. The mean age of students in the sample is 11.5 years. The sample is roughly equally split between male and female students, with male students making up 47.3 percent of the sample. With respect to household demographic characteristics, 83.2 percent of students live with their mother and 72.3 percent live with their father. Students reported having an average of 3.3 siblings, and this is slightly higher at 3.5 for students in Nacaróa. The survey also asked questions to gauge household wealth. Roughly a third of students reported that their family owned a radio (32.2 percent). A much higher share of students reported that their household owned a mobile phone (62 percent) and livestock (88 percent). On average, 41 percent of students reported owning a bicycle, but this varied significantly by district; a greater share of students in Muecate (53.2 percent) reported that their household owned a bicycle compared to students in Nacaróa (28.5 percent).

Students were also surveyed on their general academic practices. Roughly 55 percent of students reported reading books at school, and 67 percent reported having books at home. These figures imply that a sizeable proportion of students lack the proper materials to promote learning. In addition, 56.6 percent also reported doing nothing to continue learning during the COVID-19 lockdown period, which implies that many students are at similar or lower levels of learning compared to before the lockdown. A very small proportion of students speak Portuguese at home (3 percent), which could affect performance because this is the language used in most school materials. Roughly half of students in Nacaróa say that their parents help them with homework, and this is substantially higher for students in Muecate (60.6 percent). Roughly a quarter of students also report belonging to a reading camp (27.5 percent). Notably, 56 percent of students reported that they did not do anything during the school closures.

Students in the sample attended an average of 3.6 out of the last 5 days of school. The data also showed that 31.3 percent of students have worked in the last year to support themselves or their families, which could be affecting attendance if work is during school hours. According to deputy school directors, 7.3 percent of students have missed more than 10 days of school in the last year due to illness. These statistics highlight the negative impact that COVID-19 has had in this population. Learning loss due to closures will be difficult to offset with alternating attendance schedules.

Nutrition outcomes among students are poor with 41.2 percent of students reporting being hungry during the school day. They also report consuming an average of only 2.3 meals per day, and an average of 1.1 of those meals are eaten at school per day. Roughly two third of students (62.5 percent) also reported going hungry because of COVID-19. This suggests that many students are not getting enough meals and proper nutrition, which can have detrimental effects on health and academic performance.



**Table 4. Summary statistics for grade 4 students**

	Nacarora (1)		Muecate (2)		Total		Difference
	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2)-(1)
Child's Age	710	11.35 (2.066)	784	11.58 (1.933)	1494	11.47 (2)	.227
Proportion male	720	0.46 (.498)	800	0.49 (.5)	1520	0.47 (.499)	.033
School Attendance Over Last 5 School Days	722	3.61 (1.166)	799	3.62 (1.196)	1521	3.61 (1.181)	.005
Miss More than 10 Days a Year Due to Illness	722	0.08 (.133)	800	0.07 (.08)	1522	0.07 (.109)	-.013
Child Works (I=Yes)	722	0.32 (.467)	799	0.31 (.462)	1521	0.31 (.464)	-.012
Average Number of Meals Per Day	722	2.23 (.755)	799	2.36 (.663)	1521	2.30 (.711)	.13**
Number of Meals eaten at School Per Day	722	1.09 (.467)	799	1.04 (.269)	1521	1.06 (.377)	-.051**
Hungry During the School Day (I=Yes)	722	0.43 (.496)	799	0.40 (.489)	1521	0.41 (.493)	-.037
Has Books at Home to Read (I=Yes)	722	0.69 (.464)	799	0.65 (.476)	1521	0.67 (.47)	-.032
Reads books at school (I=Yes)	722	0.54 (.499)	799	0.57 (.495)	1521	0.56 (.497)	.032
Parents Help with Homework (I=Yes)	722	0.51 (.5)	799	0.61 (.489)	1521	0.56 (.497)	.099***
Belongs to a Reading Camp (I=Yes)	722	0.30 (.459)	799	0.25 (.435)	1521	0.28 (.447)	-.048
Went Hungry Due to COVID (I=Yes)	722	0.61 (.489)	799	0.64 (.48)	1521	0.63 (.484)	.033
Lives with Mother (I=Yes)	722	0.85	799	0.81	1521	0.83	-.044**

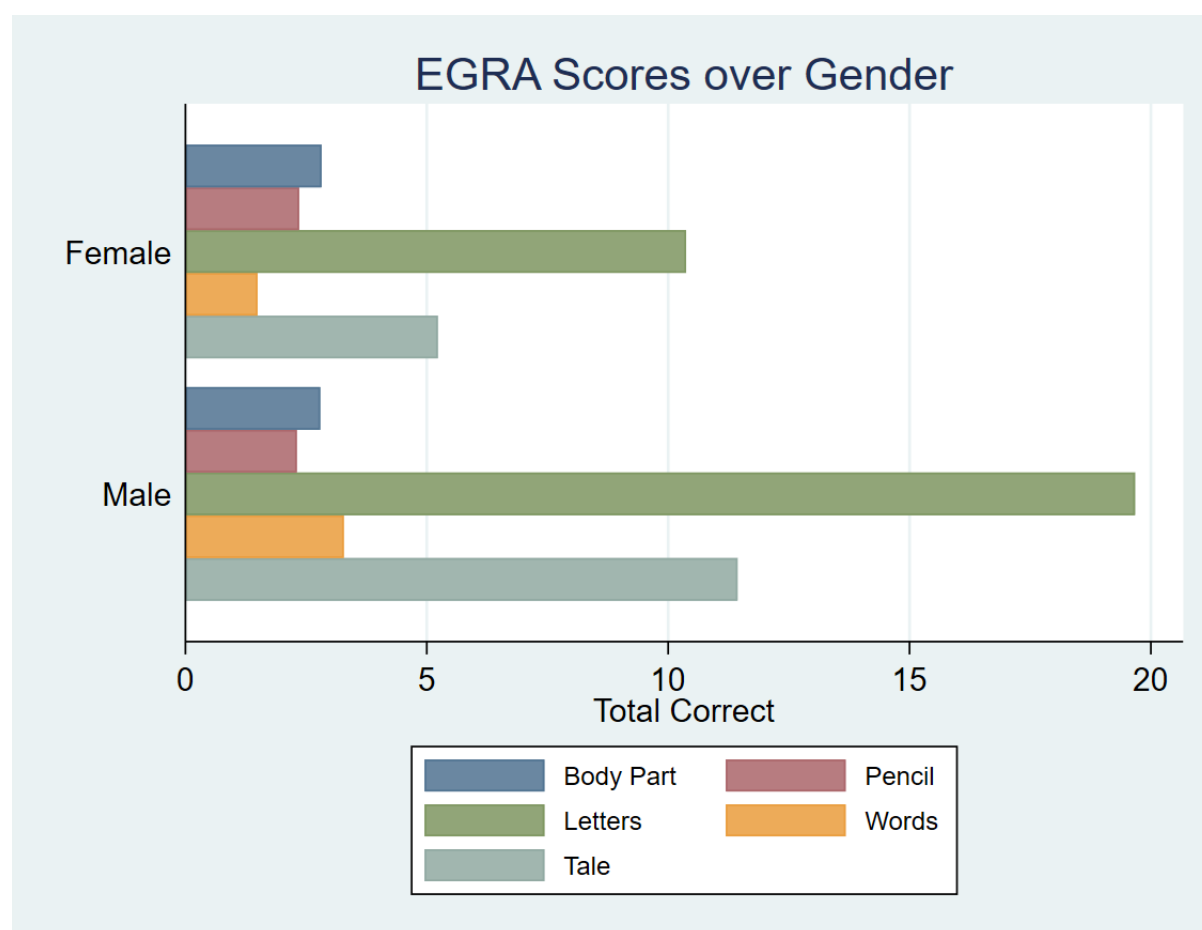
	Nacarua (1)		Muecate (2)		Total		Difference
	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2)-(1)
Lives with Father (I=Yes)	722	0.75 (.353)	799	0.70 (.392)	1521	0.72 (.374)	-.046*
Number of Siblings	722	3.46 (2.193)	799	3.11 (2.119)	1521	3.27 (2.161)	-.352**
Speaks Portuguese at Home (I=Yes)	722	0.03 (.183)	800	0.03 (.156)	1522	0.03 (.169)	-.01
Reads books at school (I=Yes)	722	0.54 (.499)	799	0.57 (.495)	1521	0.56 (.497)	.032
Did Nothing to Continue Learning during COVID	722	0.56 (.496)	800	0.57 (.496)	1522	0.57 (.496)	.004
Family Has Radio (I=Yes)	722	0.32 (.468)	799	0.32 (.467)	1521	0.32 (.467)	-.001
Family Has Livestock (I=Yes)	722	0.88 (.329)	799	0.88 (.325)	1521	0.88 (.327)	.003
Family Has Mobile Phone (I=Yes)	722	0.62 (.486)	799	0.62 (.485)	1521	0.62 (.485)	.003
Family Has Bicycle (I=Yes)	722	0.29 (.452)	799	0.53 (.499)	1521	0.41 (.493)	.247***
Miss More than 10 Days a Year Due to Illness	722	0.08 (.133)	800	0.07 (.08)	1522	0.07 (.109)	-.013

## Grade 4 EGRA Performance

Table 5 contains sample and district means for Grade 4 Student EGRA subtask scores, as well as the proportion of students that scored zero on each EGRA subtask. EGRA performance among students was generally low, and average scores were lower than 50 percent in each subtask. However, there were significant district differences in performance, and this primarily stemmed from higher performance by students in Nacarua. The average score on the Oral Vocabulary subtask was 2.8 points (out of 8), but students in Nacarua (3.4 points) scored significantly higher than students in Muecate (2.3 points). A higher proportion of students in Muecate (31.5 percent) also scored zero in this section compared to students in Nacarua (18.3 percent). Students in Nacarua (2.7 points) also performed better in the Concepts about Printed Material subtask compared to Students in Muecate (2.0 points, out of 6), and a lower proportion scored zero on this subtask and the Letter Identification subtasks. Students in Nacarua also scored higher in the Oral Reading Fluency and Listening Comprehension subtasks.

Overall, male students outperform their female counterparts. While male and female students have roughly equal scores on oral vocabulary and concepts about printed materials, male children score much higher in letter name identification and reading, familiar words reading, listening comprehension, oral reading fluency, and reading comprehension. These sub-tasks are higher-order reading skills, so there is a risk that female students may fall further behind.

**Figure 2. EGRA Scores by Gender**



**Table 5. Summary Statistics of Grade 4 EGRA Scores**

	Nacarora (1)		Muecate (2)		Total		Difference
	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2)-(1)
Correct Answers on Oral Vocabulary (out of 8)	722	3.41 (2.637)	799	2.27 (2.06)	1521	2.81 (2.419)	-1.139***
Percent Scored Zero on Oral Vocabulary	722	0.18 (.387)	799	0.32 (.465)	1521	0.25 (.435)	.133***
Correct Answers on Concepts about Printed Materials (out of 6)	722	2.70 (2.194)	799	2.00 (1.979)	1521	2.33 (2.112)	-0.698
Percent Scored Zero on Concepts about Printed Materials	722	0.23 (.42)	799	0.36 (.481)	1521	0.30 (.458)	.134***
Correct Answers on Letters Identification (out of 100)	722	16.19 (24.369)	799	13.54 (22.144)	1521	14.80 (23.257)	-2.652
Percent Scored Zero on Letter Identification	722	0.25 (.435)	799	0.34 (.473)	1521	0.30 (.457)	.084**
Correct Answers on Familiar Words Reading (out of 30)	722	2.54 (4.914)	799	2.19 (4.827)	1521	2.36 (4.87)	-.355
Percent Scored Zero on Familiar Words Reading	722	0.39 (.488)	799	0.45 (.497)	1521	0.42 (.493)	.058
Correct Answers on Reading Comprehension (out of 4)	722	0.09 (.36)	799	0.09 (.351)	1521	0.09 (.355)	-.009
Percent Scored Zero on Reading Comprehension	722	0.93 (.263)	799	0.93 (.251)	1521	0.93 (.257)	.007
Correct Answers on Oral Reading Fluency (out of 161)	722	9.86 (26.932)	800	6.79 (22.137)	1522	8.25 (24.568)	-3.067*
Percent Scored Zero on Oral Reading Fluency	722	0.00 (0)	800	0.00 (0)	1522	0.00 (0)	0
Correct Answers on Listening Comprehension (out of 4)	722	0.65 (1.03)	799	0.49 (.849)	1521	0.56 -0.94	-.159**
Percent Scored Zero on Listening Comprehension	722	0.62 (.485)	799	0.68 (.466)	1521	0.65 (.476)	.06

## Grade 4 Teachers

Table 6 reports on demographics statistics of teachers and shows that 75 percent of teachers in the sample are female, and their average age is 31.8 years. Roughly three quarters of teachers possess a high school education (78.8 percent) and have been teaching for an average of 7.8 years. In addition, 29.4 percent of teachers report speaking Portuguese at home, a much higher proportion than students, suggesting some degree of mismatch in teacher and student language abilities. In Nacarua, 75.3 percent of teachers report attending ZIP meetings at least half the time, while this is significantly less for teachers in Muecate (49.3 percent).

Teachers also reported on their teaching practices. On average, 68.8 percent of teachers group students according to ability or in other ways, but this varies significantly by district. In Muecate, 62.7 percent of teachers group students, while 75 percent of teachers in Nacarua do this. Very few teachers also use Portuguese as their primary language of instruction (3.8 percent), but majority reported that Portuguese was the secondary language of instruction (80.6 percent). This is unsurprising as few grade 4 students report speaking Portuguese at home (3 percent), and teachers may want to instruct children in the local dialect instead. In addition, 98.8 percent of teachers reported that reading materials were in Portuguese, which may present issues because the class materials need to be translated orally from Portuguese to the primary language of instruction for students to understand.

Teachers also report on the availability of classroom resources. The data shows that 72.5 percent of teachers have access to the grade 4 reading curriculum. In addition, 69.4 percent received reading materials, and 23.8 percent report that receiving these materials from World Vision or USDA. This roughly matches up with student reporting on availability of reading materials as only 55.7 percent of grade 4 students reported having reading materials at school. Qualitative data also corroborates these accounts, with teacher focus groups describing measures taken to deal with insufficient materials. For example, some teachers reported writing relevant lessons on the board so that students without books could copy them, or having students share textbooks. Although these measures are creative, they were not without their challenges, and both students and teachers expressed their frustrations during the qualitative study.

Teachers were asked questions on their experience with COVID-19 and how it affected different areas of their lives. Very few teachers (0.6 percent) reported losing a family member as a result of COVID-19, while 6.9 percent reported losing income. Some teachers (12.5 percent) also reported that it is harder to manage children as a result of the pandemic. When asked if their schools could practice COVID-19 prevention methods, 57.1 percent of teachers in Nacarua agreed, while teachers in Muecate expressed marginally more confidence (69.9 percent). None of the COVID-19 shock measures were statistically significantly different across districts, suggesting that both were equally negatively affected.

**Table 6. Summary Statistics of Teachers**

	Nacarora (1)		Muecate (2)		Total		Difference
	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2)-(1)
Teacher is male	77	0.19 (.399)	83	0.30 (.462)	160	0.25 (.434)	.106
Teacher's Age	77	32.08 (7.462)	83	31.58 (6.867)	160	31.82 (7.141)	-.5
Teacher has High School Education	77	0.82 (.388)	83	0.76 (.43)	160	0.79 (.41)	-.059
Years as a Teacher	77	8.21 (6.214)	83	7.43 (6.504)	160	7.81 -6.36	-.774
Teacher Groups Students by Ability	77	0.10 (.307)	83	0.06 (.239)	160	0.08 (.274)	-.044
Teacher Groups Students in any way	77	0.75 (.434)	83	0.63 (.487)	160	0.69 (.465)	-.127*
Teacher has Access to Grade 4 Reading Curriculum (I=Yes)	77	0.78 (.417)	83	0.67 (.471)	160	0.73 (.448)	-.105
Teacher Received Reading Materials	77	0.71 (.455)	83	0.67 (.471)	160	0.69 (.462)	-.04
Teacher Received Reading Materials from WV/USDA	77	0.27 (.448)	83	0.20 (.406)	160	0.24 (.427)	-.068
Teacher Participated in Literacy Boost (I=Yes)	77	0.66 (.476)	83	0.60 (.492)	160	0.63 (.484)	-.06
Number of Grade 4 Male Students in Class	77	36.32 (34.049)	83	34.69 (33.787)	160	35.48 (33.817)	-1.638
Number of Grade 4 Females Students in Class	77	30.79 (14.958)	83	29.23 (29.348)	160	29.98 (23.489)	-1.563
School Has Tools to Practice COVID Protection (I=Yes)	77	0.57 (.498)	83	0.70 (.462)	160	0.64 (.482)	.127*
Speaks Portuguese at Home (I=Yes)	77	0.30	83	0.29	160	0.29	-.01

	Nacaroa (1)		Muecate (2)		Total		Difference
	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2)-(1)
Portuguese is Primary Language in Class (1=Yes)	77	(.461) 0.03	83	(.456) 0.05	160	(.457) 0.04	.022
Portuguese is Secondary Language in Class (1=Yes)	77	(.16) 0.86	83	(.215) 0.76	160	(.191) 0.81	-.098
Attends ZIP Meetings Half the time or More	77	(.352) 0.75	83	(.43) 0.49	160	(.396) 0.62	-.259***
Males in 2nd Grade before COVID	76	(.434) 40.32	83	(.503) 37.37	159	(.487) 38.78	-2.942
Females in 2nd Grade before COVID	76	(39.907) 35.45	83	(32.406) 32.96	159	(36.099) 34.15	-2.484
COVID Effects: Lost Family Member	77	(19.27) 0.00	83	(28.139) 0.01	160	(24.264) 0.01	.012
COVID Effects: Lost Income	77	(0) 0.06	83	(.11) 0.07	160	(.079) 0.07	.007
COVID Effects: Hard to Manage Children	77	(.248) 0.10	83	(.261) 0.14	160	(.254) 0.13	.041
COVID Children Activities: Studying	77	(.307) 0.66	83	(.354) 0.67	160	(.332) 0.67	.012
Reading Materials in Portuguese	77	(.476) 0.99	83	(.471) 0.99	160	(.472) 0.99	.001
Reading Materials in Emakhuwa	77	(.114) 0.01	83	(.11) 0.02	160	(.111) 0.02	.011
Grade 4 Teacher Attended 80% or more of Previous 5 days	77	(.114) 0.71	83	(.154) 0.80	160	(.136) 0.76	.081
		(.455)		(.406)		(.431)	

## Deputy School Directors

Table 7 shows that the majority of directors are male (86.9 percent) and have a high school education (89.4 percent). The mean age for directors in Muecate is on 36.3 years, while mean age is 40.5 years for directors in Naracoa. Most directors are experienced and active in teaching, with 95 percent reporting teaching at their school and having 14.2 years of teaching experience on average. However, average teaching experience varied by district, with directors in Nacaroa (16.2 years) having almost 4 additional years of experience than directors in Muecate (12.4 years). Participation in directorial duties was also high with all directors providing feedback to teachers on their performance, and 96.3 percent observing classes.

Directors were also surveyed on the status of school facilities. All schools had latrines, and 93.8 percent of latrines were gender segregated. Only 3.1 percent of directors reported that their school had access to electricity. All directors reported that the school had a school council, and 99.4 percent report that parents are part of the school council. In terms of availability of resources, most directors (94.4 percent) reported that the school has reading materials that grade 4 students can use, and 96.9 percent report that the school has participated in the Literacy Boost program in the past.

Directors also report on teacher and student enrolment and attendance. In Muecate, there were 242.8 male students and 212.6 female students enrolled, while the numbers were significantly higher in Nacaroa at 310.3 male students and 276.4 female students. Statistics also showed that among grade 4 students, there was an average of 40.7 male students per school. Among female grade 4 students, average enrolment varied from 42.1 in Nacaroa to 32.4 in Muecate. Directors also reported their schools having an average of 8 permanent and temporary teachers.



**Table 7. Summary Statistics of Deputy School Directors**

	Nacarua (1)		Muecate (2)		Total		Difference
	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2)-(1)
Director's Age	77	40.52 (7.579)	83	36.35 (6.691)	160	38.36 (7.41)	-4.17***
Director is male	77	0.88 (.323)	83	0.86 (.354)	160	0.87 (.339)	-.028
Director has High School Education	77	0.92 (.27)	83	0.87 (.341)	160	0.89 (.309)	-.055
Director Teaches at School (I=Yes)	77	0.97 (.16)	83	0.93 (.261)	160	0.95 (.219)	-.046
Years Working Full-Time at Current School	77	3.36 (2.345)	83	4.13 (2.704)	160	3.76 (2.559)	.769*
Years as a Teacher	77	16.25 (7.408)	83	12.39 (5.541)	160	14.24 (6.768)	-3.861***
Director observes classes (I=Yes)	77	0.97 (.16)	83	0.95 (.215)	160	0.96 (.191)	-.022
Director Provides Feedback (I=Yes)	75	1.00 (0)	79	1.00 (0)	154	1.00 (0)	0
School has Latrine (I=Yes)	77	1.00 (0)	83	1.00 (0)	160	1.00 (0)	0
Latrines are Gender Separated (I=Yes)	77	0.99 (.114)	83	0.89 (.313)	160	0.94 (.243)	-.095**
School Has Electricity (I=Yes)	77	0.05 (.223)	83	0.01 (.11)	160	0.03 (.175)	-.04
School Operates in Shifts (I=Yes)	77	0.94 (.248)	83	0.95 (.215)	160	0.94 (.231)	.017
School Has a School Council (I=Yes)	77	1.00 (0)	83	1.00 (0)	160	1.00 (0)	0
Parents are Part of School Council (I=Yes)	77	0.99	83	1.00	160	0.99	.013

	Nacarora (1)		Muecate (2)		Total		Difference
	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2)-(1)
School has been visited by health worker (1=Yes)	77	0.87 (.114)	83	0.89 (0)	160	0.88 (.079)	.021
School Collects Fees (1=Yes)	77	0.04 (.338)	83	0.02 (.313)	160	0.03 (.325)	-.015
School Received Reading Materials	77	0.95 (.195)	83	0.94 (.154)	160	0.94 (.175)	-.008
School Participated in Literacy Boost (1=Yes)	77	0.97 (.223)	83	0.96 (.239)	160	0.97 (.231)	-.01
Number of Male Students in School	77	310.39 (.16)	83	242.77 (.188)	160	275.31 (.175)	-67.619**
Number of Female Students in School	77	276.42 (220.529)	83	212.59 (174.468)	160	243.31 (200.232)	-63.825**
Number of Male Students in Grade 4	77	44.81 (204.596)	83	37.31 (155.461)	160	40.92 (183.018)	-7.492
Number of Female Students in Grade 4	77	42.06 (32.647)	83	32.42 (28.456)	160	37.06 (30.678)	-9.643
Number of Teachers in School	77	8.83 (34.431)	83	7.31 (25.765)	160	8.04 (30.535)	-1.518
		(8.435)		(5.268)		(6.993)	

## 5.2 Summary of Performance Indicators

Table 8 below displays the baseline values of the performance indicators to be tracked by World Vision's program monitoring and evaluation system. Note that only school-level indicators were measured, not community or parent indicators. Additionally, Grade 4 is taken to be representative of the school for indicators such as rates of attendance or teaching practices. Note that not all baseline values are zero since ECT1 and ECT2 programming were previously taking place.

**Table 8. Performance Indicators**

Result	Indicator		Baseline
<b>1-Improved Literacy of School-Aged Children</b>	Percentage of students who by the end of two grades of primary schooling demonstrate that they can read and understand the meaning of grade level text	Total	13.2 (1523)
		Male	18.5 (703)
		Female	8.1 (793)
<b>2-Improved quality of literacy instruction</b>	Percentage of teachers in target schools who demonstrate improved literacy instruction as identified by supervisors, mentors or coaches (Defined as: Percentage of teachers who split students into groups (by ability or to work together with a textbook or other learning materials)).		68.8 (160)
<b>3-More Consistent Teacher Attendance</b>	Percentage of <b>Grade 4 teachers</b> who attend and teach school at least 80% of scheduled days per school year		75.6 (160)
<b>4-Better Access to School Supplies and Materials</b>	Percentage of teachers who received textbooks and other teaching and learning materials provided as a result of USDA assistance		46.8 (111)
	Percentage of teachers who received teaching and learning material from any source		69.4 (160)
	Percentage of classrooms with literacy instruction materials (textbooks, workbooks) sufficient for effective instruction		82.2 (151)
<b>5-Improved Attentiveness of Students</b>	Percentage of students who feel more attentive because of the food they get at school		73.6 (1,523)
<b>6-Reduced Short-Term Hunger</b>	Percentage of students in target schools who indicate that they are not hungry during school day		58.8 (1523)
<b>7-Increased Access to Food (School Feeding)</b>	Number of school aged children receiving daily school meals (breakfast, snack, lunch) as a result of USDA assistance	Total	74,734
		Male	39,850

Result	Indicator		Baseline
		Female	34,884
<b>8-Improved Student Attendance</b>	Percentage of students that attend school at least 80% of the school days	Total	43.8 (1,519)
	Percentage of students that attend school at least 80% of the school days	Male	44.7 (718)
	Percentage of students that attend school at least 80% of the school days	Female	43.2 (801)
	Average student attendance rate in USDA supported classrooms/schools	Total	53.9 (5,752)
	Average student attendance rate in USDA supported classrooms/schools	Male	51.6 (2,934)
	Average student attendance rate in USDA supported classrooms/schools	Female	56.3 (2,818)
<b>9-Reduced Health Related Absences</b>	Percent of students in target schools who miss more than 10 school days/year due to illness	Total	29.4 (1,523)
		Male	24.6 (703)
		Female	33 (793)
<b>10-Improved Knowledge of Health and Hygiene Practices</b>	Percentage of children in target communities who can identify at least 3 important health/hygiene practices		31.6 (1,523)
	Percent of students targeted by the hygiene promotion program who know at least three (3) of the five (5) critical times to wash hands	Total	31.5 (1519)
	Percent of students targeted by the hygiene promotion program who know at least three (3) of the five (5) critical times to wash hands	Male	28.6 (718)
	Percent of students targeted by the hygiene promotion program who know at least three (3) of the five (5) critical times to wash hands	Female	34.2 (801)
<b>11-Increased Use of Health and Dietary Practices</b>	Percent of target beneficiaries who use appropriate hand washing practices	Total	89.3 (1523)
		Male	88.7 (703)
		Female	90.3 (793)

Result	Indicator		Baseline
	Percentage of school-aged children receiving a minimum acceptable diet (measured by eating at least 3 meals per day).	Total	39.4 (1521)
		Male	38.0 (718)
		Female	40.6 (801)
<b><i>12-Increased Knowledge of Safe Food Preparation and Storage Practices</i></b>	Percentage of food preparers at target schools that can identify at least 3 key practices aimed at safe food preparation	Total	99.4 (160)
<b><i>13-Increased Access to Clean Water and Sanitation Services</i></b>	Number of schools using an improved water source	Total	129 (159)
	Number of schools with improved sanitary facilities	Total	159 (159)

### 4.3 Baseline Balance

We compare means of baseline characteristics and test to see if there are any significant differences in means across treatment arms. Since respondents were randomly assigned to treatment arms, there should be no significant difference in means between treatment arms. However, in samples of practical size, sampling error can result in differences in means for some variables by chance.

Tables 9-11 show the balance tests for the Grade 4 Students, Grade 4 Teachers, and Deputy Director characteristics. The first three columns of each table show the means and standard deviations of characteristics in the control (School Meals), treatment 1 (School Meals + Teacher Training), and treatment 2 (School Meals + Teacher Training + Community Participation) groups respectively. Columns 4-5 show mean of each variable across all groups. Columns 6-8 displays the p-value from a test of whether the mean of the variable is statistically different between the control and treatment 1 (column 6), between treatment 1 and treatment 2 (column 7), and between treatments 1 and 2 (column 8). The final column displays the p-value from a test for statistically significant differences across all three groups.

In the final row of the tables, additional tests are reported. The first is a p-value for a test of whether all the characteristics *within* the control group are jointly statistically significant. The next two report the same statistic for treatment groups 1 and 3, respectively. Lastly, the last column in the last row reports the p-value for statistically significant differences across all three groups **and** all the variables. It is this last statistic that is of most interest.

## Grade 4 Students

Demographic statistics show that grade 4 students in treatment 1 were older than children in treatment 2 and control groups. Student gender was also imbalanced, as there were more male students in treatment 1 than in the control group.

There was also some imbalance in some EGRA test scores across the three groups, although these differences were only marginally significant. For example, students in treatment 1 outperformed children in treatment 2 and control groups in the Familiar Words Reading and Oral Reading Fluency subtasks. A higher proportion of students in the control group scored zero on the Familiar Words Reading subtask compared to students in the other treatment arms. In the Reading Comprehension subtask, students in treatment 1 also scored higher and fewer of them scored zero compared to students in treatment 2. A lower proportion of students in treatment 1 also scored zero on the Listening Comprehension subtask compared to students in the control group.<sup>4</sup>

Tests of differences for other baseline characteristics show that the average number of meals and school attendance over the last week are balanced across treatment arms. However, there are significant differences for other characteristics. For example, a higher percentage of students in treatment 1 have worked in the past 12 months compared to students in treatment 2 and control groups, although this difference is again only marginally significant. A higher proportion of students in the control group have also missed more than 10 school days in the year compared to students in treatment 1. Additionally, although less students in treatment 1 have books to read at home than in the students in the control group, more students in treatment 1 read books at school than students in treatment 2.

We also estimate the combined effect of all baseline characteristics on the three treatment arms. The results show that two out of three individual joint covariate tests (all covariates within an experimental arm) are highly significant, indicating that some or all covariates have a non-zero effect on treatment status. The overall joint covariate test, which aggregates the results of all our balance tests across all treatment arms, is also significant at the 5 percent level.

Overall, many baseline characteristics were unbalanced across experimental arms. These covariates are important, and we will ensure to control for these differences by adding these baseline variables in the regression during the impact evaluation endline analysis in 2023. By doing so, we will be able to ensure that our results are accurate.

---

<sup>4</sup> The variable “Percent Scored Zero on Oral Reading Fluency” was dropped from the balance analysis because no students scored zero in this subtask.

**Table 9. Baseline balance, student characteristics**

	School Meals		+ Teacher Training		+ Community Action		Total		(4)	(5)	(6)	
	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2) - (1)	(3) - (1)	(2) - (3)	(4) = (5) = (6)
Child's Age	513	11.242 (1.851)	516	11.783 (2.237)	465	11.37 (1.833)	1494	11.469 (2)	.541**	.128	.413*	.081*
Child's Biological Sex (1=Male)	516	.434 (.496)	527	.507 (.5)	477	.478 (.5)	1520	.473 (.499)	.073**	.044	.029	.131
Correct Answers on Oral Vocabulary (out of 8)	515	2.707 (2.424)	529	2.968 (2.453)	477	2.757 (2.37)	1521	2.813 (2.419)	.261	.05	.211	.615
Percent Scored Zero on Oral Vocabulary	515	.28 (.449)	529	.227 (.419)	477	.252 (.434)	1521	.252 (.435)	-.053	-.028	-.025	0.559
Correct Answers on Concepts about Printed Materials (out of 6)	515	2.291 (2.074)	529	2.359 (2.152)	477	2.352 (2.112)	1521	2.334 (2.112)	.068	.061	.007	.927
Percent Scored Zero on Concepts about Printed Materials	515	.299 (.458)	529	.308 (.462)	477	.289 (.454)	1521	.299 (.458)	.009	-.01	.019	.925
Correct Answers on Letters Identification (out of 100)	515	12.971 (21.418)	529	16.346 (23.751)	477	15.05 (24.49)	1521	14.797 (23.257)	3.375	2.079	1.296	.299
Percent Scored Zero on Letter Identification	515	.324 (.469)	529	.268 (.444)	477	.302 (.46)	1521	.298 (.457)	-.056	-.022	-.033	.412
Correct Answers on Familiar Words Reading (out of 30)	515	2.052 (4.495)	529	2.849 (5.685)	477	2.138 (4.202)	1521	2.356 (4.87)	.796*	.086	.71*	.145
Percent Scored Zero on Familiar Words Reading	515	.474 (.5)	529	.391 (.489)	477	.388 (.488)	1521	.418 (.493)	-.082*	-.086*	.003	.141
Correct Answers on Reading Comprehension (out of 4)	515	.083	529	.115	477	.067	1521	.089	.032	-.016	.048*	.197

	School Meals		+ Teacher Training		+ Community Action		Total		(4)	(5)	(6)	
	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2) - (1)	(3) - (1)	(2) - (3)	(4) = (5) = (6)
Percent Scored Zero on Reading Comprehension	515	.932 (.334) (.252)	529	.911 (.408) (.285)	477	0.945 (.31) (.227)	1521	.929 (.355) (.257)	-.021	.013	-.034*	.203
Correct Answers on Oral Reading Fluency (out of 161)	516	6.797 (21.513)	529	10.826 (28.797)	477	6.95 (22.288)	1522	8.245 (24.568)	4.03*	.153	3.876*	.111
Correct Answers on Listening Comprehension (out of 4)	515	.493 (.903)	529	.641 (.996)	477	.549 (.919)	1521	0.562 -0.943	.148	.056	.092	.293
Percent Scored Zero on Listening Comprehension	515	0.693 (.462)	529	.611 (.488)	477	.658 (.475)	1521	.654 (.476)	-.083*	-.035	-.048	.186
School Attendance Over Last 5 School Days	515	3.598 (1.183)	529	3.652 (1.219)	477	3.591 (1.137)	1521	3.615 (1.181)	.054	-.007	.061	.918
Miss More than 10 Days a Year Due to Illness	516	.083 (.087)	529	.053 (.067)	477	.084 (.154)	1522	.073 (.109)	-.03**	.002	-.031	.101
Child Works (1=Yes)	515	.291 (.455)	529	.361 (.481)	477	.285 (.452)	1521	.314 (.464)	.07*	-.006	.076*	.096*
Average Number of Meals Per Day	515	2.299 -0.697	529	2.308 (.757)	477	2.3 (.674)	1521	2.302 (.711)	.009	.001	.008	.984
Hungry During the School Day (1=Yes)	515	.404 (.491)	529	.427 (.495)	477	.407 (.492)	1521	.413 (.493)	.023	.003	.021	0.831
Has Books at Home to Read (1=Yes)	515	.724 (.447)	529	.62 (.486)	477	.667 (.472)	1521	.67 (.47)	-.104***	-.058	-.047	.033**
Reads books at school (1=Yes)	515	0.575 (.495)	529	.597 (.491)	477	.493 (.5)	1521	0.557 (.497)	.023	-.082*	.105**	.048**



	School Meals		+ Teacher Training		+ Community Action		Total		(4)	(5)	(6)	
	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2) - (1)	(3) - (1)	(2) - (3)	(4) = (5) = (6)
Parents Help with Homework (I=Yes)	515	.551 (.498)	529	.586 (.493)	477	.537 (.499)	1521	0.559 (.497)	.035	-.015	.049	.389
Belongs to a Reading Camp (I=Yes)	515	.274 (.446)	529	.314 (.464)	477	.235 (.424)	1521	.275 (.447)	.04	-.039	.079*	.149
Went Hungry Due to COVID (I=Yes)	515	.662 (.473)	529	.629 (.483)	477	.581 (.494)	1521	.625 (.484)	-.033	-.081*	.049	.217
Lives with Mother (I=Yes)	515	.841 (.366)	529	0.832 (.374)	477	0.822 (.383)	1521	0.832 (.374)	-.009	-.019	.01	0.81
Lives with Father (I=Yes)	515	.726 (.446)	529	.732 (.444)	477	.709 (.455)	1521	.723 (.448)	.005	-.018	.023	.74
Speaks Portuguese at Home (I=Yes)	516	.017 (.131)	529	.04 (.195)	477	.031 (.175)	1522	.03 (.169)	.022	.014	.008	.196
F Stat P-Value									0***	.224	.038**	.02**

## **Grade 4 Teachers**

Results in Table 10 show that baseline characteristics for grade 4 teachers are mostly balanced across treatment arms, with few statistically significant differences across the experimental groups. Teachers in treatment 1 have more teaching experience than teachers in treatment 2. In terms of teaching style, a higher proportion of teachers in treatment 2 group students by ability than teachers in the control group. A higher proportion of teachers in treatment 2 have access to the Grade 4 reading curriculum compared to teachers in treatment 2. Other characteristics are balanced across treatment arms. All three individual joint covariate tests (within experimental arms) are significant at the 5 percent level, indicating that some or all the covariates have a non-zero effect on treatment status. However, the p-value from the overall joint covariate test (all characteristics across all arms) is only marginally statistically significant. Once again, we will control for teacher characteristics in our analysis at the impact evaluation endline in 2023.

## **School Deputy Directors**

Results in Table 11 show that baseline characteristics for deputy directors are mostly balanced across treatment arms. Directors in treatment 2 have marginally less experience than teachers in treatment 1 and control groups. Other characteristics are balanced across treatment arms. Only one out of three individual joint covariates tests are significant at the 5 percent level, and the overall joint covariate test is statistically insignificant, signaling adequate balance. Nevertheless, we will also control for director characteristics in our analysis at the impact evaluation endline in 2023.

**Table 10. Baseline balance, teacher characteristics**

	School Meals		+ Teacher Training		+ Community Action		Total		(4)	(5)	(6)	
	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2) - (1)	(3) - (1)	(2) - (3)	(4) = (5) = (6)
Teacher's Biological Sex (1=Male)	54	.315 (.469)	56	.179 (.386)	50	.26 (.443)	160	.25 (.434)	-.136	-.055	-.081	.255
Teacher's Age	54	31.407 (6.566)	56	32.786 (8.583)	50	31.18 (5.872)	160	31.819 (7.141)	1.378	-.227	1.606	.507
Teacher has High School Education	54	.759 (.432)	56	.804 (.401)	50	.8 (.404)	160	.788 (.41)	.044	.041	.004	0.808
Years as a Teacher	54	7.444 (4.729)	56	9.125 (8.195)	50	6.72 (5.315)	160	7.806 (-6.358)	1.681	-.724	2.405*	.193
Teacher Groups Students by Ability	54	.13 (.339)	56	.089 (.288)	50	.02 (.141)	160	.081 (.274)	-.04	-.11**	.069	.046**
Teacher Groups Students in any way	54	.667 (.476)	56	.732 (.447)	50	.66 (.479)	160	0.688 (.465)	.065	-.007	.072	.644
Teacher has Access to Grade 4 Reading Curriculum (1=Yes)	54	.704 (.461)	56	0.821 (.386)	50	.64 (.485)	160	.725 (.448)	.118	-.064	.181**	.083*
Teacher Received Reading Materials	54	.722 (.452)	56	.643 (.483)	50	.72 (.454)	160	0.694 (.462)	-.079	-.002	-.077	.622
Teacher Received Reading Materials from WV/USDA	54	.241 (.432)	56	.286 (.456)	50	.18 (.388)	160	.237 (.427)	.045	-.061	.106	.44
Teacher Participated in Literacy Boost (1=Yes)	54	.593 (.496)	56	.679 (.471)	50	.62 (.49)	160	.631 (.484)	.086	.027	.059	.655
Number of Grade 4 Male Students in Class	54	40.333 (40.544)	56	34.911 (39.255)	50	30.86 (12.449)	160	35.475 (33.817)	-5.423	-9.473	4.051	.223
Number of Grade 4 Females Students in Class	54	36.593	56	26	50	27.3	160	29.981	-	10.593**	-9.293	.125

		School Meals		+ Teacher Training		+ Community Action		Total	(4)	(5)	(6)	
		(34.459)		(14.847)		(13.598)		(23.489)				
School Has Tools to Practice COVID Protection (1=Yes)	54	.648 (.482)	56	.679 (.471)	50	.58 (.499)	160	.637 (.482)	.03	-.068	.099	.545
Grade 4 Teacher Attended 80% or more of Previous 5 days	54	.833 (.376)	56	.714 (.456)	50	.72 (.454)	160	.756 (.431)	-.119	-.113	-.006	.24
F Stat P-Value									0***	.002***	.024**	.059*

**Table 11. Baseline balance, deputy school director characteristics**

	School Meals		+ Teacher Training		+ Community Action		Total		(4)	(5)	(6)	
	N	Mean/SE	N	Mean/SE	N	Mean/SE	N	Mean/SE	(2) - (1)	(3) - (1)	(2) - (3)	(4) = (5) = (6)
Director's Age	54	39.37 (7.713)	56	38.679 (7.143)	50	36.9 (7.291)	160	38.356 (7.41)	-0.692	-2.47	1.779	.237
Director's Biological Sex (1=Male)	54	.852 (.359)	56	.875 (.334)	50	.88 (.328)	160	.869 (.339)	.023	.028	-.005	.918
Director has High School Education	54	.889 (.317)	56	.893 (.312)	50	.9 (.303)	160	.894 (.309)	.004	.011	-.007	.976
Director Teaches at School (1=Yes)	54	.981 (.136)	56	.929 (.26)	50	0.94 (.24)	160	0.95 (.219)	-.053	-.041	-.011	.273
Years as a Teacher	54	15.056 (7.196)	56	14.946 (7.044)	50	12.58 (5.729)	160	14.244 (6.768)	-.109	-2.476*	2.366*	.081*
Director observes classes (1=Yes)	54	0.944 (.231)	56	.982 (.134)	50	.96 (.198)	160	.963 (.191)	.038	.016	.022	.496
School Received Reading Materials	54	0.944 (.231)	56	.911 (.288)	50	.98 (.141)	160	0.944 (.231)	-.034	.036	-.069	.245
School Participated in Literacy Boost (1=Yes)	54	.981 (.136)	56	.964 (.187)	50	.96 (.198)	160	.969 (.175)	-.017	-.021	.004	.759
Number of Male Students in Grade 4	54	43.407 (33.4)	56	41.786 (36.976)	50	37.26 (17.07)	160	40.919 (30.678)	-1.622	-6.147	4.526	.341
Number of Female Students in Grade 4	54	40.019 (29.657)	56	36.857 (37.997)	50	34.1 (20.853)	160	37.063 (30.535)	-3.161	-5.919	2.757	.419
Number of Teachers in School	54	7.926 (6.081)	56	8.946 (-9.499)	50	7.16 (3.956)	160	8.044 (6.993)	1.021	-.766	1.786*	.211
F Stat P-Value									.135	.087*	0***	.168

### 4.3 Determinants of Reading

Tables 12 and 13 below show the correlation matrix between students' performance in the Reading Fluency subtest and the variables related to students' background, teachers' profile, and school management at individual and school level. The results show that most of the variables with statistically significant correlations are related to student background variables. Variables related to teacher characteristics and the deputy school director profile are not significant.

Both levels, individual and school, student age, books at home, student gender and speaking Portuguese at home are significantly correlated with Reading Fluency. Having books at home shows the highest correlation at school level ( $r = 0.25$  at school level and  $0.10$  at individual level see Table 10), followed by speaking Portuguese at home ( $0.21$  at school level see Table 10), student age ( $r = 0.165$  school level and  $0.081$  at individual level), and gender at the individual level ( $0.13$ , see Table 10).

**Table 12. Correlation between performance in reading fluency and social, economic and linguistic variables**

		Oral Reading Fluency	
		Individual	School
Student Age	Pearson Correlation	.081**	.165*
	Sig. (2-tailed)	0.002	0.037
	N	1502	159
Household Asset Index	Pearson Correlation	0.046	0.102
	Sig. (2-tailed)	0.071	0.201
	N	1520	160
Books at home (1=yes 0=no)	Pearson Correlation	.103**	.250**
	Sig. (2-tailed)	0.00	0.001
	N	1520	160
Gender (1 -male 0 - female)	Pearson Correlation	.130**	0.096
	Sig. (2-tailed)	0.00	0.228
	N	1518	160
Number of meals consumed per day	Pearson Correlation	0.032	-0.019
	Sig. (2-tailed)	0.212	0.807
	N	1520	160
Speaking Portuguese at home (1=yes 0 - no)	Pearson Correlation	.053*	.209**
	Sig. (2-tailed)	0.038	0.008
	N	1520	160
** Correlation is significant at the 0.01 level (2-tailed).			
* Correlation is significant at the 0.05 level (2-tailed).			

Table 13 shows the correlation between teachers' characteristics and the deputy school directors' profile. None of the teachers' characteristics variables is significantly correlated with Reading Fluency. A similar pattern of the results is seen in the variables related to the deputy school directors' profiles. One can argue that the schools are similar in terms of teacher and deputy school director characteristics. Another hypothesis could be that the factors related to educational disruption around the school closures due to the pandemic dominate here, engulfing any effect of an individual teacher or deputy school director.

**Table 13. Correlation between reading fluency and teachers' characteristics and school directors' profiles**

			Reading Fluency
Teachers profile	_gender (1-F 0-M)	Pearson Correlation	-0.147
		Sig. (2-tailed)	0.063
		N	160
	mean age	Pearson Correlation	-0.051
		Sig. (2-tailed)	0.521
		N	160
	level of education	Pearson Correlation	-0.082
		Sig. (2-tailed)	0.301
		N	159
	years of experience	Pearson Correlation	-0.091
		Sig. (2-tailed)	0.253
		N	160
	Teach_participative an active method (1-yes 0-No)	Pearson Correlation	0.123
		Sig. (2-tailed)	0.121
		N	160
	teacher participation in Zips Meetings (frequency)	Pearson Correlation	0.143
		Sig. (2-tailed)	0.072
		N	160
school directors profile-Management	school director gender (1-Male 0-Female)	Pearson Correlation	0.078
		Sig. (2-tailed)	0.328
		N	160
	school director Age	Pearson Correlation	-0.018
		Sig. (2-tailed)	0.822
		N	160
	director years of experience in school	Pearson Correlation	-0.016
		Sig. (2-tailed)	0.842
		N	160
	director_c05b_mean years of experience in school	Pearson Correlation	-0.009
		Sig. (2-tailed)	0.908
		N	160
	classroom observation ( frequency)	Pearson Correlation	-0.046
		Sig. (2-tailed)	0.569
		N	154
	Index_school_Infractr_(water, sanitation , electricity )	Pearson Correlation	0.037
		Sig. (2-tailed)	0.644
		N	160
	school council meeting (frequency)	Pearson Correlation	0.089
		Sig. (2-tailed)	0.263
		N	160
	Reading campus (1-yes o-No)	Pearson Correlation	-0.01
		Sig. (2-tailed)	0.899
		N	155
	** Correlation is significant at the 0.01 level (2-tailed).		
	* Correlation is significant at the 0.05 level (2-tailed).		

## Predictors of Student Reading Fluency Using Hierarchical Linear Model (HLM)

The HLM analysis followed three steps. In the first step of the analysis, a null model of HLM was run in order to obtain the amount of variance explained at each level of the hierarchy (Raudenbush and Bryk, 2002). The null model was the simplest model because it contained only the dependent variable (for this study, students' oral reading fluency) and no predictor variables were specified at any level (generating the mean and standard deviation of the variable). The second step was to build up the student-level model or the so-called "unconditional" model at Level-1. This involved adding student-level predictors only to the model. The objective of this step was to examine which student-level variables had statistically significant ( $p < 0.05$  level) effects on the outcome variable. In the next step school variables were added to the model. Variables were only added if they were statistically significantly correlated with reading in the bivariate models.

The data shows that the intra-class correlation (0.09), the correlation in reading scores among students in the same class, see Table 14, is lower when compared with other studies carried out in Mozambique. For instance, large scale assessment studies such as SACMEQ on Grade 6 the intra-class correlation value was around 0.42 in 2013 (Lauchande, 2015), and in the National Evaluation on Grade 3 was 0.47 in 2016 (INDE/MINEDH, 2017). One could argue that the huge difference could be attributed to the similarity of schools within the two districts, when compared with huge variation across the country.

**Table 14. Estimation of the variance components in the empty model**

Variables		Reading	
		Estimate	SE
Grand mean		7.5	0.84
Variance	Student Level	544.6	20.8
Components	School Level	55.7	12.8
Intra-class correlation		0.09	

Table 15 below shows results of the HLM model. At the student level the regression coefficients of age, gender, and books at home are statistically significantly correlated with Reading Fluency ( $p < 0.05$ ). The three predictors in the model accounted for 16.1% of Reading Fluency variation at the school level and 2.6% at individual level. When variables are aggregated at school level, the amount of explained variation increases to 21%, even though most of the variables are not significant. The low share of variance has been explained by the variables we collected suggesting that perhaps there are some other variables that are salient in this context.

In summary, other factors being equal, students who reported having books at home scored 2.7 points (11% of SD) higher than those who reported not having books at home. While keeping the other variables constant, male students scored 6 points (24% of SD) higher than female students (12% of SD). The bivariate analysis and multivariate regression analysis highlight the significance of socioeconomic factors of having and reading books at home. These correlations have also been found in other studies on Primary Education conducted in Mozambique (e.g. Raupp et al., 2013, 2016; Turney et al., 2018) and in other contexts. Many studies have found that school inputs, including teacher education, school facilities, textbooks, and other



inputs have not resulted in higher learning levels (Glewwe and Muralidharan, 2016). Rather, factors such as teacher pedagogy seem to matter most, followed by the home environment. In this case, teacher pedagogical techniques such as splitting students into groups for activities, separating students into groups by ability levels, and using a variety of classroom materials were not significantly correlated with reading scores. It is possible that this result is due to very low levels of these practices and these materials in this context – especially after the COVID-19 school closures.

**Table 15. Estimates for two-level models-reading**

		Null model		Student level		Student/School level	
	<b>Fixed Effects</b>	Coef	SE	Coef	SE	Coef	SE
	Intercept	7.5**	0.84	-2.4-	3.38	-8.0	5.3
<b>Student Level</b>	Age			0.59*	0.28	0.32	0.37
	Gender (1-male ,0 -female)			6.0**	1.2	6.2**	1.29
	Speaking Portuguese at home (1-yes 0-no)			5.9	3.6	3.38	3.8
	Books at Home (1- yes 0-no)			2.17**	0.62	1.7**	0.67
<b>School level</b>	Mean Age					0.71	0.57
	Mean gender					0.34	4.4
	Mean Portuguese home					20.3	11.6
	Mean books Home					3.4	1.8
<b>Random effects</b>							
<b>Variance component</b>	Residual/Student level	544.6**	20.0	530.1**	20.4	529.9**	20.4
	Intercept/School level	55.7**	12.8	46.7**	11.9	43.9**	11.6
<b>Proportion of explained variance</b>	Student level			2.6%		2.7*	
	School level			16.1%		21%	

\*\* significant at 0.01; \* significant at 0.05

## 5 Conclusion and Recommendations

The baseline survey shows that students' EGRA performance is generally poor, with average scores that are less than 50% overall. As argued throughout the report, these poor results can, at least in part, be attributed to the detrimental effects of COVID-19, as schools were closed for about a year. The fact that school level factors such as teacher characteristics and deputy school director profile/management are not relevant in explaining student reading ability reinforces the argument that the negative effects of COVID-19 enguled any effect of an individual teacher or deputy school director.

Despite COVID-19, overall EGRA results obtained in this baseline survey are consistent with those reported in other studies in Mozambique (e.g. Raupp et al., 2013, 2016; Turney et al., 2018). This suggests that other explanatory factors should also be explored and addressed. For this particular study the language of instruction and socioeconomic and demographic factors seem to be the prime influential factors of literacy achievement.

Literacy instruction is conducted in Portuguese in a context where only 3% of the students surveyed reported speaking this language at home (3.5% in Nacarua and 2.5% in Muecate). This means that, as it is common in rural settings in Mozambique, most of these students have their first contact with Portuguese when they start schooling and have scarce access to this language outside their homes. Consequently, at the same time they are struggling to learn the language, they are acquiring basic reading skills. It is certainly due to this language barrier that only 3.8% of the teachers surveyed reported using Portuguese as their primary language of instruction. That is, about 94% of grade 4 teachers report that they use Macua, the students' home language as the primary language of instruction instead of Portuguese, the official language of teaching and learning in the schools in this study. The correlation analysis conducted in this study also shows how a lack of Portuguese competence is key factor of low achievement in literacy. As reported, the results intrinsically correlated with poor results in EGRA. Stakeholders interviewed also noted the potential importance of the introduction of mother tongue based bilingual education in Muecate and Nacarua in enhancing children's opportunities to learning.

The analysis of the relationship between reading performance and socioeconomic factors such as having books at home and having someone at home who reads to the child confirms the importance of a print rich home environment in the children's reading development. This reinforces the ECT3 objectives of creating conditions for children to read at school and in their communities, by creating school and community literacy rich environments (libraries and book banks) as well as setting reading clubs in the children's communities.

One of the findings of this study is that there is correlation between EGRA results and gender. The pattern captured is that boys performed better than girls in reading ability. This may be a reflection of a general lack of sensibility to gender inequalities in schools and also in the communities, in particular in matrilineal settings such as Nampula. In contexts such as these, classroom interactional procedures and practices tend to favor boys. This suggests that ECT3 could reinforce teachers' and parents' awareness on gender inequalities and ensure that girls have equal opportunity as boys for reading practice and growth in class and in homes/communities.

## Impact Evaluation Recommendations

Here, we provide recommendations for the impact evaluation going forward.

1. Recognition of the negative impacts that COVID-19 has had on learning will be important when considering the effects of UL. It is possible that effect sizes may be small, but the interpretation would not mean that the intervention didn't necessarily work. At the time of the IFPRI endline survey in 2023, the research team will carefully collect data on COVID-19 effects and will carefully interpret any results accordingly.
2. As a result of the above, the targets for the ECT3 program may need to be revised downwards. USDA can work with WV to revise ECT3 targets downwards. The program will not be able to meet the targets otherwise (due to COVID). This can be done immediately after the baseline report is approved to ensure no delays in the understanding of goals. The differences between proposal baseline figures and the actual baseline survey figures can be used as a proxy for setting the new targets. For example, if a target is set at 50% but the baseline value is currently 10%, a lower target is more likely to be feasible and achievable.
3. Data from the IFPRI impact evaluation endline in 2023 can be supplemented by World Vision's ongoing monitoring data from the MEQA system and the program performance final evaluation in 2024 to ensure a fuller picture of the effects of the program. The benefit from having process and routine data available at final program performance evaluation is to enrich the findings from the school surveys. Data on classroom observations conducted by either World Vision or government staff would be especially useful. WV should prioritize setting up the MEQA system so that information is available by the final program performance evaluation. In addition to the MEQA system, qualitative interviews from the program final evaluation in 2024 with program staff should also be conducted to get a more nuanced view of rollout and processes.
4. An additional interview with reading camp leaders should be added to the program final evaluation data collection in 2024. Since reading camps are a large focus of the evaluation, data from reading camp leaders will be vital to understanding and interpreting effects on reading.
5. Throughout the remainder of the project, understanding what support is needed in the community will be vital. Understanding why participation levels may be low at both community events and reading camps will complement evidence on teacher effort in the classroom. World Vision could lead this analysis alongside CESC and the community structures CESC is supporting.
6. Similarly, information on WASH training and understanding from the WASH-UP Clubs would be useful. This will enable the program to build more capacity among students in WASH practices and behaviours and will help with the sustainability of these practices beyond the classroom to the household level. Some Wash-Up activities can be included in the reading camp activities to bridge different aspects of the interventions, so they complement each other.
7. Finally, since most students do not speak Portuguese at home, training more officials and teachers in Bi-Lingual Education so that a bilingual approach can be used in the classroom and in the Reading Clubs is essential to building the reading skills that this project aims to achieve.

## References

- Abadzi, H. (2011). Reading Fluency Measurements in EFA Partner Countries: Outcomes and Improvement Prospects, EFA, FTI Secretariat Update September 2011. In: <https://documents.worldbank.org/curated/en/925221468179361979/pdf/797780WPOreadi0Box0379789B00PUBLIC0.pdf>
- Akresh, R., D. Halim, and M. Kleemans. (2018). "Long-Term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia." *NBER Working Paper No. w25265*, National Bureau of Economic Research, Cambridge, MA, USA.
- Alderman, Harold, and Donald Bundy. (2012). School Feeding Programs and Development: Are We Framing the Question Correctly?, *World Bank Research Observer* 27 (2): 204–21.
- Aurino, E., Gelli, A., Adamba, C., Osei-Akoto, I., and Alderman, H. (2020). Food for thought? Experimental Evidence on the Learning Impacts of a Large-Scale School Feeding Program. *Journal of Human Resources*, 1019.
- Barrera-Osorio, F., Gertler, P., Nakajima, N., and Patrinos, H. (2020). *Promoting Parental Involvement in Schools: Evidence From Two Randomized Experiments* (Working Paper No. 28040, National Bureau of Economic Research, Cambridge, MA, USA).
- Beg, S., Fitzpatrick, A., Lucas, A., Tsinigo, E., and Atimone, H. (2019). Strengthening teacher accountability to reach all students (STARS). World Bank SIEF). Milestone, 3.
- Creemers, B. (2002). From school effectiveness and school improvement to effective school improvement: Background, theoretical analysis, and outline of the empirical study. *Educational Research and Evaluation*, 8(4), 343 – 362.
- Daane, M.C., Campbell, J.R., Grigg, W.S., Goodman, M. J. & Oranje, A. (2005) *Fourth-grade Students Reading Aloud: NAEP 2002 Special Study of Oral Reading (NCES 2006-469)*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, DC: Government Printing Office. In: <https://nces.ed.gov/nationsreportcard/pdf/studies/2006469.pdf>.
- Drake, Lesley, Meena Fernandes, Elisabetta Aurino, Josephine Kiamba, Boitshepo Giyose, Carmen Burbano, Harold Alderman, Lu Mai, Arlene Mitchell, and Aulo Gelli. (2017). "School Feeding Programs in Middle Childhood and Adolescence." In *Disease Control Priorities 3*, edited by D. Bundy, N. De Silva, S. Horton, D. Jamison, and G.C. Patton. Washington, D.C.: The World Bank. [http://dcp-3.org/sites/default/files/chapters/DCP3 CAHD\\_Ch 12.pdf](http://dcp-3.org/sites/default/files/chapters/DCP3 CAHD_Ch 12.pdf).
- Glewwe, P., and Muralidharan, K. (2016). Chapter 10 - Improving Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications. In E. A. Hanushek, S. Machin, and L. Woessmann (Eds.), *Handbook of the Economics of Education* (Vol. 5, pp. 653–743). Elsevier.
- INDE/MINED. (2003). *Plano Curricular do Ensino Básico*. Maputo: INDE/MINED.
- INDE/MINEDH. (2017). *Avaliação Nacional da 3ª Classe*. Maputo: INDE/MINEDH.
- Lee, V., Zuze, T. & Ross, K. (2005). School Effectiveness in 14 Sub-Saharan African Countries: Links, with 6<sup>th</sup> Graders Reading Achievement in *Studies in Educational Evaluation*, 31 pp 207-246.

- Lauchande, C. (2015) Contextual Factors Influencing School Effectiveness in Mozambique. <https://www.up.ac.mz/2015-08-21-07-21-15/.../44-revista-udziwi>.
- Lord, F. M. (1980). Practical applications of item response theory. Hillsdale, NJ.
- Maluccio, J., Hoddinott, J., Behrman, J., Martorell, R., Quisumbing, A., & Stein, A. (2009). The Impact of Improving Nutrition during Early Childhood on Education among Guatemalan Adults. *The Economic Journal*, 119(537), 734-763.
- McKenzie, J., Karisa, A., Kahonde, C., & Tesni, S. (2021) Review of UDL in Low-and Middle-Income Countries.
- Mendez, M. A., and Adair, L. S. (1999). Severity and Timing of Stunting in the First Two Years of Life Affect Performance on Cognitive Tests in Late Childhood. *The Journal of Nutrition*, 129(8), 1555–1562.
- Muralidharan, K. (2017). Chapter 3—Field Experiments in Education in Developing Countries. In A. V. Banerjee and E. Duflo (Eds.), *Handbook of Economic Field Experiments*, Vol. 2, pp. 323–385. North-Holland.
- Popova, A., Evans, D. K., Breeding, M. E., and Arancibia, V. (2021). Teacher Professional Development around the World: The Gap between Evidence and Practice. *The World Bank Research Observer*, 36(1), 1-30.
- RTI International. (2015). Early Grade Reading Assessment (EGRA) Toolkit, 2<sup>nd</sup> Edition. Washington, DC: United States Agency for International Development.
- Raudenbush, S. W. & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods, Second Edition. Newbury Park, CA: Sage
- Raupp, M., Newman, B. & Revés, L. (2013) Impact Evaluation for the USAID/Aprender a Ler Project in Mozambique: Baseline Report, Revised in June.
- Raupp, M., Newman, B., Revés, L., Lauchande, C., Alan, E. & Jordan, M. (2016). Impact Evaluation for the USAID / Aprender a Ler Project in Mozambique: Final Report, May.
- Secretariado Técnico de Segurança Alimentar e Nutricional (STSAN), 2018. Relatório de Avaliação Sazonal da Nutrição. Maputo: Ministério de Agricultura e Segurança Alimentar
- Soemantri, A. G., Pollitt, E., and Kim, I. (1985). Iron deficiency anemia and educational achievement. *The American Journal of Clinical Nutrition*, 42(6), 1221–1228.
- Turney, A., Noyes, D., Dhliwayo, T. M., Machkasov, Y., Van de Waal, W. & Hua, H. (2018). Effectiveness Evaluation. Early Grade Reading Assessment & Supplementary Tools. Baseline Report – Mozambique (Draft). Maputo: World Education.

## Annex 1

### ECT2 Final Evaluation Indicators Table (The Indicator Actuals Achieved (FE) serves as the Baseline for ECT3)

MGD	Result	Indicator Gender/Type	Unit	Baseline	MTE	FE	Final Targets	Level of Achieve- ment (%)	Status	
	SOI - Improved Literacy of School-Aged Children									
I	Improved Literacy of School Aged Children	Percentage of students who by the end of two grades of primary schooling demonstrate that they can read and understand the meaning of grade level text (male/female)	Male	%	0.0	34%	18.5	45.0	41.1	Below target
			Female	%	0.0	31%	8.1	45.0	18.0	Below target
			Total	%	25.1	32.0	13.2	45.0	29.3	Below target
		Number of individuals benefiting directly from USDA – funded interventions	Male	Nr.	34088.0	36544.0	43700	36544.0	119.6	Above target
			Female	Nr.	29628.0	31900.0	38495	31900.0	120.7	Above target
			Continuing	Nr	63716.0	68444.0	82195	68444.0	120.1	Above target
I.1	Improved Quality of Literacy Instruction	Percentage of teachers in target schools who demonstrate improved literacy instruction as identified by supervisors, mentors or coaches		%	74.8	100.0	68.8	100.0	68.8	Below target
I.1.1	More Consistent Teacher Attendance	Percentage of teachers in target schools who attend and teach school at least 80% of scheduled days per school year		%	90.0	96.0	75.6	90.0	84.0	Below target
I.1.2	Better Access to School Supplies and Materials	Percentage of Teachers who received textbooks and other teaching and learning materials provided as a result of USDA assistance		%	0.0	100	46.8	100.0	46.8	Below target
		Number of textbooks and other teaching and learning materials provided as a result of USDA assistance		Nr.	N/A	1634.0	90906	2400	3787.8	Above target
		Percentage of classrooms with literacy instruction materials (text books, work books) sufficient for effective instruction		%	80.0	90.0	82.2	100.0	82.2	Below target
I.1.3	Improved Literacy Instructional Materials	Number of classrooms with literacy instructional materials (textbooks, workbooks,) sufficient for effective instruction*		Nr.	337.0	406		450.0		
I.1.4	Increased Skills and Knowledge of Teachers	Number of teachers/educators/teaching assistants in targeted schools who demonstrate use of new and quality teaching techniques or tools as a result of USDA assistance †		Nr.	0	566.0		566.0		
		Number of teachers/educators/teaching assistants trained or certified as a result of USDA assistance		Nr.	0	708.0	773	708.0	109.2	Above target
I.1.5	Increased Skills and Knowledge of School Administrators	Number of school administrators and officials in target schools who demonstrate use of new techniques or tools as a result of USDA assistance		Nr.	0	106.0	146	240.0	60.8	Below target
		Number of school administrators and officials trained or certified as a result of USDA assistance		Nr.	300.0	282.0	282	300.0	94.0	Below target

MGD	Result	Indicator Gender/Type		Unit	Baseline	MTE	FE	Final Targets	Level of Achieve- ment (%)	Status
I.2	Improved Attentiveness	Percentage of students in target schools identified by their teachers as attentive during class/instruction*		%	98.0	90.0		98.0		
		Percentage of students who feel more attentive because of the food they get at school †		%			73.6			
I.2.1	Reduced Short-Term Hunger	Percentage of students in target schools who indicate that they are not hungry during school day		%	52.6	100.0	58.8	100.0	58.8	Below target
I.2.1.1	Increased Access to Food (School Feeding)	Number of individuals receiving take-home rations as a result of USDA assistant		Nr.	0.0	728.0	765	728.0	105.1	Above target
		Number of take-home rations provided as a result of USDA assistant		Nr.	0.0	527,289.0 0	7520000	9726666.7	77.3	Below target
		Number of school aged children receiving daily school meals (breakfast, snack, lunch) as a result of USDA assistance	Male	Nr.	32775.0	35084.0	39,850	35084.0	113.6	Above target
			Female	Nr.	24726.0	30722.0	34,884	30722.0	113.5	Above target
			Continuing	Nr.	57501.0	65806.0	74,734	65806.0	113.6	Above target
		Number of daily school meals (breakfast, snacks, lunch) provided to school aged children as a result of USDA assistance		Nr.	12,615,812	22,174,98 0	39,644,345	51,093,460	77.6	Below target
		Number of social assistance beneficiaries participating in productive safety nets as result of USDA assistance	Male	Nr.	N/A	36246.0	43,391	36246.0	119.7	Above target
			Female	Nr.	N/A	31499.0	38,066	31499.0	120.8	Above target
			Continuing	Nr.	N/A	67745.0	81,457	67745.0	120.2	Above target
I.3	Improved Student Attendance	Percentage of students that attend school at least 80% of the school days	Male	%	96.7	85.2	44.7	98.0	45.6	Below target
			Female	%	84.8	74.8	43.2	92.0	46.9	Below target
			Continuing	%	90.8	80.0	43.95	95.0	46.26	Below target
I.3.1	Increased Economic and Cultural Incentives (Or Decreased Disincentives)	Number of schools benefiting from saving groups social funds as result of USDA assistance		Nr.	N/A	146.0	160	120.0	133.3	Above target
I.3.2	Reduced Health Related Absences	Percent of students in target schools who miss more than 10 school days/year due to illness	Male	%	N/A	N/A	24.6	10	246	Above target
		Percent of students in target schools who miss more than 10 school days/year due to illness	Female	%	N/A	N/A	33	10	330	Above target
		Percent of students in target schools who miss more than 10 school days/year due to illness	Continuing	%	2.1	2.0	29.4	10	294	Above target
I.3.3	Improved School Infrastructure	Number of educational facilities (i.e., school buildings, classroom, and latrines)	Schools	Nr.	11.0	17.0	48	15.0	320	Above target
			Storerooms	Nr.	150.0	48.0	95	90.0	105.6	Above target
			Latrines	Nr.	440.0	1430.0	1595	1595.0	100	On target

MGD	Result	Indicator Gender/Type		Unit	Baseline	MTE	FE	Final Targets	Level of Achieve- ment (%)	Status
		rehabilitated/const ructed as a result of USDA assistance	Canteens	Nr.	150.0	123.0	250	150.0	166.7	Above target
			Wells	Nr.	N/A	36.0	43	48.0	89.6	Below target
			Small Water	Nr.	N/A	12	12	12	100	On target
1.3.4	Increased Student Enrollment	Number of students enrolled in school receiving USDA assistance	Male	Nr.	32775.0	35084.0	42891	35084.0	122.3	Above target
			Female	Nr.	24726.0	30722.0	37623	30722.0	122.5	Above target
			Continuing	Nr.	57501.0	65806.0	80514	65806.0	122.4	Above target
1.4.1	Increased Capacity of Government Institutions	Number of district and Provincial MINED officials who know standard operating procedures and tools for management and oversight of school feeding programs and Literacy Boost		Nr.	0	48.0	48	48.0	100	On target
1.4.3	Increased Government Support	Percentage of Education coordination meetings where school feeding is part of the agenda		%	0	31.0	100	100.0	100	On target
1.4.4	Increased Engagement of Local Organizations and Community Groups	Number of community groups that report having access to meaningful opportunities for feedback about project implementation		Nr.	0.0	37.0	41	37.0	110.8	Above target
		Number of Parent-Teacher Association (PTAs) or similar “school” governance structures supported as a result of USDA assistance		Nr.	0.0	160.0	160	160.0	100	On target
		Number of public - private partnership formed as result of USDA assistance (nutrition)		Nr.	0.0	80.0	80	80.0	100	On target
	SO2-Increased Use of Health & Dietary Practices									
2	Increased Use of Health and Dietary Practices	Percent of target beneficiaries who use appropriate hand washing practices	Male	%	0.0	94.0	88.7	95.0	93.4	Below target
			Female	%	0.0	95.0	90.3	95.0	95.1	Below target
			Continuing	%	0.0	94.0	89.3	95.0	94.0	Below target
		Percent of school- aged children receiving a minimum acceptable diet	Male	%	0.0	53.0	38.0	53.0	71.69	Below target
			Female	%	0.0	47.0	40.6	47.0	86.38	Below target
			Continuing	%	0.0	100	39.4	100	39.4	Below target
2.1	Increased Knowledge of Health and Hygiene Practices	Percentage of children in target communities who can identify at least 3 important health/hygiene practices		%	61.7	86.0	31.6	90.0	35.1	Below target
2.2	Increased Knowledge of Safe food prep and Storage Practices	Percentage of food preparers at target schools who can identify at least 3 key practices aimed at safe food preparation		%	0	100.0	99.4	100.0	99.4	Below target
2.3	Increased Knowledge of Nutrition	Number of individuals trained in child health and nutrition as result of USDA assistance	Male	Nr.	N/A	482.0	520	818.0	63.6	Below target
			Female	Nr.	N/A	1005.0	1017	546.0	186.3	Above target
			Continuing	Nr.	N/A	1487.0	1537	1364.0	112.7	Above target
2.4	Increased Access to Clean Water and Sanitation Services	Number of schools using an improved water source		Nr.	150.0	160.0	129	160.0	80.6	Below target
		Number of schools with improved sanitary facilities		Nr.	40.0	130.0	159	145.0	109.7	Above target



MGD	Result	Indicator Gender/Type	Unit	Baseline	MTE	FE	Final Targets	Level of Achieve- ment (%)	Status
2.5	Increased Access to Preventive Health Interventions	Number of schools who receive at least 2 visits per year from health facility staff	Nr.	N/A	147.0	160	160.0	100	On target
		Number of students receiving de-worming medication(s)	Nr.	N/A	55949.0	71258	59411.0	119.9	Above target
2.6	Increased Access to requisite Food prep and storage tools and equipment	Number of schools with appropriate food prep and storage equipment	Nr.	150.0	160.0	160	150.0	106.7	Above target

Note: Indicators denoted with an asterisk (\*) were not measured during the endline survey. Indicators denoted by † are the indicators used as proxies for those that were not measured.