



Tanzania McGovern-Dole Food for Education Program

Midterm Evaluation

September 2019

This publication was produced at the request of the United States Department of Agriculture. It was prepared independently by Trembley Consulting

Project Concern International Tanzania, McGovern-Dole Food for Education Program Midterm Evaluation

This midline evaluation was designed to focus on multisectoral program outputs, outcomes, and impacts over the period from May 2017 to May 2019. The evaluation aims to improve PCI's understanding of its performance to date and to inform the leadership team in making improvements and adjustments in the final two years of programming to ensure optimal outcomes and sustainability.

Agreement Number: FFE-621-2016/011-00

Project Duration: 2016-2021

Implemented by: Project Concern International

Evaluation Authored by: Trembley Consulting

Table of Contents

Table of Contents	3
Acronym List	7
Executive Summary.....	8
Section 1: Mara Region Context and Project Background	15
Project Design and Key Activities.....	19
Baseline Survey Overview and Key Recommendations	22
Section 2: Midline Evaluation Purpose and Questions	24
Relevance	24
Effectiveness	24
Sustainability	25
Literacy	25
Gender	26
Section 3: Midline Evaluation Methodology and Limitations	27
Quantitative Data.....	28
Qualitative Data	30
Empirical Approach to Impact Identification	33
Limitations of the Evaluation	34
Section 4: Key Performance Findings: Relevance, Implementation, and Beneficiary Satisfaction	37
Communications and Coordination	37
Teacher Training and Literacy Activities	38
School Feeding Activities	39
Infrastructure, Sanitation, and Health	41
Improved Agricultural Practices.....	42
WE Groups	43
Section 5: Key Impact Findings Related to Strategic Objective 1: Improved Literacy of School-Age Children	45
School Enrollment.....	45
Student Attendance	47
Student Attentiveness	51
Quality of Literacy Instruction and School Management	53
Student Literacy Outcomes.....	56

Section 6: Key Impact Findings Related to Strategic Objective 2: Increased Use of Health and Dietary Practices	62
Student Knowledge of Health, Hygiene, and Nutrition Practices	62
Caregiver Knowledge of Health, Hygiene, and Nutrition Practices and Benefits of Education	64
School Water and Sanitation and Health Management Practices	65
Dietary Quality	68
Section 7: Key Findings Related to Program Sustainability	70
Parental Engagement and School Feeding Program Contributions	70
Farmer Groups	78
WE Groups	80
Health and Hygiene Activities	82
Government Roles	84
Section 8: Major Conclusions and Program Recommendations	88
Beneficiaries are satisfied with program delivery.	88
Literacy is difficult to influence in highly resource-scarce environments.	88
WE and Farmer Groups and school gardens demonstrate positive, sustainable impacts.	88
Health, nutrition, and sanitation indicators are not negative, but could be more insightful	89
Government engagement is positive but public sector ownership is lacking.	89
Heavy emphasis on parental contributions risk making vulnerable households worse off.	89
Recommendations for Program Management	91
Bibliography	94
Annex A: Results Framework	96
Annex B: Data Collection Tools List	97
Annex C: Distribution of Qualitative Extracts	99
Annex D: Midline Data Collection Timeline	101
Annex E: Methodology for Calculating Minimum Acceptable Diet (MAD)	102
Annex F: Recommendations for Endline Survey and Evaluation	103
Annex G: Original Terms of Reference	106
Annex H. Inception Report	118

List of Tables

16 Priority Program Indicators.....	13
Geographical Coverage of Food for Education Programming by Phase.....	19
School Feeding Commodities Transition Schedule.....	20
School Feeding Commodities Transition Schedule.....	20
Quantitative Data Sample Sizes	28
Quantitative Data Sample Sizes	28
EGRA Sample.....	29
Classroom Observation Sample	30
Focus Group Discussion Subjects.....	31
Key Informant Interview Subjects.....	31
Thematic Coding Scheme for Qualitative Data	32
Average Number of Students Enrolled in Schools	45
Percent of Parents with Knowledge of Primary School Educational Benefits	46
Main Benefits of Child Going to School	46
Illness-Related Absence and Regular Attendance	47
Percent of Students Indicating They Are Full During they Day.....	51
Percent of Students Off Task during Classroom Lesson	51
Percent of Head Teachers Reporting that Schools had Gardens and Agricultural Demonstration Plots...	53
Percent of FFE School Head Teachers who Received Training from PCI by Most Common Training Type	53
Classroom Displays and Materials	54
School Administrators in Target Schools who Demonstrate use of Management Practices	55
Percent of Students Scoring Above National Benchmarks in Non-Word Reading by Grade and Sex	56
Average number of words per minute on non-word Assessment	58
Percent of Students Scoring Above National Benchmarks in Oral Fluency by Grade and Sex	59
Average number of words in 63-word passage read in one minute, by Grade.....	60
Percent of Students who Demonstrate that They Can Read and Understand the Meaning of Grade-Level Text by Grade and Sex	60
Average Number of Comprehension Questions Answered Correctly by Grade.....	61
Percent of Students with Knowledge of Good Health and Hygiene and Important Nutritional/Dietary Practices	62
Percent of Parents with Knowledge of Important Preventative Health Practices	64
Number and Percentage of Schools with an Improved Water Source or Improved Sanitation Facilities..	65
Number and Percentage of Schools with an Improved Water Source or Improved Sanitation Facilities..	66

Average Number of Students per Toilet Facility at School (midline only)	66
Percent of Schools with Active Committees, Volunteer Student Aids, and Health Clubs	67
Percent of Students Receiving a Minimum Acceptable Diet (MAD).....	68
School Feeding Program Commodity Contributions from June 2018 through April 2019 by Source and Commodity Type for 50 Midline Evaluation Schools	70

List of Figures

Seasonal Volatility in Average Tanzanian Maize Prices, January 2017 – March 2019.....	16
Difference-in-Difference Impact Identification.....	33
Distribution and Average Number of Non-Words Read in One Minute, Grade 4	57
Community Contributions of Maize and Beans by Month in 50 Midline Evaluation Schools	73
Agricultural Calendar in Tanzania	73

Acronym List

COSTECH	Commission for Science and Technology
DPO	District Planning Officer
EGRA	Early Grade Reading Assessment
EQUIP	Education Quality Improvement Program Tanzania
FFE	McGovern-Dole Food for Education Program
IRB	Internal Review Board
FGD	Focus Group Discussions
KII	Key Informant Interview
LGA	Local Government Authority
MAD	Minimum Acceptable Diet
MGD	McGovern-Dole Program
NA	Not Applicable
NBS	National Bureau for Statistics
NGO	Non-Governmental Organization
PCI	Project Concern International
PORALG	Prime Minister's Office of Regional and Local Governance
PTP	Parent-Teacher Partnership [Committee]
REO	Regional Education Officer
SWASH	School Water, Sanitation, and Hygiene
TOR	Terms of Reference
ToT	Training of Trainers
TZS	Tanzanian Shillings
UNICEF	United Nations Children's Fund
USDA	United States Department of Agriculture
VSA	Volunteer Student Aids
WE	Women's Empowerment

Executive Summary

Project Concern International (PCI) has continuously implemented McGovern-Dole Food for Education (FFE) programs across Mara Region since 2010. The first phase of the program was active from September 2010 through December 2014. Phase II, which ran from September 2013 through December 2016, continued work in Phase I schools while expanding into new communities. Phase III began in October 2016 and runs through December 2021. It has maintained a consistent presence in Phase I and II communities while increasing coverage further into remote areas.

The majority of households in Mara Region depend primarily, if not exclusively, on agricultural activities to meet their basic needs. Poverty is pervasive across the country; the World Bank assessed that four fifths of Tanzanians lived below the international poverty line of \$3.20 per day based on the most recent available data (2011). UNICEF has observed that 74 percent of all Tanzanian children, and 81 percent of children in rural areas, experience multidimensional poverty. Childhood welfare is strongly associated with maternal education, even when mothers have just some primary schooling.

Tanzania achieved near universal access to primary education in 2007. However, just 8 percent of second grade students in mainland Tanzania are able to read. Over a quarter of rural children aged 5 to 13 cannot read or write, and 14 percent are two or more years behind grade level in school. In January 2016, Tanzania introduced the Fee-Free Basic Education Policy covering pre-primary, primary, and lower secondary schools. The policy resulted in massive enrollment increases for Standard I in 2016 and 2017, contributing further to overcrowding and pressure on school infrastructure. Over 5,100 schools had Standard I classes with over 100 students in 2016 (Government of Tanzania, 2018).

FFE III activities are now active in 231 school communities and had directly benefitted 206,388 school children and adults by March 2019. PCI estimates that FFE III will reach a total of over 440,000 direct and indirect beneficiaries over the life of the project.

The multisectoral program includes a wide variety of complementary initiatives designed to improve student literacy. The primary intervention has been to provide a once daily school meal five days per week in new schools. The meal is comprised of rice, pinto beans, and oil, which until recently has been provided by the U.S. Department of Agriculture (USDA), supplemented by community contributions and crops grown on school grounds).

In addition to school feeding, PCI supports an array of activities which work in concert to improve student attendance and attention, teacher and instructional quality, school infrastructure and sanitation, and the sustainability of all program components after the project closes in 2021. The technical design of the program is organized around the achievement of two strategic objectives: to improve literacy of school-age children and to increase use of health and dietary practices. PCI works to improve literacy by augmenting the quality of literacy instruction, increasing school attendance, and improving student attentiveness. It also supports Women's Empowerment (WE)

savings and loan groups and farmer groups to relax household economic constraints for school and other expenses and to increase local food production capacity.

This midline evaluation was designed to focus on multisectoral program outputs, outcomes, and impacts over the period from May 2017 to May 2019. The evaluation aims to improve PCI's understanding of its performance to date and to inform the leadership team in making improvements and adjustments in the final two years of programming to ensure optimal outcomes and sustainability. The evaluation's objectives are as follows:

- To assess the relevance of the project's technical strategy and implementation approach, including the validity of assumptions made during project design
- To measure progress on key indicators during the two years since the baseline survey was conducted and to identify both enablers and constraints impacting that progress
- To provide recommendations for mid-course corrections and performance strengthening
- To assess sustainability efforts to date
- To document lessons learned, challenges, and unanticipated outcomes
- To provide recommendations for areas of focus within the final evaluation, in terms of both substantive focus and evaluation process

The evaluation drew from extensive quantitative and qualitative data collection and analysis that relied upon comparisons with a set of quasi-control¹ non-intervention schools to focus on three key topics: program relevance, effectiveness, and sustainability. Community leaders and members alike agreed that program activities addressed local needs while citing a daunting array of challenges in the education sector. Regional, district, and ward officials provided unanimously positive feedback on PCI's efforts to coordinate and collaborate with government leaders in their work.

Students, their parents, and teachers overwhelmingly expressed satisfaction with school feeding activities. They found commodity management and food preparation processes to be well organized and the food to be of high quality. Students liked the food's taste, and in some cases requested larger serving sizes.

To illuminate program effectiveness, trends in student attendance, attention, and literacy were examined. Two measures of school attendance were recorded – one sourced from parental reporting and the other based on classroom record keeping. Neither data source yielded statistically significant differences in attendance between treatment and comparison schools. The challenges of chronic absenteeism and truancy surfaced frequently in the qualitative data. Irregular attendance was widely cited as a confounder of educational outcomes by stakeholders in both PCI program treatment and Serengeti comparison schools. Significant differences in classroom attentiveness across groups were not detected.

¹ The group of Serengeti District comparison schools is referred to as "quasi-control" because the treatment and comparison schools were not randomized.

Literacy outcomes were measured with the Early Grade Reading Assessment (EGRA), an internationally standardized methodology. The primary indicator of interest is defined as the percent of students who score above national benchmarks in non-word reading, with the national benchmark set at reading 40 devised words per minute. The proportion of second grade boys and girls who could meet the benchmark increased in treatment schools and decreased in control schools, resulting in a significant difference collectively and also for each gender separately. The percent of fourth graders achieving the same benchmark fell in both treatment and comparison schools. The mixed results of gains in 2nd grade and declines in 4th grade are less puzzling if one considers the indicator design. Measuring the percent that reach 40 words per minute means that even under ideal circumstances, this indicator ignores students who improve from reading but not up to 39 words per minute, or 41 words per minute to a higher level. Reviewing the average scores on the non-word assessment shows negative, but less dramatic results.

Improved knowledge and practice of healthy behaviors was measured among both students and parents, in quantitative and qualitative data collection efforts. No statistically significant changes between treatment and comparison groups at midline were observed. It is plausible that the lack of robust quantitative results in this area stems from inconsistent measurement across survey waves and the imprecise definition of variables. In some cases, baseline values could not be replicated². Qualitative findings did surface student knowledge of hygienic behaviors such as hand washing and maintaining a clean environment, as well as positive reviews of health clubs.

Enhanced dietary and health knowledge combined with increased access to nutritionally diverse and safely prepared food was expected to support improved student dietary diversity. A student is assessed to have a Minimum Acceptable Diet (MAD) if he or she ate three or more meals on the previous day in which food from at least four defined food groups is reported. At baseline, about 40 percent of both treatment and comparison students recorded a Minimum Acceptable Diet. Measurement differences across survey waves make midline comparisons imperfect, but the proportion of treatment school students with a MAD increased to 57 percent, a statistically significant 10 percentage points higher than in comparison schools.

Improved water sources and sanitation facilities appear to be in decline since baseline: the number of facilities with improved sanitation appear to be declining, and the number of students per functioning bathroom facility is well above international standards. While the cause of the decline is unclear, the overall poor state of washroom facilities and water sources suggest this is a worthy area of focus for the remainder of the project; however, before doing so, it is worth investigating whether past improvements are still functioning, and how to adjust implementation if they are not.

² This report is accompanied by detailed notes on instruments, analytical datasets, and annotated code to allow the endline evaluation to replicate the midline methodology and avoid future issues.

The sustainability model of the program's school feeding component is organized around parents contributing food in sufficient quantities to fully offset USDA commodity donations. In the midline survey, 95 percent of parents stated a willingness to contribute to school feeding, but only 38 percent had done so in the preceding six months. In contrast, 91 percent had contributed to related expenses such as guards and cook fees.

Over an 11-month period from June 2018 through April 2019, the 50 treatment schools included in the evaluation recorded a total of nearly 400,000 kg of new food commodity receipts. Of that total, 79 percent (by weight) was sourced from USDA while about 14 percent came from community contributions. Most of the balance came from school farms with a small portion from school gardens.

Many parents and other respondents referenced the significant challenge that drought and climate change conditions have posed for local food production, heightening food insecurity at the household level. A trend emerged that parents articulated a strong willingness to contribute food for school feeding under improved agricultural production conditions but lamented their current inability to do so. A sizable proportion of teachers, parents, and students themselves expressed concern that school feeding would not continue after PCI's withdrawal of support.

WE groups appear to be a highly successful intervention of the project. Over one in five parents (and over one in four women) in the parents' survey reported being a current or past member. Members expressed a clear sense of ownership over group management and processes. One prevalent challenge across WE groups was the correlated timing of peak loan demand, during the dry season. This led to an inability to borrow at times of greatest need due to a lack of funds in the cash box. Many focus group participants agreed that saving money at home was nearly impossible due to its accessibility, both to themselves and other family members. Group savings allowed them a more disciplined, and successful, alternative. All WE group participants were certain that the groups would continue after PCI support was withdrawn.

Farmer groups were successful in reaching a broad cross-section of community members. Over one fifth of parents interviewed in the midline survey indicated that they were current (12%) or past (9%) farmer group members. Farmers who participated in focus group discussions were generally satisfied with the quality of trainings and activities. Many described improvements in practices, crop quality, and yields. In other areas, program activities focused on maize cultivation while encouraging farmer groups to plant alternate crop varieties or use improved seed varieties. contribute other types of seeds. Some farmers observed that the soil does not sustain maize well. Multiple respondents also remarked upon the sustainability of school gardens. Evidence suggests these groups have been successful in making long-term improvements that will extend beyond the program.

Recommendations:

- Recalibrate expectations for literacy outcomes and impacts given the many prevailing challenges to quality education in Mara Region, as well as the remaining timeline of program activities. Adjust indicator to monitor average reading scores, instead of percent meeting national benchmark
- Make efficient and evidence-based choices with program resources, including human resources, in the remaining two years of implementation. WE groups and farmer groups both emerged as successful and sustainable models.
- As the baseline evaluation previously recommended, pivot focus from a household food donation model to a community food production model, guided by the input of many evaluation respondents across both treatment and comparison communities.
- Communicate to program staff, partners, and government at all levels that community food contributions are voluntary. The program was always designed to incorporate voluntary contributions but there are reports of non-voluntary contributions from the qualitative data that warrant further investigation.
- Work closely with government actors at all levels to jointly devise non-coercive sustainability plans supported by public monetary and non-monetary resources (such as the provision of public lands for school food production).
- Explore lessons for program sustainability drawing from the experiences of Phase I and Phase II schools. On a host of metrics, strong gains for Phase III schools coincide with declines or stagnation of activity levels among earlier phase schools. If local ownership of activities has not taken root as anticipated in communities with active programming since 2010 or 2014, seek understanding of the underlying reasons. In its influence of national school feeding policy, PCI would add value by advocating for clear and distinct budgetary allocations to cover commodity procurement, management, and preparation costs (including kitchen infrastructure) with limited or no transferability to other education sector expenditures. It should also focus efforts on disseminating both successes and challenges that have emerged in the examination of its own programming to support the Government of Tanzania's school feeding momentum while avoiding potential pitfalls

Table 1: Priority Program Indicators

	Group	Baseline	Midline	1 st Difference ³	2 nd Difference
MGD #26					
1a. Percent of students that score above national benchmarks in non-word reading (2 nd graders)	Treat	5%	13%	8%	14.6%* ⁴
	Control	9.7%	3%	-6.6%	
1b. Percent of students that score above national benchmarks in non-word reading (4 th graders)	Treat	17.6%	7.3%	-10.3%	-6.4*
	Control	9.7%	5.7%	-3.9%	
2a. Percent of students that score above national benchmarks in oral fluency reading (2 nd graders)	Treat	19.3%	3.3%	-16%	8.5%*
	Control	25.3%	0.8%	-24.5%	
2b. Percent of students that score above national benchmarks in oral fluency reading (4 th graders)	Treat	45%	26.7%	-18.3%	-12.1%*
	Control	35%	28.8%	-6.2%	
3a. Percent of students who demonstrate that they can read and understand the meaning of grade level text (2 nd graders)	Treat	7.3%	8%	0.7%	14.5%*
	Control	16.3%	2.5%	-13.8%	
3b. Percent of students who demonstrate that they can read and understand the meaning of grade level text (4 th graders)	Treat	25%	32.2%	7.2%	-7.7%*
	Control	22%	36.9%	14.9%	
4. Percent of students who were “off task” at the beginning, middle, or end of a classroom lesson	Treat	3.7%	5.0%	1.3%	1.4%
	Control	4.6%	4.5%	-0.1%	
5. Percent of students who indicate that they are “full” during the school day	Treat	87.9%	51.8%	-36.1%	7.9%*
	Control	91%	47%	-44%	
6. Percent of students who were absent from school due to illness in the past six months (two-week reference period, parental report)	Treat	4.6%†	6.3%	1.7%	-3.1%
	Control	4.2%†	9.0%	4.8%	
7. Percent of school-age children receiving a minimum acceptable diet	Treat	40.9%†	56.8%	15.9%	10.8%*
	Control	41.3%†	46.4%	5.1%	
8. Percent of students that can identify 3 good health/hygiene practices	Treat	66.6%†	17.3%	-49.3%	10.8%*

³ First differences are the change in indicator values over time within treatment or control groups; for example improvement in the treatment group between baseline and midline. Second differences are the gap between improvement in the treatment group and improvement in the control group (the difference of first differences).

⁴ Asterisks (*) indicate statistical difference, tested at the 95 percent confidence level.

	Control	75.5%†	15.4%	-60.1%	
9. Percent of Mothers that can identify 3 important preventative health practices	Treat	40.5%	26%	-14.5%	-4.3%
	Control	41.5%	31.3%	-10.2%	
10. Percent of Students that can identify 3 important nutritional/dietary recommendations	Treat	36.9%	46.8%	9.9%	1.3%
	Control	39.4%	48%	8.6%	
11. Percent of Parents that can identify 3 important preventative health practices	Treat	68.3%	39.8%	-28.5%	-8%
	Control	64.3%	43.8%	-20.5%	
12. Percent of Parents/Guardians who can name 3 benefits of completing primary school	Treat	47%	96%	49%	-5.3%
	Control	42%	96.3%	54.3%	
13. Average Number of Students Enrolled in School through May 2017/2019	Treat	623	726	103	-113
	Control	676	893	217	
MGD #1 14. Percentage of Students regularly attending school (80%) in last six months	Treat	82.9%†	72%	-10.9%	6.5%
	Control	81.1%†	63.7%	-17.4%	
MGD #22 15. Percentage of Schools using an Improved Water Source	Treat	76%†	65% (149)	-11%	
	Control	32%†	∅		
MGD #23 16. Percentage of Schools with Improved Sanitation Facilities	Treat	86.5%†	79% (183)	-7.5%	
	Control	80.6%†	∅		
MGD #3 Number of school administrators who demonstrate use of new techniques or tools	Treat	66.0%	82.0%	16.0%	6.0%
	Control	84.0%	94.0%	10.0%	
MGD #4 Number of educators who demonstrate use of new and quality teaching techniques or tools	Treat	71.7%	71.2%	-0.5%	-2.0%
	Control	72.0%	73.5%	+1.5%	

Note: Values marked with a † are taken from the baseline report. Given a lack of information on how baseline values were calculated, all baseline values were recalculated using the methods used at midline for comparability with these exceptions, due to unknown baseline sources or changes in the instruments†. Values marked with a * denote statistically significant difference. Values marked with a ∅ are for indicators 15 and 16, where project management data are used at midline, and therefore are not available in control schools.

Section 1: Mara Region Context and Project Background

Mara Region, one of 31 administrative regions in Tanzania, occupies the northernmost stretch of the country's Kenyan border, abutting Lake Victoria from the east. The most recent national census,⁵ (Government of Tanzania 2012), in 2012, recorded a population of 1.74 million residents across seven administration districts: Musoma Urban, Musoma Rural, Bunda, Butiama, Rorya, Tarime, and Serengeti. Four percent of mainland Tanzania's population resided in the region at that time. The United Nations has estimated an annual population growth rate of 3 percent nationally between 2012 and 2018 (World Bank 2019). This rate of growth implies a current regional population of approximately 2.14 million residents in 2019.

PCI programming is active in Musoma, Bunda, and Butiama, which had a combined population of nearly 890,000 inhabitants in 2012, or 1.09 million as projected to 2019. Neighboring Serengeti, which has been adopted throughout this analysis as a comparison district, had just under 250,000 people when the census was administered, implying a current population of approximately 307,000.

Poverty is pervasive across the country; the World Bank assessed that four fifths of Tanzanians lived below the international poverty line of \$3.20 per day based on the most recent available data (World Bank, 2019). A 2016 UNICEF study which drew from the 2012/13 National Panel Survey observed that 74 percent of all Tanzanian children, and 81 percent of children in rural areas, experience multidimensional poverty. To meet this classification, these children confront poverty in at least three of the following dimensions: nutrition, health, protection, education, information, sanitation, water, and housing (UNICEF, 2018). The study further reported that parental education is the strongest predictor of childhood poverty. Childhood welfare was strongly associated with maternal education, even when mothers have just some primary schooling.

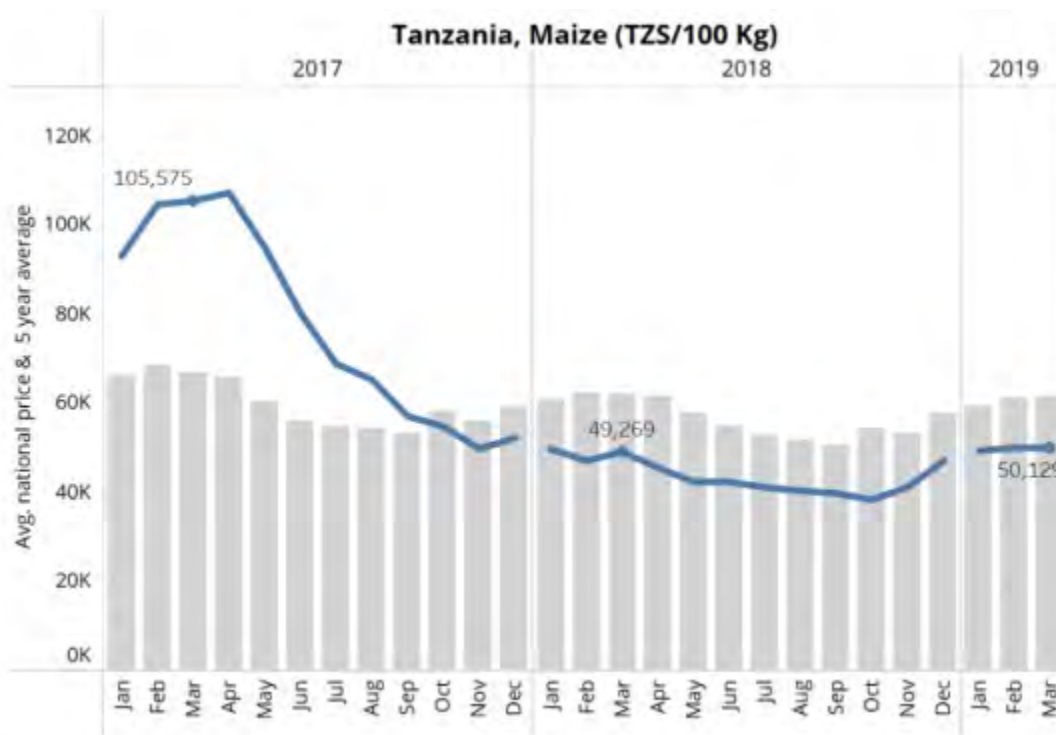
Livelihoods and Household Economics

The majority of households in Mara Region depend primarily, if not exclusively, on agricultural activities to meet their basic needs. A 2018 assessment found that 90 percent of residents rely on agriculture, with crop production concentrated on cassava, maize, sorghum, finger millet, beans, and sweet potatoes (UNICEF, 2018). Cotton is the principle cash crop. Fishing is also prominent, employing 40 percent of the region's labor force according to one source (World Bank, 2019). Seasonal and inter-annual food price volatility correlates with the planting and harvest calendar, making it challenging for smallholder farmers to smooth household consumption throughout the year in the absence of inclusive financial services and reliable storage facilities. National average prices presented in Figure 1 mask starker local market variance.⁶ Grey bars represent the 5-year average price while the blue line tracks the most recent 12 monthly average prices per 100 kilograms of maize. Based on September 2019

⁶ FEWS Net no longer collects data on regional Tanzanian markets.

conversion rates, the average March 2019 maize price in Tanzania was comparable to average prices in Zambia and South Africa, and less expensive than maize in Mozambique.

Figure 1: Seasonal volatility in Tanzanian Maize Prices, January 2017 – March 2019 (World Food Program 2019)



The region is prone to recurrent drought as well as flooding, and the staple crop, maize, is particularly drought-sensitive. The 2016-17 Agriculture Sample Survey reported that just 2.5 percent of total planted area nationwide was irrigated (Tanzania National Bureau of Statistics, n.d.). In Mara, irrigation reaches only 0.6 percent of agricultural land (Ibid.). The rapid spread of Fall Armyworm in the past two years has introduced an additional threat to regional food security (SouthernAfrican Development Community, 2018).

Education

Tanzania achieved near universal access to primary education in 2007. However, in the decade that followed, primary school enrollment fell while challenges to instructional quality were pervasive, especially in rural areas (UNICEF, 2019). The 2014 primary school leaving survey revealed that just 8 percent of second grade students in mainland Tanzania could read properly (Ibid.). UNICEF finds that 27 percent of rural children aged 5 to 13 could not read or write, and that 14 percent were two or more years behind grade level in school.

In January 2016, Tanzania introduced the Fee-Free Basic Education Policy covering pre-primary, primary, and lower secondary schools. The policy resulted in massive enrollment increases for Standard I in 2016 and 2017, contributing further to overcrowding and pressure on school infrastructure. Over 5,100 schools had Standard I classes with over 100 students in 2016 (Government of Tanzania, 2018). The gross intake ratio for Standard I was 140 percent in 2016

but it is expected to fall to 115 percent by 2020 (Government of Tanzania, 2018). The government has estimated a need to construct at least 45,000 additional classrooms to accommodate increased enrollment among a growing population (Government of Tanzania, 2018). In Mara Region, additional challenges include long distances between schools and the households they serve, teacher shortages, ethnic differences between students and teachers which can result in violence against teachers and frequent transfers, and linguistic differences between students' native languages and the Swahili curriculum.

In the 2017-2018 Fiscal Year, the education sector accounted for 15 percent of the total national budget and 3.9 percent of GDP (UNICEF, 2018).⁷ Overall funding levels fall short of the 20 percent Global Partnership for Education target, which Tanzania had met in 2014/2015. Real funding levels in 2017/18 fell from the previous year as well. Recurrent spending comprises 74 percent of the budget while development investments make up its balance. A stark 47 percent of the development portion of the budget is consumed by higher education student loans.

Schools are primarily funded with capitation grants from the national government, with transfer amounts determined by enrollment numbers. But the Mara Regional Education Officer reported that primary schools across the region received less than 30 percent of projected grants in 2014 (Government of Tanzania, 2018). According to the World Bank, 32 percent of schools in Tanzania did not receive their full capitation grant in 2017, suggesting these issues have not been fully addressed (World Bank, 2017).

Amidst persistent resource constraints in both schools and the households they serve, Mara students drop out at alarming rates as their education progresses. Enrollment numbers fall precipitously at the point of transition to secondary school. UNICEF reports that in 2014, 46,121 Standard VII students registered for the Primary School Leaving Examination (Government of Tanzania, 2018). Of those, just 18,385 joined lower secondary school Form I – reflecting a staggering transitional dropout rate of 60 percent.

The Ministry of Education, Science, and Technology's Education Sector Development Plan 2016/17 – 2020/21 describes two key policy initiatives that distinguish it from previous five-year plans (Government of Tanzania, 2018). The first is a national commitment to providing 12 years of free and compulsory basic education to the entire population. The second is to expand technical and vocational education and training opportunities to support workforce development. The Plan goes on to articulate budgetary trade-offs associated with these two commitments, noting:

It becomes impossible to pursue simultaneously within this five-year period some other potential policy initiatives, such as expansion of the Higher Education sub-sector or **provision of centrally funded free meals in primary schools, without significantly increasing the proportion of** national resources allocated to the education sector. (Government of Tanzania, 2018)

⁷ All data in this paragraph come from the same source.

The plan considers demand as well as supply-side factors to achieve its objective. To increase demand, it calls for collaboration with Local Government Authorities (LGAs) “to implement cost-reduction/compensation/incentive mechanisms to attract pupils from families in greatest financial difficulty to school and retain them, including through cash-transfers, direct subsidies or school feeding programs.” (Government of Tanzania, 2018)

A set of six priority programs is laid out to address education access, quality, and management. The first of these pertains to access and equity in basic and secondary education. Within a set of initiatives focused on primary and lower secondary education, multiple strategies address household contributions in service of the objective that **“main education costs borne by households are reduced by half for all pupils, and further still for those from poor and disadvantaged backgrounds.”** (Government of Tanzania, 2018) The strategies involve increasing capitation grants to cover the cost of learning materials for all students and LGAs working with disadvantaged communities “to provide porridge and lunches for those in greatest need, as well as other indirect costs of schooling (e.g. sanitary pads for girls, uniforms, exercise books, transport costs).” (Government of Tanzania, 2018) They also entail the development of community-based solutions to provide learning materials and meals at school. The intended result of these strategies is to “reduce the burden of non-fee education costs on the poorest families, removing a disincentive to enrollment or retention.” (Government of Tanzania, 2018)

Health, Nutrition, and Sanitation

The burden of disease contributes to poor health and education outcomes for children across the country. One hundred percent of Tanzania’s mainland population is at risk of contracting malaria and the incidence of infection was estimated at 114 per 1,000 people in 2014. Malaria is a primary contributor to under-five mortality, which has fallen significantly over the past decade but remains at 67 per 1,000 live births (USAID, 2018).

The prevalence of stunting among children under five, that is the proportion of children whose height measures at or below two standard deviations below the (months of) age-specific mean for a healthy child, is a commonly utilized metric for food and nutritional insecurity. Stunting is associated with cognitive as well as physical deficits which persist throughout the life course, impacting not only educational outcomes but also those related to labor supply, family formation, and morbidity and mortality at older ages. In 2015, 34 percent of Tanzanian children were stunted. Twelve percent of children were severely stunted, meaning their height for age fell at or below 3 standard deviations below the mean of a healthy reference population (Government of Tanzania, 2016). While stunting levels remain alarmingly high, they have also fallen quite substantially in recent years, from 48 percent in 1999. Mara region fares somewhat better on this indicator than the country as a whole: 29 percent of Mara Region children under 5 were stunted in 2015 (Government of Tanzania, 2016).

Access to clean, potable water for both household and home agricultural use is highly limited, as is access to sanitation facilities. The UNICEF child poverty study found a 77 percent

prevalence of sanitation poverty, based on both unimproved or shared sanitation facilities and inadequate stool disposal, among 5 to 13-year-old children, the highest of any poverty domain for this age group. In rural areas, 81 percent of children in this age range had unimproved or shared sanitation (Government of Tanzania, 2018). Tanzania has experienced intermittent cholera outbreaks since at least 1974, when the first epidemic was reported (CDC, 2017). Ongoing infections have been recorded between 2015 and 2019, with a disproportionate share of cases in Mara Region (UN Children's Fund, 2019).

Project Design and Key Activities

Project Concern International (PCI) has continuously implemented McGovern-Dole Food for Education (FFE) programs across Mara Region since 2010. The first phase of the program was active from September 2010 through December 2014. Phase II, which ran from September 2013 through December 2016, continued work in Phase I schools while expanding into new communities. Phase III began in October 2016 and runs through December 2021. It has maintained a consistent presence in Phase I and II communities while increasing coverage further into remote areas.

FFE III activities are now active in 231 school communities⁸ and had directly benefitted 206,388 school children and adults by March 2019. PCI estimates that FFE III will reach a total of 442,743 beneficiaries (287,938 direct beneficiaries) over the life of the project. This target number includes over 265,000 students across 45 wards of Musoma, Bunda, and Butiama.

Table 2: Geographical Coverage of Food for Education Programming by Phase⁹

Phase	Schools by District	Total Schools
FFE I	Bunda (51), Butiama (9), Musoma (60)	120
FFE II	Bunda (61), Butiama (21), Musoma (73)	156 (originally 134 schools but 22 in Musoma were split during the project, 31 new schools)
FFE III	Bunda (101), Butiama (57), Musoma (73)	231 (75 new schools in Bunda and Butiama)

The multisectoral program includes a wide variety of complementary initiatives designed to improve student literacy. The primary intervention is to provide a once daily school meal five days per week in new schools. The meal is comprised of rice, pinto beans, and oil provided by the U.S. Department of Agriculture (USDA), supplemented by community contributions and crops grown on school grounds).

⁸ A further 16 schools in Butiama district have been targeted for a two-year program (2017-2019) in which food commodities, including maize, beans, and oil, are procured locally rather than imported from the U.S. The latter program, *Chakula Chetu* (Our Food), falls outside the scope of this evaluation.

⁹ District values from Baseline Report; Totals from Midline Statement of Work

PCI has supported the establishment and management of appropriate kitchen and storage facilities in all schools. Meals are prepared by cooks from the community whose compensation is furnished directly by parents. All school feeding activities are implemented in collaboration with school committees that play a key role in mobilizing parental contributions.

Over the life of FFE III, the frequency of meals provided by PCI is gradually reduced and communities are expected to take on a growing role to manage and sustain the program. For FFE I and II schools, the transition to community-led, USDA commodity-independent school feeding is accelerated relative to FFE III schools.

Table 3: School Feeding Commodities Transition Schedule¹⁰

		Year 1	Year 2	Year 3	Year 4	Year 5
Phase I and II Continuing Schools (156)	USDA-Provided Commodities	3 days/week	3	2	2	0
	Community/School-Provided Commodities	2	2	3	4	5
Phase III New Schools (75)	USDA-Provided Commodities	0	4	3	2	2
	Community/School-Provided Commodities	Stove and kitchen preparation	1	2	3	3

In Phase I and II schools, PCI now provides meals two days per week while the community is expected to provide inputs for meals on the other three days. In Phase III schools, PCI now provides meals three days per week while communities are expected to contribute food for two days. This school feeding transition plan will, however, need to be adjusted to account for the lack of USDA imported food.

In addition to school feeding, PCI supports an array of activities which work in concert to improve student attendance and attention, teacher and instructional quality, school infrastructure and sanitation, and the sustainability of all program components after the project closes in 2021. The technical design of the program is organized around the achievement of two strategic objectives and three key intermediate results:

Strategic Objective 1: Improved Literacy of School-Age Children

Intermediate Result 1.1: Improved Quality of Literacy Instruction

Intermediate Result 1.2: Improved Student Attentiveness

Intermediate Result 1.3: Improved Student Attendance

Strategic Objective 2: Increased Use of Health and Dietary Practices

¹⁰ Power Point presentation provided by Godfrey Matumu, April 2019

The program's full Results Framework, which includes additional Foundational Results is provided in Annex A.

- Education and Literacy

In the component of education and literacy, PCI aims to support different activities which promote literacy in new and interactive methods. It works to improve the quality of literacy instruction through the promotion of teacher attendance and recognition, provision of school supplies, most notably reference books. It further invests in the establishment of libraries, distribution and production of reading materials, training of teachers and school administrators and raising awareness on the importance of education

- Health, Nutrition, and Sanitation

PCI supports infrastructure for health, including access to safe water and sanitation facilities. It has constructed latrines for boys and girls in addition to changing rooms that provide adolescent girls with a safe and private place to manage menstrual hygiene. It supports teachers to improve the content and instructional quality of health and nutrition education as well as to promote behavior change through school health clubs. Health and WASH campaigns at the ward level simultaneously reach parents across these communities.

- Agriculture

The project encompasses two major agricultural activities: schools gardens and farmer field schools with demonstration plots. School gardens are managed by students, serving both as educational models and sources of vegetables for school meals. Students learn about nutritionally diverse crops and are encouraged to replicate gardening activities in their homes. Farmer field schools are organized around farmer groups, comprised of parents and other community members, that seek to improve cultivation practices and yields. Farmer groups are guided to contribute 20 percent of their harvests to the school feeding program and are able to sell the balance. PCI has also sought to link farmers with market service providers, including seed companies, while training them on business skills in addition to agronomic practices. Farmers also participate in the women's empowerment (WE) village saving and loan associations to make more efficient use of their collective financial resources (see below).

- Household Economics

PCI has led the establishment of women's empowerment (WE) groups to promote better household financial management, basic business skills and the access of women to decision making at the household and community levels. These groups focus on informal savings and lending activities which allow them to better smooth income and consumption throughout

the predictably uneven agricultural calendar. Members are better able to save earnings from higher income months corresponding with major harvest periods and to borrow funds year-round for either investment or consumption purposes. The groups are also intended to elevate the economic influence and decision-making power of women within their households. Still, WE groups are heterogeneous by design, with both female and male participants. As of June 2019, three quarters of WE group members were women (7,785 participants) and one quarter were men (2,544 participants).

Beyond these major program components, PCI has initiated early child development groups that train the parents to be their child's first teacher and better meet the care and developmental needs of young children, thus setting them up for future success.

Across each of these areas, PCI works closely with school leadership and local, district, and regional government to coordinate efforts so that activities may continue under local stewardship once the FFE III program closes in 2021. Expectations of community "graduation" are made clear from the outset. In the teacher training space, it collaborates with universities and teacher training colleges as well as local NGOs.

Baseline Survey Overview and Key Recommendations

A baseline evaluation for the project was conducted in mid-2017, with quantitative and qualitative data collection completed during the month of May. In addition to sampling schools from the three program districts of Musoma, Bunda, and Butiama, the baseline captured data from schools in neighboring Serengeti district to serve as a comparison group given Serengeti's similar economic and demographic profile. The comparison between program and non-program schools is useful to estimate programmatic contributions to changes in outcomes of interest, as explained in greater detail in Section 3.

The baseline evaluation elevated the following key recommendations:

- **Sustain school meals.** The analysis suggested that community engagement levels should be closely monitored to assess any increases in contribution levels. To support community ownership of the program, it recommended study tours between communities with the aim of exposing leaders to successful models. It further recommended an increased reliance upon both school gardens and demonstration plots for agricultural production to support school feeding.
- **Support/sustain school infrastructure.** The baseline emphasized the monitoring and maintenance of SWASH infrastructure by SWASH committees to ensure that infrastructure improvements, particularly latrines, remained in usable condition.
- **Focus on teacher training.** The report cited the pressure of increased enrollment on limited teacher numbers and capacity. It recommended partnership with state and non-

state actors to scale up support for teacher training programs and cascading training of trainer models.

- **Scale up WE groups.** The potential for women’s improved socio-economic status within the household was cited as an impetus to focus efforts on the expansion of Women’s Empowerment groups.
- **Sustain engagement with local government officials.** The analysis suggested that PCI Food for Education focal persons should engage at national, regional, district, ward, and village levels of government.
- **Reflect on program sustainability.** “At a macro perspective, the study team recommends PCI to reflect on the dialogue of sustainability within all components of the FFE program with improving government and community engagement as a key focus area for phase III of the FFE program. Increasing community outreach and behavior change communication can be targeted to improve community participation with the overall objective of long-term program sustainability” (Sambhodi Limited, 2017).¹¹

¹¹ Herein referred to as the baseline report.

Section 2: Midline Evaluation Purpose and Questions

The purpose of the midline evaluation is to improve PCI's understanding of its performance to date and to inform the leadership team in making improvements and adjustments in the final two years of programming to ensure optimal outcomes and sustainability. The evaluation's objectives are as follows:

- To assess the relevance of the project's technical strategy and implementation approach, including the validity of assumptions made during project design
- To measure progress on key indicators during the two years since the baseline survey was conducted and to identify both enablers and constraints impacting that progress
- To provide recommendations for mid-course corrections and performance strengthening
- To assess sustainability efforts to date
- To document lessons learned, challenges, and unanticipated outcomes
- To provide recommendations for areas of focus within the final evaluation, in terms of both substantive focus and evaluation process

The evaluation focuses on three key topics across both quantitative and qualitative analysis: program relevance, effectiveness, and sustainability.

Relevance

The analysis of relevance explores the alignment between program objectives, activities designed to support those objectives, and the needs of beneficiary communities. Questions in the realm of program relevance include the following:

- To what extent is the results framework and program approach valid for the local context?
- To what degree have the needs of the beneficiary population been included in the design of the program?
- Which program strategies implemented have been more effective and relevant in contributing to program results? Which were the least effective and least relevant?
- How has implementation and accomplishment of program activities been carried out at the technical /programmatic, organizational and contractual levels?
- To what extent did PCI address the baseline evaluation recommendations and lessons learned?
- Does the project design, structure, logic or management need to be modified to facilitate the achievement of desired outcomes and impact?

Effectiveness

- Has the program achieved the expected outcomes for literacy, teaching quality, student attention and attendance, and nutrition?

- To what degree have health, hygiene and nutrition behaviors and practices promoted by the project been adopted by the targeted beneficiaries?
- What evidence exists to demonstrate the relationship between project activities and outcomes related to literacy, health and nutrition?
- Has participation in WE groups led to increased school involvement and improved student outcomes (particularly for school attendance)?
- Has the project met targets for key performance indicators? If targets are not being achieved, why not?
- What is the level of satisfaction of the direct beneficiaries or participants?

Sustainability

- How well have the projects developed and maintained collaborative relationships with appropriate government agencies, local partners, USDA, and other key stakeholders?
- What efforts has PCI made to transition school feeding activities to local actors (local government, private sector, federal government agencies, parents)? How successful have those efforts been?
- Does the participation of mothers and/or other family members in the WE groups contribute to the sustainability of the program and program outcomes?
- What progress has the project made to ensure the sustainability of adequate human, financial, infrastructure, capital, and other resources?
- To what extent have partners, governments, and local communities internalized and implemented efforts to sustain school feeding activities, such as providing in-kind support for school feeding activities?
- What specific local, regional, governmental, and school stakeholder institutional capacities have been developed to foster host-country engagement and accountability. What additional capacities should be developed for which stakeholders?
- What elements or considerations should be incorporated into public policy in the local governments to sustain a school feeding program?

Literacy

Given the project's highest-level objective of improving literacy, special attention is paid to literacy interventions and outcomes. The following questions are considered:

- What changes in reading comprehension are observable since the baseline evaluation?
- Is there a correlation between literacy outcomes and characteristics of the student, teacher and school? For example, what is the relationship between the amount of literacy instruction in the classroom, quality of literacy instruction, class size, sex of the teacher, hunger level and nutrition of the student, and access to libraries and reading corners?
- Have teachers improved literacy instruction practices in the classroom as a result of participating in teacher training on literacy? For example, are there changes in

classroom organization, literacy instruction content, teaching methods, or use of different teaching materials?

Gender

Because many project activities take explicit account of gender roles, dynamics, and challenges in their design, the contribution of gender to desired outcomes, and the disaggregation of those outcomes by gender, is also explored.

- To what extent are project activities reaching female and male beneficiaries equally?
- To what extent is the project actively addressing gender-specific barriers and constraints to education and community engagement?
- What differential outcomes and impacts, if any, are project activities having on women and girls? Men and boys?
- Has household decision making changed among households with WE group members, particularly related to decisions around their children's education, and importance of girls attending school?

Section 3: Midline Evaluation Methodology and Limitations

The evaluation is based on a quasi-experimental, mixed-methods approach to assess both performance and impact indicators. A local firm, IPSOS Tanzania, led all quantitative and qualitative data collection activities. Trembley Consulting worked in collaboration with PCI and IPSOS Tanzania to provide input throughout that process and was responsible for quality control, data analysis, and report preparation. Before the commencement of data collection, the study received a research permit from the Government of Tanzania's National Bureau of Statistics.

Qualitative and quantitative data collection methods, sampling decisions, and tools development were conducted simultaneously to ensure the complementarity of quantitative and qualitative data. Each type of data also serves to verify and enrich the interpretation of the other type such that the joint analysis may comprehensively address key research questions.

All data collection tools were drafted between April 19 and May 2, 2019. The new set of tools was informed by baseline instruments, a revised evaluation scope which elevated distinct learning priorities related to program quality and sustainability, and input from PCI technical and monitoring and evaluation team leads. Midline data collection tools matched the baseline tools to the greatest extent possible to maximize comparability between studies. At the same time, all tools were reorganized and reduced in length. It should be noted that the deliverables included with the baseline report did not, in some cases, provide sufficient information as to their exact methods of data collection and analysis. Every attempt to remain consistent was made, but there are likely some instances where complete replication was not achieved.

The quantitative data were collected at 100 schools. The same sample of schools were visited at both baseline and midline. When selected at baseline, they were both determined to be largely representative geographically, and the 50 treatment schools were stratified to proportionally represent the three phases of the project. This has the benefit of ensuring comparability over time. For example, if a new set of schools had been chosen and scores improved dramatically, a skeptic could argue that the midline schools could just on average have better scores: by comparing in the same schools over time, we can have greater certainty that any improvements are not due to sampling artifacts.

The sample size was determined by calculating the minimum detectable improvement in reading scores for various sample sizes, as reading improvements is among the most important indicators and the most difficult to detect an effect in. The parameters of the sample size calculations are outlined in Annex H. The larger the sample, the greater the statistical power, and the ability to detect smaller changes. Because of the importance of literacy outcomes, the minimum sample size was calculated in terms of an improvement in the number of words per minute. For example, per the table below, for statistical tests to be able to detect an improvement the number of words a student can read per minute by as little as six words per minute, it would be necessary to interview an average 35.7 students per school (half in grade 2 and half in grade 4), or 3,568 students in total. If it was only necessary to detect an improvement of 11 words per minute or

more, it would only be necessary to assess 7.04 students per school, or 704 students. After discussing expected learning and value-for-money, it was agreed upon to set the minimum detectable effect of an improvement of 7.3 words per minute. As such, the minimum EGRA sample size was set at 1,144 per grade, or 2,288 in total. The other sample sizes were set to be equal to the minimums set at baseline, which allowed for substantial statistical power for the other variables of interest.

Table 4: Minimum Detectable Effects base on minimum sample sizes

Improved Words per Minute	Students per School	Total Sample	Design Effect	Cluster Size	Test + Attrition	Simple Random Sample Min.
5	62.3	6,232	2.2	14.3	713.3	642
6	35.7	3,568	1.8	9.9	496.7	447
7	22.9	2,288	1.6	7.3	366.7	330
8	15.8	1,576	1.4	5.6	280.0	252
9	11.4	1,144	1.3	4.4	220.0	198
10	8.8	884	1.2	3.6	180.0	162
11	7.0	704	1.2	3.0	150.0	135

Four new quantitative tools (for parents, head teachers, grade II/IV students, and grade VI/VII students), three revised quantitative tools (classroom observation, attendance record, and school sustainability assessment), and 10 new qualitative tools were finalized during the week of quantitative and qualitative enumerator training.

Quantitative Data

Table 5: Quantitative Data Sample Sizes

SURVEYS	EXPECTED SAMPLE	ACHIEVED SAMPLE
Grade II-EGRA Assessment	1,144	1,194
Grade IV-EGRA Assessment	1,144	1,194
Grade VI	200	200
Grade VII	200	200
Parents of students	800	800
Head teacher of primary schools	100	100
School sustainability and readiness observation	100	100
Classroom observation tool	100	100
Attendance	-	800

Early Grade Reading Assessment

At baseline, 600 grade 2 (300 treatment and 300 control) and 600 grade 4 students were assessed using the EGRA. At midline, it was agreed to base the sample size on the goal of being able to detect an improvement in reading seven words per minute. Doing so would require increasing the total number of students receiving the EGRA to 1,144 students. As was done at baseline, the EGRA sample was split equally between grades 2 and 4.

Table 6: EGRA Sample

		Baseline		Midline	
		Girls	Boys	Girls	Boys
Grade 2	Treat	150	150	300	300
	Control	150	150	300	294
Grade 4	Treat	150	150	300	300
	Control	150	150	300	294

The goal of 1,114 EGRA students was executed by IPSOS by visiting the 100 schools and assessing 24 students per school: six boys and six girls from each grade (given that it is impossible to collect 11.4 students per grade). IPSOS successfully exceeded the minimum sample size by completing 1,194 grade 2 and 4 assessments. They provided data for 24 students each in 99 schools, and in one comparison school provided data for 6 grade 2 and 6 grade 4 girls. The minor discrepancy in representativeness is only one half of one percentage point in the sample and has less than a one in one hundred thousand chance of affecting significance tests.

Attendance

Collecting data on attendance utilized a slightly different strategy. Students were selected among those who were attending school that day: however, to account for absent students in the attendance records, attendance was measured by the enumerator randomly selecting two students (one boy and one girl) from the class register in grades two, four, six, and seven, and recording of their attendance during the past six months. By selecting from the register, the sample included students who were absent that day (or are chronically absent).

Parent/Caretaker Survey

As with baseline, 800 parents were surveyed during data collection: two from each grade of interest. Parents and caretakers of all students were sent letters in the days before enumerators visited the schools to obtain consent for interviewing their children. Those randomly selected to be part of the parent/caretaker survey were offered a small remuneration for traveling to the school.

Classroom Observation, School Sustainability, and Head Teacher Survey

As with the baseline, each of the 100 schools took part in one class observation, the school sustainability assessment, and an interview with the available head teacher or most senior administrator present. The only difference between the baseline and midline approach was that while at baseline the grade observed in a school varied, at midline all observations were conducted of grade 2 classrooms to ensure comparability within the midline sample.

Table 7: Classroom Observation Sample

	Baseline	Midline
<i>Grade 1</i>	20	0
<i>Grade 2</i>	65	100
<i>Grade 3</i>	7	0
<i>Grade 4</i>	8	0

Overall, the resulting sample has a robust sample size to measure changes among key indicators. Given that the same schools are used, and that the midline and baseline almost perfectly match each other in representativeness, the comparability between the two sample groups is strong.

Qualitative Data

Qualitative data were collected from key stakeholder individuals and groups across program and comparison communities as well as from government leaders in the education and agricultural sectors. Focus Group Discussions (FGDs) and Key Informant Interviews (KIIs) were led by IPSOS Tanzania.

Relative to the baseline evaluation, the midline placed a different and somewhat heavier emphasis on qualitative data collection to triangulate quantitative findings as well as to cast a prospective view towards program close out and the sustainability of its efforts. The baseline undertook 20 Focus Group Discussions, including 10 with parents across program and comparison schools, 5 with WE groups, and 5 with farmer groups. It led 53 Key Informant Interviews (KII), including 10 with program school head teachers, 20 with grade 6 students from both program and comparison schools, 10 with local government representatives, 4 with District Education Officers, 3 with District Focal Persons, and 6 with PCI officials.

At midline, a total of 37 FGDs and 27 KIIs were convened across program and comparison areas. Table 8 provides details on the composition of FGD participants. Based on the input and priorities of PCI technical leads, the decision was made to focus FGD data collection efforts on program participants and their communities in order to maximize evaluation insights for intervention sustainability. Sample sizes were determined by PCI program management, technical, and monitoring and evaluation staff in consultation with IPSOS and Trembley Consulting.

Most groups were expected to have between seven and 10 participants. Targeted teacher group discussions were somewhat smaller, soliciting input from head teachers, store teachers, health teachers, garden teachers, and library teachers. FGD discussions were designed to run between 60 and 90 minutes and were audio recorded.

Table 8: Focus Group Discussion Subjects

Group Type	FFE Program Areas	Comparison Areas	Selection Criteria
Parents	6	4	3 high performing, 3 low; 2 high and 2 low in Serengeti
WE Groups	5		3 high performing, 2 low
Farmer Groups	6		3 high performing, 3 low
Grade 7 Boys	3	1	
Grade 7 Girls	3	1	
Header Teachers with Store, School Garden, Library, and Health Teachers	5	3	
Subtotal:	28	9	
Total FGDs:	37		

Table 8 furnishes a breakdown of KII respondent types. Interviews generally ran from 40 to 70 minutes and were also audio recorded. In alignment with the FGD approach, KII sampling and substance focused predominantly on the experiences of program participants and their communities. Capturing insights to inform sustainability efforts in the last two years of project implementation was consistently emphasized.

Table 9: Key Informant Interview Subjects

	FFE Program Areas	Comparison Areas	Selection Criteria
Literacy Teacher	6		4 high performing teachers, 2 low
Village Chairperson	3		Convenience
Agricultural Extension Officer (Village)	2		1 high performing officer, 1 low
Agricultural Extension Officer (Ward)	2		1 high performing officer, 1 low
Ward Education Officer	3		Convenience
FFE Focal Person - District Planning Officers	4		Full set
District Education Officer	4	1	Full set
Subtotal:	24	1	
Regional Agricultural Officer		1	Single position
Regional Education Officer		1	Single position
Total KIIs:	27		

FGDs and KIs were transcribed in English by IPSOS. The transcripts were then coded and analyzed thematically using the QDA Miner software package. Analysis was structurally informed by the program's Results Framework (theory of change) alongside stated midline evaluation objectives, with a strong emphasis on school feeding program sustainability after 2021.

Table 10: Thematic Coding Scheme for Qualitative Data

Category	Code
Relevance	Program Meets Community Needs
Relevance	Coordination with Government or Other Programs
Relevance	Beneficiary Satisfaction
Relevance	Beneficiary Dissatisfaction
Relevance	Comparison School Activities (FGDs only)
Effectiveness	Student Attendance
Effectiveness	Student Attention
Effectiveness	Instruction/Teacher Quality
Effectiveness	Literacy
Effectiveness	Infrastructure, Facilities, Supplies
Effectiveness	Parental Engagement
Effectiveness	School Gardens
Effectiveness	Farming Improvement
Effectiveness	Nutrition
Effectiveness	Health and Hygiene
Effectiveness	WE Groups
Effectiveness	Gender
Effectiveness	Behavior Change
Effectiveness	Policy Change
Sustainability	Contribution Sustainability
Sustainability	Collaborative Relationships
Sustainability	Government Sustainability
Sustainability	Policy
Sustainability	School Management/Teaching Capacity
Sustainability	Education Outcomes
Sustainability	Farmer Group Value Add
Sustainability	WE Group Value Add
Sustainability	Challenges
Sustainability	Feedback/Suggestions for PCI

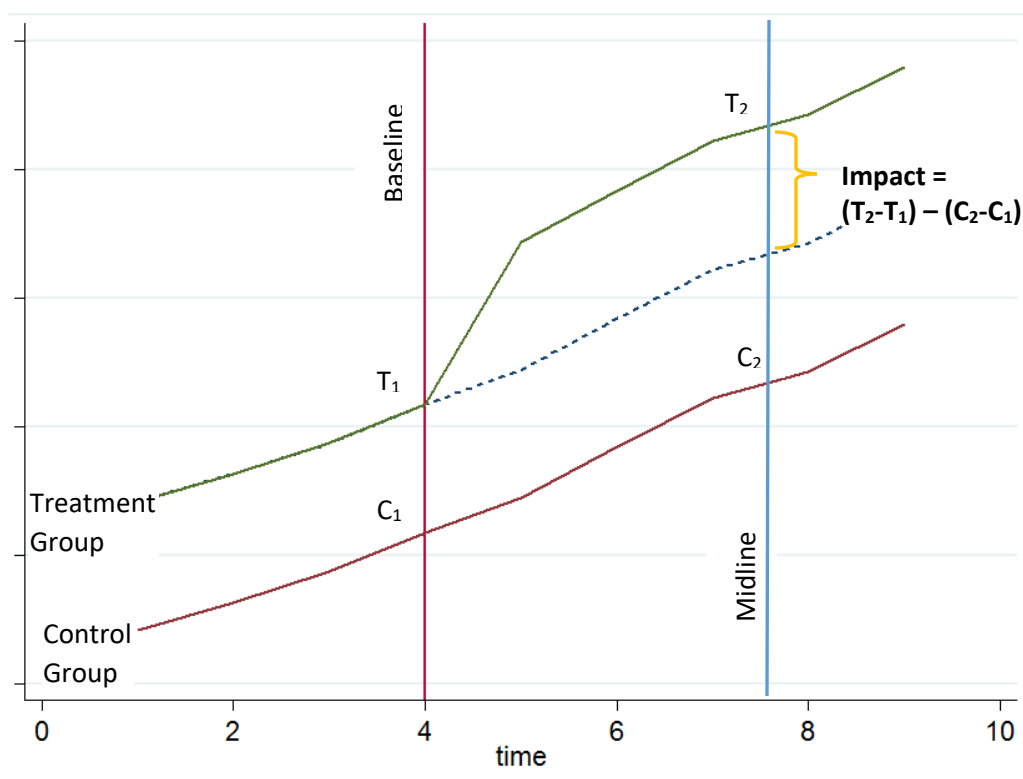
The analytical focus draws from grounded theory, an inductive and observation-led approach to inferring logical frameworks and causal pathways as articulated by research subjects. Annex C presents a detailed distribution of qualitative research extracts by code and interview type

across all transcripts. Quantitative and qualitative findings are jointly presented by topic in the following section.

Empirical Approach to Impact Identification

The quantitative component of the evaluation follows a quasi-experimental design methodology that allows for rigorous impact and attribution analysis and accounts, in part, for confounding factors with the inclusion of a comparison group. When baseline values between a treatment and comparison group are not identical but prevailing trends over time are presumed to be, a statistical approach commonly referred to “difference in difference” can be deployed to estimate program effect sizes.

Figure 2: Difference-in-Difference Impact Identification



Source: (Population Health Methods, 2019)

Figure 2 represents the assumption of identical time trends across two groups with parallel lines. Treatment and control groups record different starting values of a given variable but move in parallel over time until a program intervention is introduced, at baseline. As portrayed in this example, members of both groups exhibit gradually higher values, on average, for the variable in question over time, even in the absence of an intervention. An evaluator may

reasonably assume that literacy skills track this illustrative depiction, in that they compound over time and at a predictable rate across groups.

In the figure, an accelerating intervention is introduced at the time of a baseline survey to the treatment group only. From that point forward, the (green) treatment group follows a different trajectory for the outcome of interest. In the difference-in-difference approach, it is only the deviation from the parallel trend, rather than the full magnitude of change from baseline to midline or endline, that is attributed to program influence. That is, program impact is measured as the difference between the change, or difference, we would have expected given parallel trends and the change we actually observe.

To ensure a difference-in-difference model is valid, it is important to minimize sources of unexplained variation between time points. Using the same schools for the sample at baseline and midline ensures that time-invariant characteristics of schools will not affect the model.

Limitations of the Evaluation

This assumption of parallel trends also introduces limitations to the analysis. The comparison or counterfactual group, in this case Serengeti schools, is assumed to exhibit the “true” time trend in the outcome of interest. This assumption becomes problematic if any other factors may differentially impact literacy, or other key outcomes, in Serengeti schools, or in FFE schools. Any shifts in political leadership or economic conditions which affected Serengeti but not neighboring PCI schools, or vice versa, are unaccounted for by this approach.

Another such factor is exposure to the EQUIP or Right to Play programs. EQUIP is meant to be a regional program but 26 percent of both treatment and control schools reported receiving direct EQUIP support. One in ten control schools, and no program treatment schools, claimed support from Right to Play. Notably, head teachers of 14 percent of treatment schools also reported that they did not receive support from PCI, calling into question the reliability of information provided by respondents.

The 2017 baseline evaluation did find that the two sets of schools were highly comparable on many variables. One notable exception is that some comparison schools in Serengeti are on a reserve, in an area that does not allow farming and limits building construction. Schools serve the children of civil servants who are employed on the reserve. Teacher participants of FGDs also reported high attendance because children lived in very close proximity to the school. Thus, the social and economic conditions of these households do not appear to be comparable to the prevailing experience of treatment communities. The significant challenges associated with agricultural enterprises may only affect some comparison communities indirectly.

Serengeti focus group respondents also describe a school-based porridge feeding program that is sponsored by parents in the community. Students eat porridge at school in the morning. Flour, sugar, and money to pay cooks are provided by parents. Interview subjects also

referenced an activity called UWAWA, a component of the EQUIP project that focuses on the needs of girls in school. The UWAWA union has purchased materials for sanitary pads and sponsored trainings on menstrual hygiene in schools. As such, the true impacts of PCI activities cannot be well measured given that the comparison group has received similar interventions.

In terms of the indicators themselves, several were either calculated differently or based on questions that were asked differently between survey waves. For example, Minimum Acceptable Diet included different food category responses in the baseline survey and excluded fish which is a primary source of protein. Trembley Consulting's re-analysis of baseline data failed in multiple cases to recreate reported baseline values, or to come even within close proximity of reported values.

In the qualitative work, all interviews and discussions were held in Swahili, but the analysis was conducted in English. The English language transcripts were mostly intelligible but not high quality, and it is certain that elements of meaning and nuance were lost in translation. There is also an inherent incentive for subjects to report satisfaction and positive program outcomes in any context where data collection informs resource allocation, including this one.

Low parental engagement was cited as a challenge by PCI staff and by a wide array of KII and FGD respondents. A lack of parental awareness pertaining to some educational details poses a limitation as it makes these respondents unreliable reporters of school conditions. For example, in one Serengeti community, a focus group discussion of teachers described the presence and procedural use of the school library. A group of parents from the same community expressed discontent that the school did not have a library and suggested that adding one would be beneficial to learning outcomes.

To ensure a difference-in-difference model is valid, it is important to ensure both the data collected and the source of the data are as similar as possible at baseline, midline, and endline. Using the same schools for the sample at baseline and midline ensures that time-invariant characteristics of schools will not affect the model.

Even though the midline instruments were based on the baseline instruments, there was minimal explanation as to how each indicator was defined and calculated using that data. Attempts to reproduce the same indicators using the baseline dataset resulted in very different values. To ensure the baseline and midline indicators were comparable, standard formulas were developed to calculate each indicator, and then used to calculate both the midline indicator values and the baseline indicator values. However, in some cases, it was not possible to recalculate the data. In some cases, there was no provided data: in those cases, values from the baseline were used, and noted in the report with a †. For example, the indicators on attendance, improved water sources, and sanitary facilities has no known (or provided) data source. In some cases (such as the data on good hygiene practices), it was not possible to calculate sensible outputs from the data: while the cause is unclear, it may be due to uncertain or undocumented practices on coding missing and skip logic. When values were obtained from

the original baseline report, they were not sex-disaggregated. To prevent uncertainty about calculation of values at the endline, this report is submitted with (1) datasets that have merged all of the baseline and endline data, (2) more thorough descriptions of how indicators are calculated in the text, and (3) the syntax code used to construct the datasets and calculate all of the indicators.

Section 4: Key Performance Findings: Relevance, Implementation, and Beneficiary Satisfaction

A general consensus emerged across stakeholders that the array of FFE programming works in concert to improve the school environment. Diverse voices praised PCI for its engagement with local leadership and responsiveness to challenges articulated by community members. Overall satisfaction with the program was consistently reported. One literacy teacher in Musoma characterized the program as follows:

PCI in my society has been involved in the improvement of infrastructure mostly in water and the sanitary system. In doing so they have improved our school so much. PCI has also supplied our schools with books, for instance previously one book was shared among 20 students while now only two students share a book. PCI has also dealt with improving children's health through food provision at school. Another thing that PCI has done is educate us on gardening and on how to prevent and combat pests.

One Village Chairperson reflected, “there is no activity that PCI is doing that is not beneficial to the school or the community, it only has benefits and we are giving it accolades.” Similar sentiments were echoed by stakeholders from head teachers, to village agricultural officers, to ward education officers, to parents and students themselves.

A District Education Officer explained the benefits of PCI’s holistic, multi-sectoral approach to improved educational outcomes. He or she said, “the main purpose of PCI is to remove the obstacles that prevent a student from performing well in his or her studies. Therefore, all that PCI does, diet, building toilets, water tanks and health both aim that at the end of the day to raise academic standards.”

Community leaders and members predominantly agreed that program activities addressed the most pressing local needs while also citing a daunting array of challenges in the education sector. When asked if the project addressed salient educational issues in his or her ward, a Ward Education Officer responded:

The needs of education to the students are many. I think the project has worked on some of the needs which are very important though there a lot of needs which the organization cannot solve all of them because of the economic situation or the according to the priorities of the organization, but they work on the needs of the education. For example, students getting food while at school so the student can study properly and have a peace of mind rather than when the study while being hungry. Also, you cannot study while hungry but also the PCI does not provide food only, they have given training, books to assist us... so they work on some of the needs of education on a ward level. So, we, as spectators of education, we can ask for a chance if the time allows we talk about the infrastructures. You take a look at the teacher’s house because all these are challenges which at times the government cannot solve all the challenges without having other assistance.

Communications and Coordination

Regional, district, and ward officials provided unanimously positive feedback on PCI’s efforts to coordinate and collaborate with government leaders in their work. One District Planning Officer

referenced the lack of public resources to undertake essential activities that PCI has spearheaded:

The government appreciates all activities conducted by PCI like building toilets, providing books, providing water tanks, providing food in school and providing agricultural equipment. All these activities should be conducted by the government, but we cannot manage to conduct all of them, so we appreciate PCI for taking this up.

When asked about his or her perceptions of PCI organizational structures, one District Planning Officer said that “each department has its manager and staff who supervise their District in different activities. Therefore, PCI is well organized because they have their departments which work parallel with us to help meet all of our needs.” Ward and village-level respondents similarly expressed satisfaction with program efforts to coordinate and collaborate with local institutions and individuals.

Teacher Training and Literacy Activities

Teachers interviewed both individual and in focus group discussions valued the trainings provided to improve pedagogical practices. They reported that training content was relevant to their classroom experiences and that they were able to implement new practices in their day to day work. One literacy teacher explained:

The training was really helpful and beneficial, with lots of quality because most the things that we learnt in the trainings like the creating of teaching aids, teaching through sports and games, how to create gender balance while teaching and many more that I was able to gain while I was in the training.

Responses varied on how well those trainings, and the improved practices they encouraged, were transmitted to other teachers who did not attend. A Ward Education Officer in Musoma also expressed dissatisfaction with a focus on training lower primary teachers at the expense of teachers of class five, six, and seven. “This has negative impact since students academically grow well from lower classes but when they reach upper classes their performance deteriorates.” Another Ward Education Officer in Bunda felt that teacher trainings should not be scheduled for times when school is in session as extended teacher absence left students without instruction. PCI reports that it has not held trainings or meetings with teachers during school hours since 2017.

All stakeholders were pleased with the contribution of teaching tools and materials. Multiple respondents reported that supplied materials are useful and of high quality. Supplemental materials allow teachers to more effectively engage students while saving instruction time. One teacher in Butiama explained:

Comparing between the previous time and now, although that the books are not enough to fit one book for each student but at least three for one book. This has made our teaching easy as we don't have to write everything on the board and the students can simply do their exercises on the instructions given to them in reference to the books.

A teacher in Musoma described the utility of lesson books to improve learning outcomes. Better parental engagement in the learning process, leading to more student contact time with lesson content, was a proximal benefit of additional materials:

The students see they have benefits to them because the books are the main resource to the student so as he can know how to read and write. Because without having those materials it is difficult even to teach a student. If a student has these books, an adult can assist him to know how to read and write rather than leaving that duty to be the teacher's only.

Another teacher in Musoma highlighted the motivating influence that adequate materials had on teachers:

The PCI has been able to reach so it will be better if the project continues, so we teachers we have reduction of work because there is availability of books and we get the courage to teach, I think without the PCI we would not be here because it has encouraged us teachers to work more.

A teacher in Bunda requested additional materials, explaining that “the tools are of high quality, but the number should be increased so that they can match the number of students that we have.”

FGD respondents discussed the advantages of students borrowing library books, thereby both raising their appetite for improved reading while also engaging parents:

So, for example, we have a library leadership where there is the head teacher, a library teacher, there are two students, male and female members, parents, and partners with the parents ... they have been excited to come and borrow books even for their children at home, so we see that there is a huge benefit on the current maintenance compared to before the training.

A teacher in Musoma underscored that the combination of literacy interventions was driving results in her classroom:

On my side I see the results are straightforward, example in the beginning of the year you may have five students who are unable to read or write but after some time you may find they can read and write, also through books which were brought by the PCI it has helped students to read different stories and help them learn more.

School Feeding Activities

All student FGD respondents said that they liked the food provided by PCI at school and that it was of high quality. Parents shared this positive reception of school meals and felt that the program was well managed:

It is going on well because the moment the food is over another one is brought.

All respondents linked school meals with sharp increases in attendance. A District Education Officer reported, “the program is very successful as children love the food...The attendance is rocketed on the days when food is provided.” A farmer in Bunda added:

The project was a success because the students' attendance is improving day by day since everything is available at school and the parents give us as much as they can to support the project.

The topic of parental contributions to school feeding is taken up in detail in Section 7 of this report.

Students link school meals with their attention levels in class. A 7th grade girl in Bunda explained, “there are times you may find a student has not eaten at home and school has not cooked so when taught the brain is not at class but he or she is thinking of somewhere else.” Parents agreed. One Butiama parent said, “our children get food at school which helps their mindset and concentrate on what they are taught” and another added that “the students have grown health and strong.”

The sentiment among teachers in Butiama was similar:

Before we had these trainings and the food program at school the learning of the students was uneasy. For instance, some of them were not attending school, others falling asleep in the class and their attention was generally down. Because of the hard condition of life among most Tanzanians, the inception of this project has been important in helping the attendance of the students.

One point of dissatisfaction pertained to the portions of food served to students. A Village Chairperson in Butiama reported:

The challenge that I hear the community complaining about is that the students do not get satisfied with the food; the portion they get from PCI is too small. So as leaders we try to make them aware that the food is just to sustain them while in school and not to get them satisfied.

Another Village Chairperson in Bunda explained that provided food was insufficient both in terms of the number of days meals were served and also in terms of portions:

To say the truth, it is not enough because if it was enough, we wouldn't have needed the parents and the community to contribute, even what PCI brings is not enough. There is a time I was here, and I saw the look on the children's faces when receiving the food and it was clear that it's not enough to satisfy their hunger.

A focus group discussion of 7th grade boys in Musoma yielded an exchange which raised both the issue of portion adequacy and also of fairness:

- Interviewer: Have you once famished at school because you had taken insufficient food?
- All: Yes.
- Interviewer: Why?
- Respondent 1: Because right here they serve a little amount of food that you don't get satisfied so when in class you find yourself just falling asleep.
- Respondent 2: But the ones serving take a big share.
- Respondent 3: They also give plenty to their relatives while we have all contributed.

Later in the discussion, one participant expressed frustration that “you can find the ones who haven't contributed are the ones serving so they become satisfied.” Concerns around serving equity and the appropriate use of USDA commodities are discussed further in Section 7.

Finally, the complex dynamics of community contributions emerged as a sensitive topic for multiple respondents. A Ward Education Officer offered some insights into the intricacies of parental engagement with the program:

Sometimes there challenges because sometimes you may accept [the program] but once you implement it is when the challenges begin. For example, there is a food project for school for the students to have their meal. They have to fetch water because the organization cannot bring water. Also, for the students to eat they have to fetch firewood, and also bring salt. At first the community accepted this. Sometimes when they see their children carrying firewood the forget they are the ones who accepted the project in the first place.

Infrastructure, Sanitation, and Health

Beneficiaries expressed particular appreciation for SWASH activities, including latrine construction or improvement, handwashing stations, and the provision of water tanks. Many profiles of stakeholders, from parents to government officials, were quick to identify real value in the physical assets constructed or provided by this sector of programming. For example, a Butiama Village Chairperson reported approval of improved food management and storage facilities in the local school.

Construction of the kitchen, the food store, I can also see that we have water tanks so I can say that the project is a success and it has benefited us, we are happy about it.

A Focus Group Discussion of teachers raised satisfaction with both infrastructure improvements and provided equipment but also with the associated management practices introduced:

This project is doing well because for everything they provide they also give instructions, for instance the health equipment, the teachers concerned with health attended a seminar, so the organization is doing well.

A number of students and teachers referenced reduction in vectors of contamination and disease transmission through the uptake of regular handwashing. A District Planning Officer linked sanitation infrastructure upgrades to improved student health outcomes:

There are different changes because since PCI introduced this project to the community, they have been aware on how to contribute food to schools and pupils' performance has improved. On the same, health has also improved since were built toilets and training teachers on the use of toilets.

A head teacher reported positive outcomes as a result of sanitation activities, particularly noting shifts in handwashing behavior:

They all follow the rules that they have to wash their hands and they must so this has improved through reducing the occurrence of diseases like diarrhea, stomachache, and eliminating absenteeism.

A group of 7th grade girls in Bunda expressed satisfaction with latrines and handwashing stations and also with the level of privacy they offered:

Interviewer: Are the toilets in good condition for use?

Respondent 1: Yes

Interviewer: How?

Respondent 1: Water, soap and other stuffs are available.

Interviewer: Where do you fetch water from?

Respondent 1: We fetch water from the well

Interviewer: I am talking about the privacy of your toilets, is it easy to see a person entering?

Respondent 2: No, because our toilets have doors and gates, so it is not easy to be seen when getting in.

Interviewer: Another one, how is the condition of your toilets?

Respondent 3: It's stable

Interviewer: Why do you say it's stable?

Respondent 3: It's stable because the toilet was built recent also it has soap and clean water.

In other communities, parents expressed particular gratitude for water tanks. One said, "we thank PCI for bringing food that our children are studying well and for putting up tanks, so our children can get clean water."

The FFE III baseline evaluation emphasized the need to monitor maintenance of improved latrines to ensure they remain in usable condition, and this issue did surface in several midline discussions. Students in some schools said that toilets become dirty quickly due to lack of regular cleaning, which was sometimes mediated by a lack of access to water.

Improved Agricultural Practices

Farmer groups were successful in reaching a broad cross-section of community members. Over one fifth of parents interviewed in the midline survey indicated that they were current (12%) or past (9%) farmer group members.

Farmers who participated in focus group discussions were generally satisfied with the quality of trainings and activities. Many described improvements in practices, crop quality, and yields. A farmer in Bunda said, "it helps a lot. The reasons for learning through a lot of training is to help us to know how to use fertilizer, to protect crops and how to plant seeds."

A group of farmers in Butiama shared both shifts in knowledge and techniques and also the benefits of shared demonstration plots and the proceeds derived from them:

Respondent 1: Another benefit is that it helps farmers to know the advantages and disadvantages of using measurement in agriculture.

Respondent 2: Appropriate spacing plant population results in adequate plant population

Respondent 3: Also you can count your crops on farm before you harvest because the plants can be counted

Respondent 4: Before demonstration plot we only weed once but after we did our practical we saw how health the crops grew thus we all weed twice.

Respondent 5: I have experienced a benefit that is beyond agriculture and it is through the savings and shares that we buy. During dividends I was able to spend my share on building my house.

In Bunda, another group of farmers offered similar assessments, noting that other community members sought to learn from their positive experiences:

- Respondent 1: Before the introduction of PCI project we were using local farming methods. But now we are aware of the farming development and also, we were able to tell other people from other the villages about it.
- Respondent 2: Demonstration plots helped us to produce vegetables for our children at school and also, we have seeds of which others come to take for further cultivation for plots plantation.
- Respondent 3: Demonstration plots added something in farming activities because before we used local farming method but now days we are modernized, and other people want to learn from us after seeing the harvests

Agricultural extension officers also expanded their expertise through program activities. One extension officer in Butiama explained:

I thank it personally as though I am an expert yet there is so much that I have learnt from them especially on the soil conservation cultivation that involves the use of mulching crops like *Ngwara*. This cultivation is very advantageous. [...] Even though the scope is so wide the project has put in place a good plan of education particularly on pest control and issuing of seeds to the groups and schools where they don't prefer the use of chemical rather biological inputs.

A Musoma extension officer talked about shifts in crops grown by local groups, describing an increase in horticultural crops in particular. In other areas, program activities focused on maize cultivation while encouraging farmer groups to contribute other types of seeds (e.g. cassava, sorghum, or sweet potato). It is noteworthy in that maize is less drought-tolerant than cassava. Some farmers observed that the soil does not sustain maize well.¹² One explained, “We still cultivate maize in our demonstration plots because we have not been able to get other seeds, but we are planning to cultivate potatoes and cassava.”¹³

WE Groups

WE groups received some of the most consistently positive feedback from a wide variety of stakeholders. One District Planning Officer offered high praise:

PCI is also working with the community especially in women empowerment and we have seen positive impacts. There are people who had nothing but they joined and now they have some savings and credit groups and they got some money so we can see the physical verification; someone would tell you that they have constructed their house, or they have started their own business through the groups. So the results are very good.

Another District Planning Officer emphasized the intervention’s reach:

These groups are more than 160 in the District, right now these groups have succeeded because they help in the project towards providing food for education, and students with special needs. Also, the groups within the District have united and managed to build classes and one teacher’s office.

A third District Planning Officer shared this positive characterization, reflecting that “the presence of WE has improved the economy of women who have even attained a voice in

¹² Farmer focus group discussion, unidentified location due to source transcript file mislabeling

¹³ Farmer focus group discussion, unidentified location due to source transcript file mislabeling

decision making.” In 33 of the 50 treatment schools visited, there was at least one current WE member. Of the 400 parents of children attend beneficiary schools, 22.8 percent reported being current or past members of a WE groups. Curiously, however, 16.5 percent of parents in the control region reported being members of a WE group, as well. It is unclear if these parents misunderstood the question or are members of similar empowerment groups.

WE group members offered their own assessment of the intervention’s alignment with community needs. Members frequently cited the seasonality and volatility of agricultural livelihoods and the contribution of WE groups in smoothing their income over the calendar year. A Musoma FGD participant summarized:

In our environment, the main source of income is farming so income grows depending on the weather, if the weather is good then the income also increases. We are also grateful for these groups that were started by PCI, many women have found capital to start small businesses and it has helped in raising income in the family and the community at large.

One group member referenced increased financial autonomy within her household:

Before I joined the group, I financially depended on my husband but as soon as I joined the group, I now get my money through my hard work, so I make the money decisions by myself.

In addition to access to capital for small investments, participants used loans to meet basic needs, including school supplies for their children, and to cover shocks such as illness or deaths in the family.

Members expressed a clear sense of ownership over group management and processes. When asked if lending terms were fair, multiple respondents explained that they are all involved in determining their group’s “constitution.”

One prevalent challenge across WE groups was the correlated timing of peak loan demand, during the dry season. This led to an inability to borrow at times of greatest need due to a lack of funds in the cash box.

The scope of specific WE group activities and operations, impacts, and challenges is discussed further in Section 7.

Section 5: Key Impact Findings Related to Strategic Objective 1: Improved Literacy of School-Age Children

Sections 5 and 6 of this report rely heavily on the difference in difference approach to impact assessment described in Section 3. Interpretation of impact findings focuses on “second differences”- the differences between (1) observed changes in the treatment group between baseline and midline and (2) observed changes in the control group between baseline and midline. Observed changes in the control group are presumed to represent the time trends that would have been observed in the treatment group absent PCI interventions.

School Enrollment

Between the baseline and midline evaluations, enrolment increased substantially for both boys and girls in all schools. Enrollment increased by 16.5 percent in treatment schools and 32.1 percent in comparison schools, but the sample size of 100 schools does not provide adequate power to detect statistical significance. Despite the substantial uptick in enrollment, the Mara Regional Education Officer reported that 6,000 students remained “outside the system.” In the 100 grade 2 classes observed, on average there were 99 and 123 students registered in treatment and control schools, and 70 and 85 students in attendance.

Table 11: Average Number of Students Enrolled in Schools¹⁴

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>All Students¹⁵</i>	Treat	623	726	103	-113
	Control	676	893	217	
<i>Girls</i>	Treat	310	366	56	-57
	Control	325	439	114	
<i>Boys</i>	Treat	317	401	84	-27
	Control	356	468	112	

Parents from both treatment and control schools were surveyed about their perceptions of the benefits of schooling. The overwhelming majority of respondents said that educational benefits were equal for boys and girls, but there a few candid exceptions to that pattern were salient. One Musoma parent told the FGD interviewer:

From my view, women I do not care much due to the way that they are weak as we see at home. So I think when ... some are boys and some are girls, I like to educate boys, because the male children have confidence of what they do, but girls are the ones ... approached and manipulated and can enter into a loss, you take them to school but they bring loss to me. I prefer boys over girls.¹⁶

¹⁴ This indicator averages measurement of being off task at the beginning, middle, and end of a lesson.

¹⁵ Values for Boys and Girls do not sum to All Students because numbers faithfully represent school records, which contain some errors.

¹⁶ Quotes from interviews in report are translations provided by IPSOS. To represent as close to the original versions, we edited them as little as possible. Original: From my view, to women I do not care much due to the way

Discussion transcripts suggest that it was more comfortable for parents to discuss general cultural barriers to girls' education than to speak candidly about their personal opinions. A Serengeti parent reported:

Challenges come when the parents have a disagreement; the father wants the child to get married so that they can get cattle and the mother wants the child to study.

A District Education Officer offered the same general assessment of disproportionate challenges for girls:

Here girls are not given a priority; meaning, it is not very important for a girl to go to school. A girl can also be taken out of school to be married off while others lose their lives due to circumcision but generally if both are given equal opportunities, they can perform.

Despite these prevailing attitudes, girls are at least enrolled in primary school at approximately the same rate as boys, and gendered outcomes may instead manifest in attendance, parental support, or advancement to secondary school.

A WE group member in Butiama explained poverty-driven challenges that undermine both girls' and boys' prospects for parental support in their education:

Of course, we face many difficulties taking children to school, as for me I did not go to school so there is no way I can take my child to school because there is no money, the other thing is poverty sometimes you find you cannot even buy shoes and uniform for a child unless you work very hard like selling charcoal in order to take a child to school. Children like to go to school but it is very expensive.

When asked directly to identify three benefits of completing primary school, parents post significant gains between baseline and midline across both treatment and control groups.

Table 12: Percent of Parents with Knowledge of Primary School Educational Benefits

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>3 Benefits of Completing Primary School, All Caregivers</i>	Treat	47%	96%	49%	-5.3%
	Control	42%	96.3%	54.3%	

The underlying question for this indicator was identical in both surveys: "What do you think are the main benefits from ensuring that your child goes to school?" Prepopulated responses were the following:

Table 13: Main Benefits of Child Going to School

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>Reducing Poverty</i>	Treat	43.3%	42.3%	-1.0%	-1.0%
	Control	51.3%	51.3%	0.0%	

that they weak as we see at home. So, I think when have children some a boy and some are girls, I like to educate boys because, the male children have confidence of what they do but girls are one to be approached and manipulated can enter into loss, you take them to school but they cause bring loss to me I prefer boys children more than girls.

<i>Gaining Life Skills</i>	Treat	28.5%	56.3%	27.8%	-4.3%
	Control	24.8%	56.8%	32.0%	
<i>Gain Knowledge, Skills, and ethics</i>	Treat	34.5%	59.3%	24.8%	-3.5%
	Control	22.3%	50.5%	28.3%	
<i>Ability to express themselves</i>	Treat	28.3%	25.5%	-2.8%	0.0%
	Control	26.8%	24.0%	-2.8%	
<i>Self-awareness and self-reliance</i>	Treat	36.3%	45.8%	9.5%	-10.3%*
	Control	33.8%	53.5%	19.8%	
<i>Building capacity for further development</i>	Treat	26.8%	38.3%	11.5%	2.5%
	Control	30.5%	39.5%	9.0%	
<i>Increasing creativity and curiosity</i>	Treat	14.0%	19.3%	5.3%	-3.0%
	Control	11.3%	19.5%	8.3%	
<i>Get school meals</i>	Treat	4.5%	6.0%	1.5%	3.8%
	Control	3.8%	1.5%	-2.3%	
<i>Increased confidence</i>	Treat	5.5%	11.0%	-5.5%	-0.8%
	Control	4.8%	11.0%	-4.8%	

There is a misalignment between question wording and indicator definition in that parents are not explicitly asked to provide three distinct benefits. However, in both survey instruments, enumerators were instructed to probe until at least three responses were given. The ability of 96 percent of parents at midline, in both treatment and control groups, to provide three benefits of education is likely an artefact of that instruction, more effectively operationalized. It also reflects the tautological nature of a low-quality indicator: to measure parents who know three benefits of education, parents were instructed to provide three benefits of education.

Student Attendance

The midline quantitative instruments captured two measures of school attendance. The first indicator is based on parental reporting of their own children's absences, for any reason, over a two-week reference period. Parents are then asked for the cause of absence and only those absences reported as due to illness are recorded here. The unit of observation in this case is the parent. The second is based on classroom registers in observed classrooms, where the unit of observation is the student. This second indicator is defined to capture 80 percent or higher rates of attendance over a six-month reference period.

Table 14: Illness-Related Absence and Regular Attendance (percent of parents or students)

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>Absent due to Illness in past 2 weeks (Parent Reporting)</i>	Treat	4.6%†	6.3%	1.7%	-3.1%
	Control	4.2%†	9.0%	4.8%	
<i>Absent due to Illness in past 2 weeks, Girls (Parent Reporting)</i>	Treat	†	5.1%		
	Control	†	9.8%		

<i>Absent due to Illness in past 2 weeks, Boys (Parent Reporting)</i>	Treat	†	10.5%		
	Control	†	6.5%		
<i>Regularly Attending School (80%) in past 6 months (Classroom Register)</i>	Treat	82.9%†	72.0%	-10.9%	6.5%
	Control	81.1%†	63.7%	-17.4%	
<i>Regularly Attending School (80%) in past 6 months, Girls (Classroom Register)</i>	Treat	†	71.6%		
	Control	†	73.0%		
<i>Regularly Attending School (80%) in past 6 months, Boys (Classroom Register)</i>	Treat	†	71.0%		
	Control	†	55.8% ¹⁷		

Note: † denotes baseline values obtained from baseline report, and may have been calculated differently than midline values. In such cases, values were not sex-disaggregated in the baseline report. In this case, the original data source was not available.

Note that the desired direction of these two indicators are in opposite directions. For the first, school absence due to illness as reported by parents, a statistically significant negative second difference connotes positive program impact in that treatment school students are reported absent less frequently than control school students. For the second indicator, a positive second difference reflects positive programmatic influence in that treatment school students are more likely to attend school regularly as compared to control school students, though in both cases, attendance decreased. Unfortunately, the original data from the baseline was not available, so recalculation was not possible.

The challenges of chronic absenteeism and truancy surfaced frequently in the qualitative data. Irregular student attendance was widely cited as a confounder of educational outcomes by stakeholders in both PCI program treatment and Serengeti comparison schools. Respondents identified truancy as an area where schools and local government collaborated well. Lists of truants are provided to village councils which then contact parents. Some KII respondents in Musoma explained that boys may be disproportionately truant, moving around in pursuit of fishing activity. Others disagreed, adding that there is fishing economy work suitable for girls as well, and that patterns of absenteeism may be similar for girls and boys.

Students themselves report both household obligations, including agricultural labor supply and childcare, and also scarce household resources for school materials as primary drivers of absenteeism. A group of 7th grade boys in Butiama answered as follows when asked about the main reasons that they or classmates miss class:

Respondent 1: If it's due to sickness, I come to school and ask for the teacher's permission.

Respondent 2: When it happens, you have a funeral at home.

Respondent 3: If you had an accident.

¹⁷ Regular attendance of control group boys over the previous six months appears to be particularly low; the average attendance rate was 77% for this group, below the 80% threshold.

Respondent 4: If you miss school necessities like school uniform, socks, shoes, notebooks/ books, pens and pencils.

Interviewer: Is there any other reason for one to miss school?

Respondent 1: It happens when a child wakes up early to do all home chores until they are late and can't make it to school on time.

Pressed further on household activities that interfere with classmates' attendance, the group explained other drivers of absence caused by parental directives:

Respondent 1: They go to the farm.

Respondent 2: Parents may tell you to go and harvest cassava in the evening, but if you resist in the morning he/ she will tell you don't go to school until you have uprooting the cassava first.

Respondent 3: They tell you to go and grazing cattle.

Respondent 4: Other parents just stops you from going to school and make you stay at home.

Interviewer: Then what do you stay doing back at home?

Respondent 4: Just small house activities like cleaning the house, washing dishes, cooking, fetching water and stuffs like that.

When parents were asked about their children's recent absences in the quantitative survey, only 25 of 400 program (treatment) school parents reported an absence in the last two weeks. Of those, nearly three quarters reported that the absence was due to illness (among control group respondents, 31 of 36 parents with recent child absences attributed them to illness). Four of the 25 said their child was absent due to poor teaching quality while three reported a lack of school books or uniforms. No parents with recent child absences in either the treatment or control group said that children stayed home to undertake household chores.

However, when the full set of parents were asked about the "main reasons" that children in their child's class miss school, just under half said that illness was among the main reasons, about a third said household chores, the same share cited a lack of books or uniforms, and a quarter raised poor teaching quality or teacher absenteeism. About one in five FFE school parents and a third of Serengeti parents cited the distance to school.

A Ward Education Officer assessed that household and on-farm labor supply was the primary driver of absenteeism:

Yes, parents are the source of the children dodging classes because during agricultural season they want their children to give them a hand on the farming activities and fishing activities.

Most adult FGD respondents across discussion group types reported that agriculture comprises 80 to 100 percent of household income. The labor supply of children may be critical to sustain their efforts. One group of teachers observed:

Challenges are there because some parents see their children as sources of wealth, so you find that studies are going on at school and some students are absent because their parents have retained them to do work that creates wealth and income for them.

A parent in Bunda described the tension in his or her household between sending kids to school and providing for the family:

On my side I think it is a big challenge because if there is hunger in at home as a parent or a guardian, you end up thinking of the ways to make money and buy food and you do not think about taking children to school. So I think hunger is a challenge.

A Bunda WE group member lamented the constant inability to meet basic needs:

It's not possible for us to save money for later use because the needs are so many compared to the income we have and also there are emergencies which occur making us to use the money. In short, the money we get does not meet our demand.

A Serengeti parent described similar budget constraints that undercut the ability to buy school necessities:

It is expensive because my life depends on the rain so that I can be able to get some farm produce, so buying the sweaters and the uniforms is very expensive for me.

A WE group member in Musoma shared another experience linking the precariousness of agricultural livelihoods to the ability to provide educational inputs:

Challenges include the drought that affects our agriculture, bearing in mind that our income is entirely dependent on agriculture, so there are times a child may lack exercise books, or school uniform because of that.

In addition to fluctuations in income for school materials, endemic poverty also leads to high levels of food insecurity throughout Mara Region. Hunger itself further exacerbates poor attendance. A group of students told the midline interviewer:

Respondent 1: If you starve you can't go to school.
Respondent 2: Because if you didn't eat food at night, you'll be weak to attend school
Respondent 3: Because if you didn't eat, you cannot think of studies as you'll think of food only.

Finally, many respondents cited menstrual hygiene as a hurdle to adolescent girls' school attendance and associated improved latrines with increased attendance levels for girls in this age range. One parent in Butiama explained:

I can say the improvement of toilets has made more girls attend school because a female child when she reaches an age of maturity there is a time they undergo their menstrual period. So, when they are in that situation they need a place where they can clean up like a toilet where there should be enough water for her personal cleanness. Therefore, if the environment of the toilet is bad most girls when they are on their monthly period won't attend school.

Another parent remarked upon attendance increases for girls as a result of PCI interventions:

Now the attendance is the same because PCI built toilets with the changing room for girls and everything is there for them which was the main challenge to them at first.

A third parent referred to the positive influence of both toilets and libraries on increased attendance:

It's true through having library and toilets it has been the biggest pulling factor for students to attend because the library is used by students as part of reading different books and also the toilets, because they

are clean, and they have water, soap and other needs. Therefore, by having these tools it has been the pulling factor for students to attend school.

Student Attentiveness

While hunger was commonly cited as a reason for students to miss school, it is an equally prominent concern among students who attend. A Musoma WE group member said, “it was reported last year that many children collapsed at school due to starvation.”

The midline student survey asked students, “yesterday, did you feel hungry while at school?” However, at baseline, the question was asked differently. At both baseline and midline, students were asked about every meal they had eaten in the past school day. At baseline, they were also asked “were you still hungry after eating [meal]?” Students who reported still being hungry after eating a meal were counted as not full. However, students who were not given a meal or snack at school that day were not given an opportunity to say whether or not they were hungry during the day and aren’t represented here. In addition, this means that in the control schools, where school feeding was rare at baseline, the few students who did eat at school are the entire source of the control baseline value. For this reason, direct comparison of the baseline and midline figures may not be warranted.

Table 15: Percent of Students Indicating They Are Full During the Day

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>All Students</i>	Treat	87.9%	51.8%	-36.1%	7.9%*
	Control	91%	47%	-44%	
<i>Girls</i>	Treat	88.4%	47.2%	-41.2%	6.3%
	Control	91.2%	44.3%	-47.6%	
<i>Boys</i>	Treat	87.3%	56.4%	-30.9%	9.3%*
	Control	90.0%	49.8%	-40.2%	

The classroom observation tool recorded the percentage of students who were off-task at the beginning, middle, and end of the observed lesson. The program indicator reports on the share of grade 2 students who were off task at each point during the lesson and averages these three observations. Off-task girls and boys were also counted separately.

Table 16: Percent of Students Off Task during Classroom¹⁸

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>All Students</i>	Treat	3.7%	5.0%	1.3%	1.4%
	Control	4.6%	4.5%	-0.1%	
<i>Girls</i>	Treat	3.8%	5.2%	1.4%	1.7%
	Control	5.0%	4.7%	-0.3%	
<i>Boys</i>	Treat	2.6%	3.5%	0.9%	1.1%
	Control	3.2%	3.0%	-0.2%	

¹⁸ This indicator averages measurement of being off task at the beginning, middle, and end of a lesson.

A number of factors contribute to student inattentiveness, with overcrowding due to both a shortage of teachers and a shortage of physical classrooms regularly cited as the most prominent by teachers, parents, and students alike. Observed classroom sizes were indeed very large. At baseline, the average number of students registered in observed classes was 77 in treatment schools and 93 in control schools. By midline, those figures had risen to 99 and 123, respectively. Average attendance on the day of observation was somewhat lower, reaching 71 in treatment schools and 85 in control schools. No second differences across registration or attendance were statistically significant and both counts recorded even splits of girls and boys in both survey waves. One treatment school teacher described her experience as follows:

I will talk on the side of shortage of classroom for example my standard four class it has 113 students but the maximum number of students in a class is 45 so you find it two or three times, and this causes the teaching exercise to be difficult and this makes the students not to understand well because even if you make a group for them to share a book some will just be peeping.

Government officials agreed that overcrowding was a serious concern with adverse impacts on educational quality. A Village Chairperson commented that “you cannot improve the quality of education if you have 2 teachers and 700 or 600 students.” A Ward Education Officer said, “the number [of students] has increased but the number of teachers has not increased. Sometimes you may find one school has seven teachers or eight while having more than eight to nine hundred students.” The Serengeti District Education Officer reported a deficit of 1,151 teachers.

Attentiveness in class is meant to be the primary mediating pathway between school feeding activities and literacy impacts. Unfortunately, the quantitative data were not collected in a manner that allows for linkages between school feeding on the day in question and classroom observations of attentiveness.

At baseline, it was unclear how the educators’ demonstrated use of new and quality teaching techniques or tools was measured. The classroom observation tool includes 40 different teacher practices, most of which could be considered quality teaching techniques. They are divided into: class activities, teaching methods, assessment methods, and material use. Given that teaching and assessment methods are those most under the control of the teacher, these 20 methods were focused on as the new and quality teaching techniques or tools. One hundred percent of educators demonstrated use of at least one of these 20 techniques at both baseline and midline, so its definition as an indicator would not be meaningful. For this purpose, the indicator was calculated as the percentage of those techniques used on average by a teacher. For example, teachers in beneficiary schools at midline were observed using on average 14.24 of the 20 techniques. As a result, the indicator was calculated as $(14.24/20 =) 71.2\%$. While this calculation does not perfectly capture the definition as a number of teachers, it more effectively captures the purpose of the indicator as a measure of improved practices than simply marking that all 100 teachers use at least one improved practice.

Educators' demonstrated use of new and quality teaching techniques or tools

Treat	71.7%	71.2%	-0.5%
Control	72.0%	73.5%	+1.5%

The program has introduced school gardens and agricultural demonstration plots at school to enhance those school feeding activities. All school garden harvests are meant to supplement school meals; stakeholders reported PCI guidelines that 20 percent of demonstration plot produce should be donated to the school store. By the midline survey, 90 percent of FFE schools reported having a school garden and over a third also had demonstration plots.

Table 17: Percent of Head Teachers Reporting that Schools had Gardens and Agricultural Demonstration Plots

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>School Gardens</i>	Treat	48%	90%	42%	42%*
	Control	6%	6%	0%	
<i>Demonstration Plots</i>	Treat	36%	38%	2%	4%
	Control	24%	22%	-2%	

Phase III schools made particular strides in adopting both school gardens and demonstration plots between 2017 and 2019. At baseline, a quarter of these new program schools had gardens and 13 percent had demonstration plots, but those shares jumped to 94 and 56 percent, respectively, by the midline survey. In contrast, earlier phase schools hosted agricultural demonstration plots in just 29 percent of cases at midline, a decline from 47 percent at baseline. It is unclear why previously established plots co-located at schools appear to have grown inactive.

Quality of Literacy Instruction and School Management

Improved quality of literacy instruction is elevated as one of three program intermediate results supporting improved literacy outcomes, alongside improved student attendance and improved student attentiveness. Trainings for head teachers and literacy teachers are the primary modality for improving instructional quality. Table 16 details the most common types of trainings received by head teachers and the shares receiving those trainings across

Table 18: Percent of FFE School Head Teachers Who Report Receiving Training from PCI by Most Common Training Type (50 schools)

	Baseline	Midline
<i>Any Training</i>	64%	90%
<i>Commodity Mgmt, Safe Food Storage and Prep.</i>	22%	60%
<i>Readership Teaching Methodologies (including ToT)</i>	20%	72%

<i>Health Screening</i>	22%	32%
<i>School Gardening</i>	16%	30%

The proportion of head teachers who had received some type of training at baseline is driven by the experiences of Phase I and II schools, where FFE programming had already implemented years of activities. However, it is peculiar that 25 percent of Phase III school head teachers also reported having received PCI training by the 2017 baseline survey.

Unlike head teachers, literacy teachers were not sampled in quantitative data collection efforts so their training participation levels are not discernable from the midline data at hand. However, in the key informant interview format, many literacy teachers reported both receiving trainings and modifying their teaching techniques as a result. One teacher in a high performing school explained:

I told you directly, efficiency comes because of tools, the tools which we are used to teach children now, you teach them through observing and practices. They observe and do practice. In the training we were told to involve children in doing the work, and their work to be seen.

Another literacy teacher reflected:

Something that I did different was that I was able to engage the students while I was teaching, and at some time I even involved them in the collection of the raw materials that we would use to construct a teaching or learning aid. And at times I would request them to create a certain learning aid for me and this would increase their learning ability. [...] Now there is a huge change because after the PCI project we can say that it is the speaking class. It is different from how it was before because it was a class which does not speak.

While qualitative respondents broadly reported changes in their pedagogical approach as a result of trainings, these shifts were not reflected by the classroom observation tool. Almost no significant differences related to class organization (2 variables), instructional content (14 variables), activities (11 variables), teaching methods (12 variables), or assessment techniques (7 variables) were recorded.¹⁹

Classroom observations show similar results. In each school, one grade 2 teacher was observed: 88 percent of observations in treatment schools and 92 percent of observations in control schools, came to class with a lesson plan prepared. No significant differences at midline were detected for either the treatment or control group.

Table 19: Classroom Displays and Materials

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>Posters, Charts, or Pictures</i>	Treat	6%	68%	62%	44%
	Control	14%	32%	18%	
<i>Letters or Flash Cards</i>	Treat	50%	68%	18%	66%
	Control	86%	38%	-48%	
<i>Flip Charts</i>	Treat	36%	26%	-10%	36%
	Control	56%	10%	-46%	
<i>Alphabet Strip or Word Wall</i>	Treat	8%	30%	22%	2%
	Control	26%	24%	24%	

¹⁹ A total of three variables among this larger set of 46 exhibited any statistical significance across second differences.

<i>Student Work Displayed</i>	Treat	36%	70%	34%	50%
	Control	46%	30%	-16%	
<i>Most or All Students have a textbook or printed material</i>	Treat	34%	66%	32%	32%*
	Control	52%	52%	0%	
<i>Most or All Students have an exercise book, paper, or slate</i>	Treat	98%	98%	0%	2%
	Control	98%	96%	-2%	

While measurable differences in classroom management or instructional practices were captured by the classroom observation tool, differences in instructional materials were salient. Program classrooms exhibited posters, charts, and student work at about twice the frequency of Serengeti comparison classrooms. Between baseline and midline, the proportion of observed FFE school lessons in which most or all students had a textbook nearly doubled, while in comparison schools that proportion remained constant. Each of these differences was statistically significant and large in magnitude.

In addition to provide classroom materials, PCI has supported the construction or establishment and management of school libraries to further incentivize reading among students. At baseline, over half of Phase I and Phase II schools (56 percent) reported having a library compared to just one in eight Phase III schools. By the midline evaluation, the proportion of Phase III schools with libraries had jumped to 56 percent as well, while the share of earlier phase schools with libraries remained constant.

Quantitative tools at both baseline and midline also sought to document a broader set of good school management practices that may have indirect bearing on instructional quality. School head teachers were asked about three key documents which were physically verified by enumerators: the academic year budget, the internal performance evaluation plan, and the school development plan.

Table 20: School Administrators in Target Schools who Demonstrate use of Management Practices

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>Budget for this Academic Year</i>	Treat	46%	56%	10%	-30%*
	Control	26%	66%	40%	
<i>Internal Performance Evaluation Plan</i>	Treat	30%	50%	20%	4%
	Control	44%	60%	16%	
<i>School Development Plan</i>	Treat	40%	62%	22%	16%
	Control	78%	84%	6%	
<i>Any of the Above 3 Techniques</i>	Treat	66%	82%	16%	6%
	Control	84%	94%	10%	
<i>All of the Above 3 Techniques</i>	Treat	14%	32%	18%	-12%
	Control	16%	46%	30%	

Verified treatment school use of these practices increased by 10 percent, 20 percent, and 22 percent, respectively, in just two years. Such gains are meaningful despite their lack of statistical significance when compared to changes observed in control schools. The midline sampling strategy was designed to generate statistical power to detect differences between groups of students (as the primary sampling unit) rather than between groups of schools. With only 50 schools in each group, there is insufficient power to detect significant differences between types of schools in most instances. The significant and negative second difference in the demonstrable use in school budgets is driven by a remarkable increase in documented budgets from Serengeti schools which is beyond the scope of the evaluation to explain. It should in any case not be interpreted as a negative reflection on treatment school progress.

Student Literacy Outcomes

Literacy outcomes are measured with the Early Grade Reading Assessment (EGRA) which is made up of five subtasks designed to assess foundational reading skills: phonemic awareness, letter sound knowledge, devised word identification, oral passage reading, and reading comprehension. Key indicators of interest to the FFE program include students meeting the Government of Tanzania's national benchmarks of reading 40 devised words per minute, 50 words in a minute for oral reading fluency, and demonstrating 80 percent or higher reading comprehension.

Table 21: Percent of Students Scoring Above National Benchmarks in Non-Word Reading by Grade and Sex

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>Second Grade (All)</i>	Treat	5%	13%	8%	14.6%*
	Control	9.7%	3%	-6.6%	
<i>Second Grade Girls</i>	Treat	8.7%	13.7%	5.0%	13.3%*
	Control	10.7%	2.3%	-8.3%	
<i>Second Grade Boys</i>	Treat	1.3%	12.3%	11.0%	15.9%*
	Control	8.7%	3.7%	-4.9%	
<i>Fourth Grade (All)</i>	Treat	17.6%	7.3%	-10.3%	-6.4%*
	Control	9.7%	5.7%	-3.9%	
<i>Fourth Grade Girls</i>	Treat	16.7%	9.7%	-7.0%	-4.7%
	Control	9.3%	7.0%	-2.3%	
<i>Fourth Grade Boys</i>	Treat	18.7%	5.0%	-13.7%	-8.1%*
	Control	10.0%	4.4%	-5.6%	

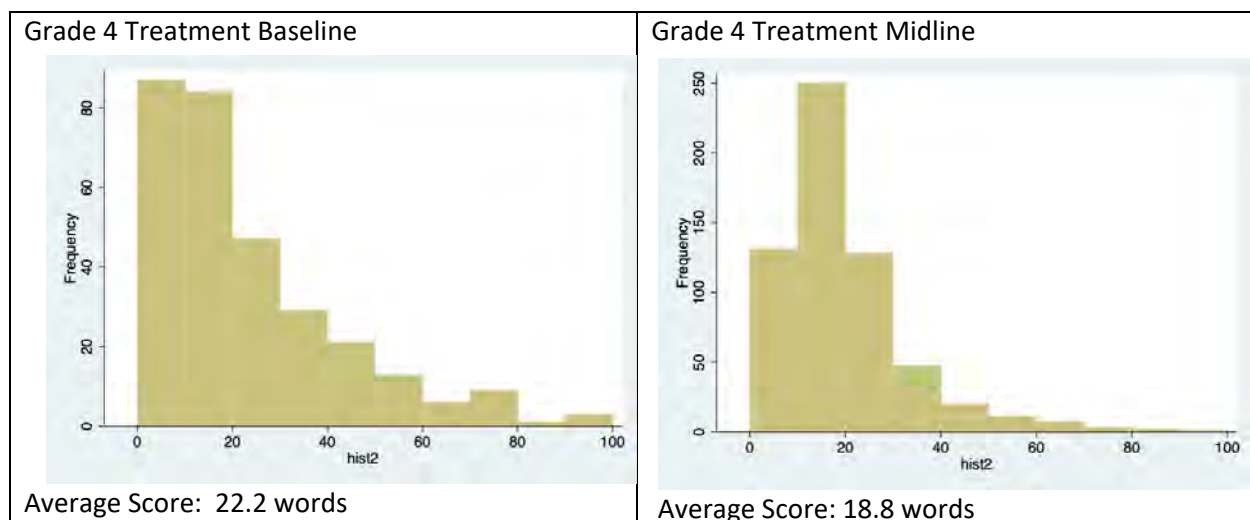
The indicator is defined as the percent of students who score above national benchmarks in non-word reading, with the national benchmark set at reading 40 devised words per minute. Invented words, instead of actual words, are used to measure a students' ability to construct words from letters without relying on words they have already memorized: it is widely accepted as a critical tool to measure the milestone of word construction.

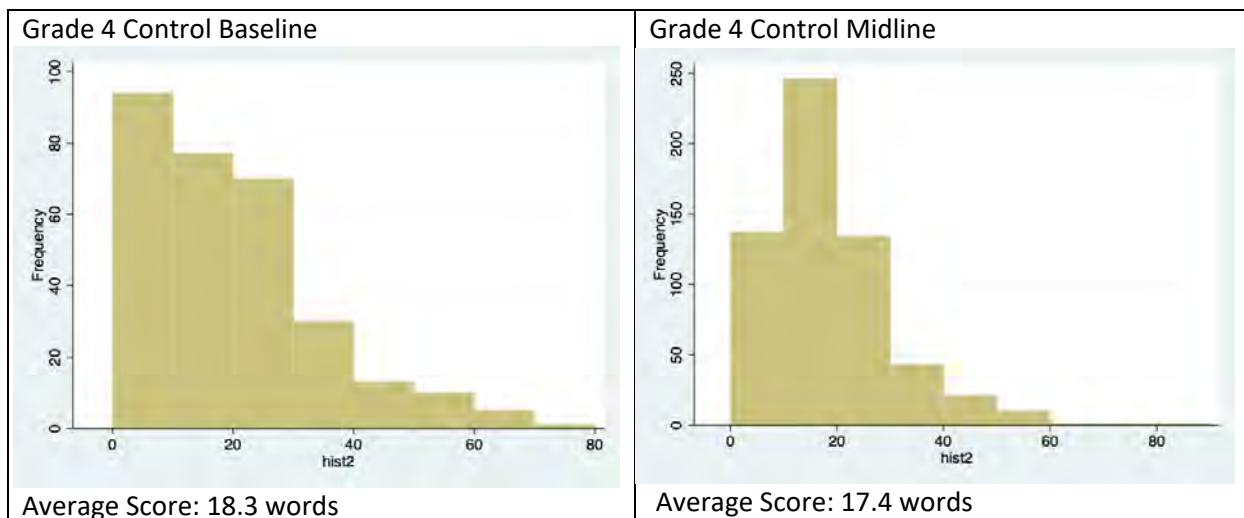
The percent of second grade boys and girls who could meet the benchmark increased in treatment schools and decreased in control schools, resulting in a significant difference collectively and for each gender.

The percent of fourth graders achieving the same benchmark fell in both treatment and comparison schools. In fact, in some cases, fewer fourth graders reached the benchmark than their grade 2 counterparts at midline. Beneficiary students' achievements fell more for beneficiary than control schools to a statistically significant degree. While beneficiary grade 4 performance fell 10.3 percentage points, control grade 4 performance fell 3.9 percentage points, a statistically significant difference.

The mixed results of gains in 2nd grade and a fall among 4th graders are less puzzling if one considers the indicator design. Measuring the percent that reach 40 words per minute means that even under ideal circumstances, this indicator ignores students who improve from reading 25 to 35 words per minute, or 41 to 51 words per minute. Reviewing the average scores on the non-word assessment shows negative, but less dramatic results.

Figure 3. Distribution and Average Number of Non-Words Read in One Minute, Grade 4





It is visible in the histograms of scores that most of the students are scoring much below the 40 words per minute threshold, and their progress cannot be captured by the current indicator definition. In both the treatment and control groups, the largest bar moves from being the students identifying 0 to 10 words per minute (the first bar) to students identifying 10 to 20 words per minute. While the average number of words fell from baseline to midline among both treatment and control groups, it is a much less dramatic change than suggested by the indicator. Using highly aspirational milestones instead of raw scores omits much of the meaningful understanding of changes in student learning.

Table 22. Average Number of Words Correctly Read Per Minute on Non-Word Assessment

		Baseline	Midline		
Second Grade(all)	Treat	15.6	19.4	3.8	11.8*
	Control	19.0	12.0	-7.0	
Fourth Grade(all)	Treat	22.2	18.8	-3.4	-2.5
	Control	18.3	17.4	-0.9	

With the understanding that subtask scores provide more information than the indicator, it is worth noting that the gains in grade 2 also appear more modest than suggested by the indicator. The average number of words increased by 3.8 words among beneficiaries and decreased by 7.0 words among comparison students, resulting in a significant difference of 11.8 words.

The agreed methodology used was to compare improvement in average grade 2 scores over time and grade 4 scores over time using a difference-in-difference analysis. An alternate perspective is to consider a cohort methodology: the grade 4 students at midline are of the

same cohort of the grade 2 students at the baseline. It is even possible some grade 2 baseline students were in the grade 4 midline sample. By comparing grade 4 midline scores to grade 2 baseline scores, a new narrative could be considered. In treatment schools, the cohort increased from 15.6 words per minute at baseline to 18.8 at midline; in control schools, the cohort decreased from 19.0 at baseline to 17.4 at midline. Given that this is not the agreed methodology in this report, it is not studied at length in this report, but is worth noting that this alternate approach can lead to different conclusions.

Table 23: Percent of Students Scoring Above National Benchmarks in Oral Fluency by Grade and Sex

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>Second Grade (All)</i>	Treat	19.3%	3.3%	-16%	8.5%*
	Control	25.3%	0.8%	-24.5%	
<i>Second Grade Girls</i>	Treat	25.3%	3.7%	-21.7%	2.3%
	Control	24.7%	0.7%	-24.7%	
<i>Second Grade Boys</i>	Treat	13.3%	3.0%	-10.3%	14.6%*
	Control	26.0%	0.01%	-25.0%	
<i>Fourth Grade (All)</i>	Treat	45%	26.7%	-18.3%	-12.1%*
	Control	35%	28.8%	-6.2%	
<i>Fourth Grade Girls</i>	Treat	47.3%	30.3%	-17.0%	-14.3%*
	Control	34.0%	31.3%	-2.7%	
<i>Fourth Grade Boys</i>	Treat	42.7%	23.0%	-19.7%	-9.9%
	Control	36.0%	26.2%	-9.8%	

Oral reading fluency follows a similar trend to non-word reading: the percentage of students meeting the national benchmark improved among second graders but fell among fourth graders. Oral reading fluency is measured by providing a student with a passage to read aloud: the enumerator counts how many words were correctly read in the first 60 seconds. The national benchmark is to read 50 words per minute. Among grade 2 students, the improvement is almost entirely driven by improvements in boys' reading. It is unclear what dynamics have driven this difference but it merits further exploration. It is worth noting that during classroom observations, the number of girls noted as off task was slightly higher than the number of boys: in treatment schools, 5.2% of girls and 3.5% of boys were identified as off task. It may be that teachers are not engaging girls as well.

The statistically significant negative second difference among fourth graders in achieving the national oral fluency benchmark is driven by differences in measured performance at baseline. In 2017, treatment school students meaningfully outperformed their control school counterparts on this measure (62 percent of treatment fourth graders completed the oral fluency subtask compared to 54 percent of comparison fourth graders). Measured performance across groups is similar at midline, and both groups recorded declines over time. The negative second difference implies that, under the assumption of equal time trends across treatment

and control groups, we would expect that the proportion of FFE school students scoring at or above this benchmark would have been 12 percentage points higher, at about 39 percent, at midline in the absence of the program. But in this case, it is more plausible that the second difference is an artefact of baseline measurement.

It is worth pointing out that while the percentage of second graders who scored above national benchmarks on oral reading fluency fell significantly between the baseline and midline surveys, the binary nature of the benchmark variable may imply a greater decline than is warranted by student performance data. National benchmarks are cut-offs, and when students are near that cut-off, a modest change in performance can result in their scores shifting just above or below the defining threshold.

Table 24. Average number of words in 63-word oral passage read in one minute

		Baseline	Midline		
<i>Second Grade(all)</i>	Treat	23.9	16.8	-6.1	11.9*
	Control	31.0	12.0	-19.0	
<i>Fourth Grade(all)</i>	Treat	38.6	34.4	-4.2	-10.3*
	Control	33.3	39.4	6.1	

Breaking down the EGRA Assessment by subtask scores demonstrates the differences between the percent of students achieving benchmarks and the average scores. Even though the percentage of students reading 40 or more words fell among grade 4 comparison students, the average number of words read actually increased from 33.3 words to 39.4 words. While fewer could read 40 or more words, the percentage of grade 4 comparison students who could not read any words fell from 13 percent to 3 percent. Among treatment grade 4 students, the proportion of students who could not read any words rose slightly from 7.0 to 8.5 percent. The decline in scores among grade 4 beneficiaries is still meaningful but is of a lesser order of magnitude decline in students meeting the national benchmark.

Table 25: Percent of Students who Demonstrate that They Can Read and Understand the Meaning of Grade-Level Text by Grade and Sex

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>Second Grade (All)</i>	Treat	7.3%	8%	0.7%	14.5%*
	Control	16.3%	2.5%	-13.8%	
<i>Second Grade Girls</i>	Treat	9.3%	8.7%	-0.7%	10.7%*
	Control	14.0%	2.7%	-11.3%	
<i>Second Grade Boys</i>	Treat	5.3%	7.3%	2.0%	18.3%*
	Control	18.7%	2.4%	-16.3%	
<i>Fourth Grade (All)</i>	Treat	25.0%	32.2%	7.2%	-7.7%*
	Control	22.0%	36.9%	14.9%	
<i>Fourth Grade Girls</i>	Treat	23.3%	32.0%	8.7%	-6.7%
	Control	21.3%	36.7%	15.3%	
	Treat	26.7%	32.3%	5.7%	

<i>Fourth Grade Boys</i>	Control	22.7%	37.1%	14.4%	-8.7%
--------------------------	---------	-------	-------	-------	--------------

Reading comprehension is the most difficult of the literacy tasks given to the student. After students are given additional time to finish reading the oral fluency passage, they are asked five questions about the story. The indicator, as described by McGovern-Dole, is the *percent 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*. The national benchmark is for students to answer 80 percent or more (four out of the five) of the comprehension questions correctly. As with the other literacy indicators, scores increased among grade 2 students and fell among grade 4 students.

Table 26: Average Number of Comprehension Questions Answered Correctly (out of 5)

		Baseline	Midline	1 st Difference	2 nd Difference
<i>Second Grade(all)</i>	Treat	1.3	1.4	0.1	1.1*
	Control	1.9	0.9	-1.0	
<i>Fourth Grade(all)</i>	Treat	2.3	2.6	0.3	-0.1*
	Control	2.3	2.7	0.4	

Also, as with the other indicators, the results are much less dramatic in terms of average scores. Although the trend of improvements in grade 2 and lagging in grade 4, the differences in improvement are smaller. Treatment second graders improved by answering 1.1 more questions correctly and Treatment fourth graders scores fell by 0.1 questions. Both differences were statistically significant.

Literacy is a complex ability to measure, and while its improvement may be the overarching goal of the project, the myriad of other factors that affect students' reading make it difficult to assign clear effects to the project. For example, observed increases in average classroom size between baseline and midline due to changes in government policy are likely to influence the classroom environment and may contribute to literacy outcomes. Further, by defining indicators in terms of national benchmark achievement, the indicators lose the nuance of measuring improvement among students whose scores were well below or well above the benchmarks.

Despite a mixed impact on literacy as evidenced by EGRA testing, some literacy teachers felt that outcomes had improved as a result of new teaching techniques. "Personally, I can say my teaching has progressed because initially at the end of the year I would have many students who cannot read, but now at the end of the year those who cannot read are few."

Section 6: Key Impact Findings Related to Strategic Objective 2: Increased Use of Health and Dietary Practices

The FFE program's second strategic objective is to increase the use of health and dietary practices. It operationalizes this goal with a two-pronged strategy, (1) enhancing knowledge of health, hygiene, nutrition, and food preparation practices while simultaneously (2) increasing access to clean water and sanitation services, preventative health interventions, and food preparation and storage tools and equipment.

Student Knowledge of Health, Hygiene, and Nutrition Practices

In the midline survey, students in both treatment and control schools exhibited a significant, parallel decline in knowledge of health and hygiene practices, slipping from 71 percent at baseline to 16 percent at midline. The control group decline is significantly less than the control group, but the reliability of the funding is undercut by unclear methods at the baseline.

Both survey waves posed the question underlying this indicator identically, "Can you name three good practices in personal hygiene?" Both also provided the same set of four precoded responses to enumerators and were instructed not to prompt the students with these options. Precoded responses included hand washing, keeping the environment clean, taking a bath, and drinking clean water. Enumerators also captured additional responses coded as "other," with specification.

Table 27: Percent of Students with Knowledge of Good Health and Hygiene and Important Nutritional/Dietary Practices

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>3 Good Health/Hygiene Practices</i>	Treat	66.6%†	17.3%	-49.3%	10.8%*
	Control	75.5%†	15.4%	-60.1%	
<i>3 Good Health/Hygiene Practices, Girls</i>	Treat	†	14.9%		
	Control	†	12.4%		
<i>3 Good Health/Hygiene Practices, Boys</i>	Treat	†	19.7%		
	Control	†	18.5%		
<i>3 Important Nutritional/Dietary Recommendations</i>	Treat	36.9%	46.8%	9.9%	1.3%
	Control	39.4%	48%	8.6%	
<i>3 Important Nutritional/Dietary Recommendations, Girls</i>	Treat	35.5%	44.7%	11.2%	5.4%
	Control	38.5%	44.3%	5.8%	
<i>3 Important Nutritional/Dietary Recommendations, Boys</i>	Treat	38.3%	50.3%	12.0%	0.6%
	Control	40.3%	51.7%	11.4%	

Note: † indicates baseline values obtained from baseline report, and may have been calculated differently than midline values. In most such cases, values were not sex-disaggregated. In all other cases, values were recalculated to ensure comparability.

The observed universal decline on this knowledge indicator suggests that the question may have been administered differently across survey waves. It is possible that enumerators suggested response options at baseline or did not fully capture “other” responses at midline. It is also possible that “other” responses were treated differently during analysis, but the overall prevalence of “other” responses is relatively low, making it implausible that this response alone drives the high magnitude of observed changes. When unique values of “other” are included in the midline data analysis to count towards correct responses, the proportions of treatment and control group students meeting indicator criteria increase only a small fraction, by 1.5 percent and 1.7 percent, respectively.

Qualitative evidence suggests that students do possess a basic understanding of good health and hygiene practices, and the connection of those practices to positive health outcomes. A group of 7th grade boys in Bunda explained:

Respondent 1: It is not only about science, you are supposed to take care of yourself in terms of cleanliness, wearing clean clothes and shoes.

Respondent 2: The teachers tell us there are many effects when one is dirty; you can get cholera and you end up with diarrhea which can lead to death.

However, the qualitative data did not surface any notable differences in basic knowledge between treatment and control schools (it also did not probe for this deeply). A group of Serengeti boys talked about health lessons at school:

Interviewer: What did you learn?

Respondent 1: Our teachers normally teach us about cleanness.

Interviewer: What do they teach you about cleanness?

Respondent 2: Our bodies and environment should be clean to avoid diseases.

When asked about the importance of cleanliness, a group of Serengeti girls explained it was important to avoid diseases, including cholera. Another 7th grade boy explained a vector of contamination at school:

You get many diseases in school because you can find that the bell for break time rings and someone has just come out of the toilet which is dirty and then finds a mango and starts eating it and then he wants to share it with somebody else and that is how we end up getting sick. That is why many diseases start from the toilet and that is also why we learn science in school.

The second main student knowledge indicator captures student awareness of important nutritional or dietary recommendations. It similarly exhibits parallel changes across treatment and control groups over time. In this case, a higher share of students demonstrate knowledge over time in both groups, increasing from about 38 to about 48 percent. No statistically significant differences are detected between treatment and control groups.

However, in this case, the questions from which the indicator is derived were not identical at baseline and midline. In both periods, the indicator is based solely on knowledge of food groups rather than a broader set of nutrition recommendations. At baseline, students were asked about *three items* that one should eat every day. Free responses were then coded to map onto the following food categories: fruits, eggs, grains, cereals, milk, flesh foods, and vegetables. Grains and cereals were erroneously coded as different food groups. In the midline survey, students were asked to name *five food groups*, and given credit for a correct response, based on indicator definition, if they could name three. Food groups at midline were changed to match USDA’s updated five food groups and adapted for local context:

- 1 = cereals, tubers, and ripe banana
- 2 = legumes and animal-sourced foods
- 3 = vegetables
- 4 = fruits
- 5 = fat, sugar, honey

The observed (positive) time trend is just as plausibly driven by the change in question wording and response options as by differential administration of the questionnaire.

Caregiver Knowledge of Health, Hygiene, and Nutrition Practices and Benefits of Education

Parental knowledge of key health and hygiene behaviors was assessed at baseline and midline with similar questions. At baseline, parents were asked, “what are the three most important health/hygiene practices that should be followed?” At midline, the English language question was revised slightly to, “what are three important health/hygiene practices that should be followed?” It is unclear if the Swahili translation was uniform across waves. In both surveys, 12 precoded values were provided:

- 1 = using toilet
- 2 = washing hands after defecation
- 3 = washing hands before meals
- 4 = washing hands after meals
- 5 = wearing clean clothes
- 6 = brushing teeth every day
- 7 = washing mouth before/after food
- 8 = taking regular baths
- 9 = eating balanced diet
- 10 = wearing shoes before going out
- 11 = eating food 3 times a day
- 12 = eating a balanced diet

Note that code 9 and code 12 both represent the same value. This was an error in the instrument but no single observation (parent respondent) in either the baseline or the midline survey recorded both codes. Rather, enumerators selected one value or the other when parents provided this answer.

Table 28: Percent of Parents with Knowledge of Important Preventative Health Practices

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>3 Important Preventative Health Practices, Mothers</i>	Treat	40.5%	26%	-14.5%	-4.3%
	Control	41.5%	31.3%	-10.2%	
<i>3 Important Preventative Health Practices, All Caregivers</i>	Treat	68.3%	39.8%	-28.5%	-8%
	Control	64.3%	43.8%	-20.5%	

Consistent with the previous health practices knowledge indicator, parental health knowledge appears to decline between the baseline and midline surveys. The magnitude of these declines is again consistent across treatment and control groups, and second differences have no significance for either mothers alone or for all caregivers. The pattern suggests that respondents may have been prompted with response options during the baseline survey or insufficiently probed in the midline survey. In either scenario, no differences are attributable to program interventions.

School Water and Sanitation and Health Management Practices

The program's second strategic objective is predicated upon both increased (demand-side) knowledge of healthy behaviors and also increased (supply-side) access to health-sustaining inputs. PCI has worked to improve school water and sanitation facilities as a supply-side modality to improve health outcomes. Specifically, its programming has supported safe and improved water sources at school and latrine construction and repair. Overall, the prevalence of improved water sources and improved sanitation facilities are both low relative to baseline, (which has declined by 11 percent) and low in terms of functioning toilets per student.

The baseline evaluation's source of information about improved water sources is unknown, as the baseline instruments did not include any data about improved water sources or sanitation facilities. For the midline, the data on schools with improved water sources and sanitation facilities were obtained from project management data. No data are available for comparison schools, so difference-in-difference tests are not possible. T-tests did not find that midline values were significantly different from baseline values in beneficiary schools. It appears that the prevalence of improved water sources and sanitation facilities is lower at midline than reported at baseline, which merits further study.

The data suggest that either too few improvements are being constructed, and/or built facilities are breaking. This may be due to the fragility (or inappropriateness) of the improved facilities or low levels of implementation. Given the lack of clarity around how this information was collected at baseline, it is difficult to assess accurately how the state of facilities has changed over time.

Table 29: Number and percentage of Schools with an Improved Water Source or Improved Sanitation Facilities

	Group	Baseline	Midline	1 st Difference
<i>Improved Water Source</i>	Treat	76%†	65% (149)	-11%
	Control	32%†	--	--
<i>Improved Sanitation Facilities</i>	Treat	86.5%†	79% (183)	-7.5%
	Control	80.6%†	--	--

Even though the number of schools constructing facilities is falling, it appears that the total number of sanitation facilities being constructed is increasing. At both baseline and midline, head teachers were asked if any sanitation facilities had been constructed or repaired for boys or girls in the past year. In treatment schools, 40 percent of head teachers reported construction or repairs at baseline, and 32 percent reported construction or repairs at midline. Among beneficiary schools, the same question's values fell from 62 percent to 48 percent. Construction prevalence fell among both groups but is higher among unsupported schools. However, the total number of facilities constructed among the 50 sampled treatment and 50 control schools are comparable over time and are increasing.

Table 30: School Bathroom Construction in Past Year

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>Schools with any bathroom construction in past year</i>	Treat	40%	32%	-8%	8%
	Control	62%	48%	-16%	
<i>Total Bathrooms Constructed in Past year, all sampled schools</i>	Treat	192	298	102	-19
	Control	158	279	121	

Even if toilet construction is improving, the actual number of working toilets per student is low. At midline, head teachers were asked how many toilets (of any type) were currently functioning, for boys and for girls. Using enrollment data, we found that there were very few functioning toilets. The World Health Organization standards for low-cost settings recommend a minimum of one toilet per 25 children in schools (Adams, Bartran, Chartier, & Sims, 2009).²⁰

Table 31: Average Number of Students per Toilet Facility at School (midline only)

Group	Midline
-------	---------

²⁰ Recommendations are one toilet per 25 girls and one toilet and one urinal per 50 boys.

<i>Girls</i>	Treat	66.7
	Control	84
<i>Boys</i>	Treat	81.4
	Control	89.5

In addition to physical infrastructure, school organizational and management practices related to health were recorded at both baseline and midline.

Given the challenges of finding an indicator that is comparable at both baseline and midline on improved toilet facilities, we recommend expanding the data collection instrument at endline to include more detailed questions on the types of toilets and their level of functionality to capture the original intent of the indicator.

Table 32: Percent of Schools with Active Committees, Health Clubs, and Volunteer Student Aids

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>School Committee or Parent Teacher Association</i>	Treat	94%	98%	4%	18%*
	Control	96%	82%	-14%	
<i>School Feeding Committee</i>	Treat	44%	68%	24%	38%*
	Control	20%	6%	-14%	
<i>SWASH Committee</i>	Treat	48%	58%	10%	22%
	Control	34%	22%	-12%	
<i>Health Club</i>	Treat	66%	94%	28%	32%*
	Control	58%	54%	-4%	
<i>Volunteer Student Aids</i>	Treat	24%	82%	58%	60%
	Control	28%	26%	-2%	

Program schools exhibited strong gains in management activities to improve the school environment overall, with significant positive differences from comparison schools. By midline, head teachers reported that 98 percent of treatment schools had School Committees or Parent Teacher Associations, nearly three in five had active SWASH Committees, 94 percent had student health clubs, and over four fifths had Volunteer Student Aids (VSAs). Caution should be exercised not to over-interpret these differences: in general, existence of these systems does not imply they are completing their function.

Among treatment schools, the 16 Phase schools in the sample posted particularly robust increases in SWASH Committee, Health Club, and VSA activity, outstripping the earlier phase schools in all three areas. Among these newest program schools, 69 percent were active with SWASH (compared to 58 percent of all treatment schools), all had active health clubs, and 94 percent utilized VSAs (compared to 53, 91, and 76 percent of Phase I and II schools, respectively).

Dietary Quality

Enhanced dietary and health knowledge combined with increased access to nutritionally diverse and safely prepared food is expected to contribute to improved student dietary diversity.

Table 33: Percent of Students Receiving a Minimum Acceptable Diet (MAD)

	Group	Baseline	Midline	1 st Difference	2 nd Difference
<i>All Students</i>	Treat	40.9%†	56.8%	15.9%	10.8%*
	Control	41.3%†	46.4%	5.1%	
<i>Girls</i>	Treat	†	59.4%		
	Control	†	46.3%		
<i>Boys</i>	Treat	†	54.1%		
	Control	†	46.5%		

Note: † indicates baseline values obtained from baseline report, and may have been calculated differently than midline values. In most such cases, values were not sex-disaggregated. In all other cases, values were recalculated to ensure comparability.

The indicator of students receiving a Minimum Acceptable Diet (MAD) is based on student reporting of what they had eaten the previous school day. A student is assessed to have a minimum acceptable diet if he or she ate three or more meals on the previous day in which they consumed food from at least four food groups. While on average girls were slightly more likely than boys to meet the minimum acceptable diet, the difference is not statistically significantly different.

At baseline, the calculation appears to have relied upon a non-standard definition of food groups, classifying grains and cereals separately. At midline, the McGovern-Dole indicator definition was used, including the following food groups (see Annex E for further details):

- Grains, roots, and tubers
- Legumes and nuts
- Dairy products
- Fish or meat
- Eggs
- Vitamin A rich foods
- Other fruits and vegetables

Baseline values could not be recreated with the data available and are instead reproduced here as reported in the baseline evaluation. Caution should thus be exercised in interpreting differences between survey waves. Within the midline data, a higher proportion of treatment school students are classified as having a minimum acceptable diet than their comparison school counterparts. The difference of approximately 10 percentage points is consistent for both girls and boys.

Section 7: Key Findings Related to Program Sustainability

PCI's Food for Education program has continuously delivered school feeding activities in Mara Region since 2010. It began Phase I in 103 schools, expanded to 156 schools in Phase II, and is now active in 231 schools across 45 wards. In two years, its programming will conclude, and beneficiary communities will need to manage school feeding and complementary activities on their own.

A core objective of this midline evaluation is to provide PCI with evidence-based guidance on how to ensure a successful transition to community ownership.

Parental Engagement and School Feeding Program Contributions

The sustainability model of the program's school feeding component is organized around parents contributing food in sufficient quantities to fully offset USDA commodity donations. In the midline survey, 95 percent of parents stated a willingness to contribute to the school feeding program, a positive reflection of parents valuing the practice of meal provision at school. However, only 38 percent of parents had made contributions in the preceding six months (at baseline, 12 percent of parents had donated food within the same reference period). In contrast, 91 percent of parents had contributed to related expenses such as fees for guards and cooks. It is important to note that the six-month recall period falls during a lean season in between harvest and the next planting, where household food is more likely to be scarce. A twelve-month recall period would better capture community contributions irrespective of seasonal fluctuations.

In 2019, PCI is in the process of transitioning school feeding programs to rely on community donations and food grown on school grounds. Its planned gradual withdrawal of USDA commodity support shifted abruptly when the Government of Tanzania declared that U.S. commodities could no longer be imported to supply school meals. The plan had involved the provision of sufficient food for two or three meals per week, depending on the phase of the school (Phase III schools, as newer program entrants, receive support for more meals each week than Phase I and II schools). Going forward, all schools will rely on community and government contributions to supply school meals.

Table 34: School Feeding Program Commodity Contributions (kg) from June 2018 through April 2019 by Source and Commodity Type for 50 Midline Evaluation Schools²¹

<i>Commodity</i>	<i>Source of Donation</i>						<i>Total (kg)</i>
	<i>USDA</i>	<i>Community (Parents)</i>	<i>School Farm</i>	<i>School Garden</i>	<i>WE Groups</i>	<i>Farmer Groups</i>	
<i>Rice</i>	225,654						225,654
<i>Maize</i>	1,240	43,790	22,858		851	325	69,064
<i>Pinto Beans</i>	68,283						68,283

²¹ The source of these data is PCI's commodity tracking system.

<i>Oil</i>	18,741						18,741
<i>Beans (local)</i>		7,623	561		61		8,245
<i>Sweet Potato</i>		2,289	1,352				3,641
<i>Vegetables</i>	183	107	400	2,690			3,380
<i>Sorghum</i>		195	325				519
<i>Salt</i>		352			40		392
<i>Oil (local)</i>	62	307		8	3		380
<i>Rice (local)</i>		187					187
<i>Millet</i>		152					152
<i>Fruit</i>			51				51
Total	314,163	55,002	25,546	2,698	955	325	398,688

PCI provided school store commodity tracking data for the same 50 schools, which were included in the baseline and midline evaluation and reflecting proportional shares of schools across three phases of programming. Over an 11-month period from June 2018 through April 2019, these schools recorded a total of nearly 400,000 kg of new food commodity receipts. Of that total, 79 percent (by weight) was sourced from USDA while about 14 percent came from community contributions. Most of the balance came from school farms with a small portion from school gardens. The recorded contributions from WE Groups and Farmer Groups were marginal, but this categorization likely does not reflect household/individually-sourced contributions from group members. If schools, on average, serve three meals per week with USDA food and two with community-sourced food, a sustainable model would imply that the share of new commodities sourced from USDA should not exceed 60 percent of the total. School store record data are consistent with reported experiences that schools have been unable to provide meals every day.

Students from a designated “high performing” treatment school knew well which days meals were provided:

Interviewer: How many times in a week, do they provide you with this food?

Respondent 1: They provide us with PCI’s food 3 times a week.

Interviewer: When is that?

All Respondents: Monday, Tuesday, and Thursday.

Interviewer: What do you eat on Wednesday?

Respondent 2: We don’t eat.

Reports from other schools were strikingly similar. A District Planning Officer from a high performing district surmised that, among the 20 district schools engaged with the program, “about five” managed to feed students on (any) days when PCI did not provide meals. A Planning Officer from a different district explained:

The project was going well but the challenge was that the community members were not contributing food on time as agreed.

A village chairperson in a high performing school elaborated, “challenges are many, first the response from the parents is very low and collection is hard because they don’t bring on time and other times you will find that there is no food.”

The chairperson of a different village was asked whether local parents are contributing to the program. “The community doesn’t contribute,” he responded. When pressed on why that is the case, he said, “Like I had told you earlier, the economic situation is bad.” He elaborated:

To say the truth it may not be sustainable because our state of food in general is that people have been blaming the weather. If you call for a meeting and do the math like asking the parents to bring 4 kilos to school, it is a problem, you will talk until you feel like crying, you tell people that these are our children and we should sacrifice and bring food and another may tell you that you are saying so because you are the Chairperson.

A literacy teacher told a midline interviewer that “most people are reluctant to contribute, they are not participating, they get involved while forced.” Other teachers expressed frustration that parents weren’t agreeing to contribute food.

Still there is no awareness about education and also they do not know the importance of education that is why when you tell the parents to contribute for a child so they can study it becomes difficult also when you call them in order to discuss how to contribute for a child no body attends, we once called parents they all only came ten of them and only one was willing to contribute others did not agree.

When one parent was asked if it is possible to continue to feed children at school, they responded:

Respondent: It's impossible because our income is low since even at home is not satisfied.

Interviewer: Why?

Respondent: We will get into conflict because every parent wants child to come back at home to take food, so community will get into conflict.

Some students explained that pressure to contribute to school feeding is itself a driver of absenteeism. One 7th grade boy in Musoma reported that “a parent may tell you to stay at home to avoid being asked about food contributions at school.” His classmates added:

Respondent 1: The parents may also state their unwillingness to contribute food at school out of their economic hardship and the fact that even if they contribute the child won’t get satisfied.

Respondent 2: They say to us that they can’t contribute the food that we may eat twice at school and at home, so the child should just eat at home.

Most teachers attributed the lack of contributions to parents not valuing the positive contributions of meals to learning outcomes. But parents’ appreciation of school feeding is reflected in both quantitative and qualitative data, and 94 percent of those interviewed reported that school feeding activities have increased their children’s willingness to attend school. Despite these positive attitudes, they are consistently unable to offset the food supply deficit left by USDA commodity reduction.

Among the 151 parents at midline who reported having made food contributions, 97 percent had donated maize and 67 percent had donated beans. The only other type of food donation reported by parents was rice, which was contributed by 5 individuals (3 percent).

Much of eastern Mara Region is a net deficit maize area, relying on imports from nearby regions and neighboring Kenya (FEWS NET, 2018). An examination of parental contributions over the calendar year reveals how closely they track the biannual harvest.

The National Education Sector Development Plan 2016/17-2020/21 to ensure that “main education costs borne by households are reduced by half for all pupils, and further still for those from poor and disadvantaged backgrounds” (Government of Tanzania, 2018). The pressured solicitation of contributions observed in some cases, an unintended outcome of the program, works against the policy edict to “reduce the burden of non-fee education costs on the poorest families, removing a disincentive to enrollment or retention” (Government of Tanzania, 2018).

Figure 4: Community Contributions of Maize and Beans by Month in 50 Midline Evaluation Schools

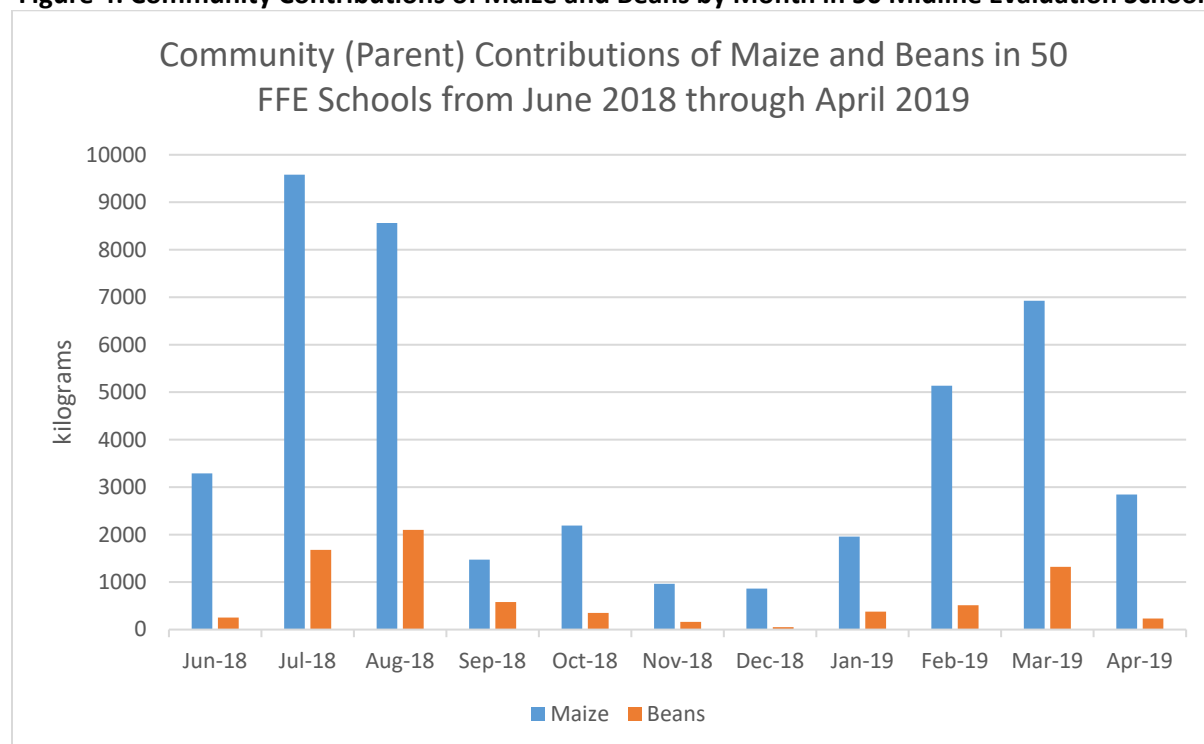
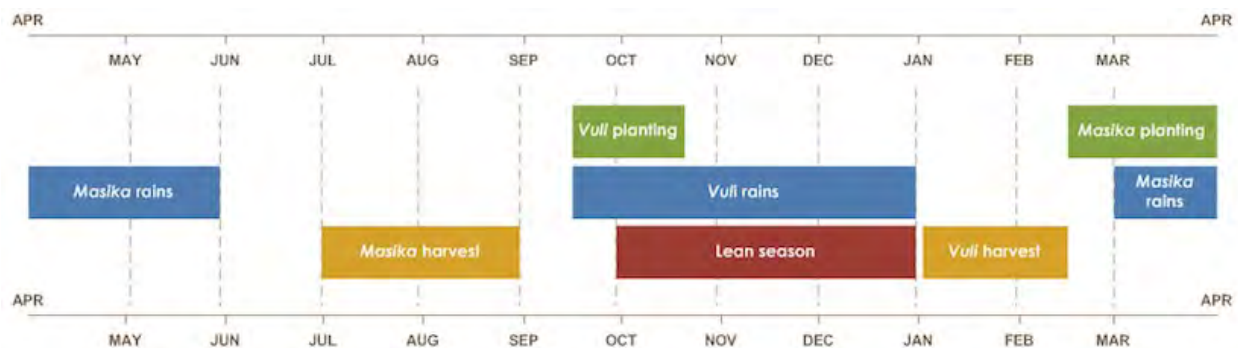


Figure 5: Agricultural Calendar in Tanzania



Source: (FEWS NET, 2013)

The strong correlation between parental contribution levels and harvest periods is consistent with the high levels of parental willingness to contribute, cited above, when food is available. In the absence of scarcity, we would expect to observe a more consistent level of community donations on a monthly basis when school is in session.

A District Planning Officer articulated the distinction between parental willingness and ability to contribute:

Parents are supposed to cover for those days that PCI does not provide food, this is something that is happening in some of the schools and in others the weather challenge that is facing the parents is making it difficult to implement their responsibility though they are willing.

A focus group discussion among parents in a high performing Bunda school led to the following exchange:

- Interviewer: When this program is over and the parents and the school have to carry the responsibility to feed the children at school, do you think you will be able to?
- Respondent 1: I think it will be a challenge because it has been hard for some parents to contribute even for the two days that we are contributing now.
- Respondent 2: On my side I think it will be impossible because of the low financial ability of the people of this area and the fact that we are all headed to the tough times.

Many parents and other respondents referenced the significant challenge that drought and climate change conditions have posed for local food production, heightening food insecurity at the household level. Over 80 percent of parents who responded to the midline quantitative survey reported that agriculture or agricultural labor was the primary source of income for their households. When a Bunda teacher was asked if there are people who should be contributing but are not doing so, he or she answered, “parents, but the most reason is the drought hence scarcity of food making parents unable to contribute food.”

According to PCI’s agriculture team, both 2018 and 2019 experienced extensive drought during the typical rainy season, resulting in poor harvest yield. In addition, the emergence of Fall Armyworm in 2017 has progressively become more of a pest to Tanzanian fields, destroying more maize crop with each year. Both drought and Fall Armyworm had a detrimental impact on household harvests and likely on their contributions to schools as well.

Other respondents said that continued parental contributions would be “possible” but struggled to elaborate upon their answers. This inability to explain how it would be more possible to contribute in the future than it is today suggests respondents’ desire to affirm what the program was asking of them, again calling into question the distinction between aspirations and ability. Indeed, a pattern emerged in the qualitative data whereby subjects subtly distinguished between their willingness to contribute and their real prospects for doing so. Subjects emphasized parents’ “readiness” to donate, conditioned on circumstances beyond their control: “If the condition is good, we will contribute without hesitation.” One WE group described their initial plans to contribute to school feeding:

- Respondent: We discussed an agenda that was cultivating our PCI farm and when we harvest maize and beans, we bring it to school for the students.
- Interviewer: Have you ever done that?
- Respondent: We have never done as you can see the weather the way it has been.

Other subjects who responded that communities could sustain school feeding programs emphasized a policy or community mandate to require contributions and increase social pressure on non-contributors. In these responses, the implication was a lack of “readiness” rather than risk exposure due to exogenous conditions. While both types of responses offer a positive assessment of contribution sustainability, the discordance in their underlying problem assessments is illuminating.

This pattern of knowing a “right” answer and agreeing that it should, aspirationally, reflect reality was mirrored in discussions of school costs. A Butiama parent group had the following conversation:

- Interviewer: Did you meet any difficulties on sending a child to school?
- All: No
- Interviewer: Why not?
- Respondent 1: Because it’s my duty to send my child to school.
- Respondent 2: There’s no difficulty because when you give birth to a child it’s your obligation to send him to school.

Serengeti parents also expressed this understanding, “there are not any difficulties, it is you as a parent to take your child to school and a teacher to welcome them.” But just a moment later in that same conversation, a participant observed that “some parents can’t afford to buy their children school uniforms.”

One group of FFE program parents explained that children eat two days per week in their school, indicating that the program had been active since 2013 or earlier. While community contributions were already insufficient to feed children for the remaining three days, parents said that they will be able to contribute in the future *because it is important*:

Interviewer: So, do you think parents of this school will be able to contribute enough food through other ways like labor force or money to pay the chefs to satisfy the needs of this project of providing food in the coming days?

Respondent: Yes

Interviewer: Why?

Respondent: Because we know the importance food to the school for our children.

A participant in a different group similarly explained:

I think it will be possible because the contributions that we offer are for our children to study well in school. We contribute food because there are students coming from far away and they need to eat food in order to help them in their studies.

Other respondents were less optimistic about school feeding sustainability that relies on parental contributions. A focus group of farmers was asked, “will the parents be able to produce sufficient food to feed the schools in the future?” They responded:

Respondent 1: It is possible depending on the rainfall which is in our area unreliable for instance this year we have experienced rainfall from May which has affected our harvests.

Respondent 2: Not to a 100% because the parents' response is not adequate.

Respondent 3: They are not capable, for instance this year there was no harvests at all. That makes it hard for the parents to give food to the school when their homes are starving.

Asked the same question, a parent in Butiama was blunt:

On my side, I think it is impossible because of the weather condition. In our community when there is drought you may cultivate and because of sun the crops shrivel, so in short it is impossible because of the weather conditions.

Students were quick to echo parental concerns about household food insecurity. One group of boys explained:

Interviewer: Who do you think is responsible for contributing food?

Respondent: The parents.

Interviewer: So why are the parents not contributing?

Respondent: They think that they will lack food at home if they contribute it to school.

A group of girls responded to the same line of questioning:

Interviewer: When the PCI project ends would parents or the society take over the responsibility as donors of food?

Respondent: No.

Interviewer: Why?

Respondent: They are not able.

Interviewer: What if each person contributes a little, what do you say?

Respondent: They do not have enough food.

A literacy teacher was similarly pessimistic, lamenting that “there is no light in the future whenever this project comes to an end.”

Other stakeholders remained focused on the need for ongoing sensitization and community education to increase donations. One said:

I think the program will not continue since the villagers show unwillingness to give a hand into the project. Leaving the entire burden onto the PCI groups, we may end up seeing the children being served only porridge. There is need for education among the villagers on this area.

Any continued efforts to persuade community members to increase contributions will, however, compete against long-established food insecurity coping strategies. One successful potato farmer explained the allocation of his harvest as follows:

I harvested 5 sacks from which I used 3 sacks for food and I sold 2 sacks to get myself out of some situations. Of the 3 sacks that I had stored for food, I contributed 1 sack to siblings, relatives, and friends.

FGD respondents frequently expressed greater comfort in sharing their agricultural produce within kinship networks than donating it to the school, where they found the distribution untrustworthy. A farmer group discussion in Musoma recorded the following exchange:

Interviewer: Is this project successful or not?

Respondent 1: I see this project is not successful because others see this as a foolish idea taking their food to school. Many people don't have their trust in the project.

Interviewer: Why do you think people fails to trust the project?

Respondent 1: It's peoples' mind-set.

Respondent 2: They don't want to donate or thinks PCI will provide food.

Respondent 3: They say that; teachers will use the food for their own needs.

A group of 7th grade girls said, “We love the food, but it is not enough because the teachers take some food to their homes.” When this group was subsequently asked about participation in school gardening activities, they noted again with frustration that teachers take vegetables from the gardens to their homes. Parents from the same community also complained that teachers arranged for students to work in their own private farms.

Finally, in addition to absolute food scarcity, harvest management preferences, and mistrust of school feeding administration, many parents and students expressed the sense that school is meant to be free and that the solicitation of contributions contradicts that edict.²² A 7th grade boy in Bunda observed:

Most of our parents say that the president has said that education is free so when you go home and tell her that you have been asked to make some contributions to the school, the parent will not understand you. Others even think really hard before buying an exercise book for their children. Education is not so free.

²² This common community understanding of the free education policy is inconsistent with the policy in question. While there are no longer school fees, indirect costs to households remain.

A 7th grade girl in Butiama repeated, “they say that the president has said that education is free, so they are not supposed to contribute for food.”

A District Planning Officer assessed:

I would say the major challenge is politics. For instance, whenever we go around emphasizing on the importance of contributing food to the parents; one of them happens to call the MP who discourages the exercise on the ground that the President has declared a free education without any contributions.

A group of teachers from a high performing school in Butiama were candid about their expectations for sustainable school feeding:

Interviewer: When the PCI project terminates the contract could you continue with the donation?

Respondent: No.

Interviewer: Why not?

Respondent: Right now, they hesitate to contribute and others recommend that their children eat at home so when it is over we just return to our normal activities.

A Bunda WE group member said, “I believe that after PCI ends the program no pupil will get food at school.” Many students shared this expectation. A group of 7th grade boys in Musoma expressed skepticism about the prospects for continuity:

Interviewer: After PCI’s project is completed; parents and the community will be responsible to feed the schools. Do you think this will be possible?

All: No.

Interviewer: Why have you all said no that fast?

Respondent 1: If right now teachers direct students to collect contributions from their parents but they still refuse to contribute, so what more if PCI is away?

Respondent 2: Parents are also claiming that it is better if the former system of going back at home at 2 is returned; that we shouldn’t be taking food at school.

Farmer Groups

Farmers repeatedly cited the many interconnected challenges of their agricultural enterprises while lamenting the lack of alternative livelihoods. “You must undertake agriculture in order to live,” one explained. “There is no life without agriculture.”

As described in Section 4, focus group respondents found PCI trainings useful and relevant. They also found them impactful. One participant explained how improved practices drove up yields:

Interviewer: Are you working on what you were taught in shamba darasa (demonstration plot)?

Respondent: Yes, because until now I have prepared a place where I cultivate spinach and collard greens.

Interviewer: The lessons from *shamba darasa*, have they improved the production of crops?

Respondent: Yes

Interviewer: How?

Respondent: Before I used to cultivate without considering the principles and due to this the yields were very little but after been trained in the *shamba darasa*. I follow up on the principles and I harvest enough crops.

Many Focus Group respondents described yields improving by 15 to 20 percent or more:

Before we were getting four to five sacks of maize after harvesting but after learning about demonstration plot we started getting six to seven sacks.

Other participants described new practices to improve and maintain soil fertility:

From different PCI seminars we were trained on how to retain soil fertility by planting different kind of crops. For example, I can plant beans then maize then potatoes, that way I can fertilize the soil.

Early successes also attracted new group members: “After seeing the increase in the produce from members of the group I too joined, actually the benefit we have experienced is the increase in the yields.” A Butiama parent reported:

There’s a huge difference and I suggest that everyone should receive the training and work on them. Ever since I grew my own demonstration plot at home I had so much profit despite the climatic challenge.

Farmer Groups appear to drive meaningful outcomes for household food security in at least some cases. One Bunda farmer described his positive experience shortening the lean season:

There is a lot of advantages that I got after joining the group. I used to keep crops randomly but after learning through different seminars now I am saving food in the nice place. The food I used to eat for six months now I can eat for eight months.

Others were candid about remaining challenges to successful agricultural outcomes. A Bunda FGD respondent explained:

I personally find some challenges especially when trying to use the skills I learnt in the demonstration plot to my plot at home. The demonstration plot is small compared to my plot, this makes the cost of cultivating in my plot higher than the one estimated in the demonstration plot. Therefore you find it better to employ local means rather than the improved means we learnt.

Demonstration plots vary in size but it is noteworthy that PCI advocates for plots of 4,000 square meters to replicate real farm conditions as much as possible. Many respondents, across farmer and other group discussions, expressed concern that they could not appropriately manage drought conditions:

Another thing we would like to be taught is how to deal with drought. We are doing our best and the way we have received training or teaching, but we fail in production because of the drought. We would like the project to help group members to get the equipment that can help us get water easily.

Some FGD participants also complained about other community members stealing crops from demonstration plots, often to feed their livestock. The Regional Agricultural Officer emphasized that even after production is improved, storage practices are inadequate:

After harvesting, they use local storage facilities known as '*vihenge*'- primitive granaries. So, storage structure is still a problem because people do not have permanent storage where crops can be stored until the time to sell them especially during lean period.

To address this challenge, PCI has promoted Perdue Improved Crop Storage (PICS) bags to help farmers store varying harvest sizes.

In terms of farmer group contributions to school feeding program sustainability, schools' records/data presented above reflect overall low levels of farmer group contributions. Multiple respondents reported that farmer groups contribute 20 percent of harvests for school feeding based on PCI guidance. School records suggest that the 20 percent rule may not be consistently followed.

WE Groups

In addition to reporting high levels of program satisfaction, WE group members shared many personal success stories. One simply said:

When I was not in the group, I was broke but when I joined the group, I get capital and I progress myself economically.

Several participants referenced secondary health benefits of improved financial liquidity. One reported:

I didn't have a cow, but I bought one after joining the group and now my health is better because of the milk.

Two other members (in the same group) described an improved ability to provide nutritious food for their children:

My children used to drink white porridge, but now days they take millet, soya, sardine porridge that improves their health.

For example, me, if am lucky to get a good loan, I will buy meat that day and rice so that my children can be happy and later I will decide on how to use the remaining loan.

Many others described using loan funds for school expenses such as fees, clothing, and exercise books. One WE FGD respondent reported, "For me, when I borrow, I will buy school uniform for the student so that he can be able to go to school and the rest I will use for the household needs."

A Butiama member described how the group savings structure allowed participants to overcome the difficulties of saving at home:

My challenge is that I cannot save in the house because when I get a problem I will use it but for the group, we have to sit down and discuss, so it's better than saving on your own.

In addition to enjoying better savings outcomes, WE group members were able to make a diverse array of productive investments. One Butiama member described his or her experience:

I borrow money from [group name] because I am a small pastoralist. I have 10 cattle, so I borrow from the group to buy veterinary medicine that increases milk production which I will sell at a good price and be able to pay back the loan. So my cattle will benefit, and I will have paid back the loan.

Some groups had also added social funds and collective efforts to provide for members in the event of an economic shock:

The community kitty is there, and we also added a method like if a member is bereaved we go and help to collect firewood, cooking, fetch water, we were not taught that, but we just thought that we should have togetherness.

All WE group participants were certain that the groups would continue after PCI support was withdrawn. This exchange from a Musoma group is illustrative of many others:

Interviewer: Will your group continue to exist after PCI completes its work?

All: Yes.

Interviewer: How?

Respondent 1: Because we have been educated by PCI on how to save money and borrow so we will maintain even after PCI completes its project.

Interviewer: Anyone else?

Respondent 2: Because we already have the procedures and we have seen the benefits, so it will be easy for us to progress using the current procedures.

Despite laudable successes across savings activities, productive investments, improved care practices, and group sustainability, linkages to school feeding support were uneven. One District Planning Officer pointed out that, in his or her district, “these groups have succeeded because they help in the project towards providing food for education, and students with special needs. Also, the groups within the District have united and managed to build classes and one teacher’s office.”

A focus group in a different district reflected less close ties. An interviewer asked about connections between group activities and local education:

Interviewer: Do you see any relation between your group and the children’s education in school?

Respondent: Yes.

Interviewer: How?

Respondent: Because we are PCI group and PCI cares a lot about education that’s why in community discussions we discuss how to farm and take the produce to school.

Like many groups, they felt connected to the program because they discussed it. But when asked if as a group they actually contributed to school feeding in any way, they clarified that they did not. Some respondents gave the impression that school feeding activities were not central to the

group's mandate or the goals of its members. Other groups described closer connections. For example, a Butiama group member reported:

We cooperate with the school. If there is any challenge with the school we solve it. That is why we were able to buy them cups.

Another group member offered their *willingness* (see discussion of parental contributions above): "We are ready but the weather is not conducive for farming, but we are willing to contribute." The school store commodities management data compiled from the 50 sampled schools showed that just 0.2% of all new food contributions over an 11-month period came from WE groups. However, it is worth noting that only 15 of 45 treatment wards have active WE group activities. In addition, WE groups may positively influence community contributions by encouraging parent group members to contribute from their own households rather than as a collective. If individual WE group members contribute from their households, this is captured by PCI under general community contributions.

When asked directly about their level of responsibility to sustain school feeding, a group in Musoma that was designated as high performing responded as follows:

Interviewer: Is it a responsibility of the group to take onto the project when PCI's tenure comes to an end?

ALL: No.

Interviewer: Why not?

Respondent 1: What if I don't have a child at school and I am a member of the group.

Interviewer: Whose is it then?

Respondent: Parents'.

Interviewer: Why?

Respondent: Because the group has not attained funds to feed children.

Interviewer: Why is it not a responsibility of the group?

Respondent: We will be able to do that when we are a strong institution with reliable funds.

Interviewer: I would like you to give me answers that are relevant to this time, that if PCI is ended will the group take onto the project?

Respondent: We won't because we were not given directives; that we will be taking onto the project.

Interviewer: I need your opinion.

Respondent: It leads me to question myself on what good it is if you spend 2 million on feeding children then it is over.

Health and Hygiene Activities

The baseline report elevated a recommendation to monitor the maintenance and cleaning of SWASH infrastructure, particularly latrines, to ensure that they remained usable. Midline data collection did reveal that maintenance and cleanliness of latrines remained a concern. A lack of water at school for cleaning purposes was cited as a constraint in some cases.

A teacher in Butiama explained that he or she had gained knowledge that would improve students' health going forward:

I can say that I have learned a lot about health and how to protect the society from getting diseases. Before this, at schools there was no First Aid Kit but now students are being provided with the first aid service soon after injuries, severe headache and the environment is always super clean. Children were taught to wash their hands after using the toilets and before eating too.

Many focus group students elaborated on the importance of hand washing, regular bathing, wearing clean clothes and shoes, and maintaining a clean environment. But across qualitative data sources, no discernible distinction emerged between treatment and comparison student responses. Students from both groups repeatedly referenced the risk of cholera infection.

FFE program and comparison Serengeti schools also both reported sensitization activities related to puberty and menstrual hygiene management. For example, a teacher focus group in Butiama elicited the following characterization of health club activities:

In the health clubs students learn varieties of things. Just recently the teacher was teaching on the transition to the puberty age therefore the one who participates in the club gets to learn a lot about the puberty changes. This helps the students to understand what changes their bodies are undergoing the possible effects of the changes if the child does not take good care of himself.

PCI has also provided menstrual pads to schools to share with students. But a conversation with a group of teachers in Serengeti suggested that their efforts were not dissimilar, absent PCI support:

Respondent 1: As a health teacher, we provide education before and during the menstruation, so as to make them understand the situation better. However we provide some equipment to help them during the menstruation period.

Interviewer: Another else?

Respondent 2: Yes. The other thing is that we have pads, soap, underwear, and water for making sure they are good at all time.

One treatment District Education Officer observed that sanitary pads provided by PCI were in high demand and supplies were often exhausted. The ongoing availability of menstrual hygiene products should be factored into project close-out.

The evaluation did not focus on direct questions to probe health-oriented attitudinal and behavior change. This may be corrected in the endline survey by honing the instruments to better gather information on these areas

Schools and other centers of PCI intervention, such as WE and farmer groups, may also play a catalytic role in changing community norms that is difficult to observe directly. One Ward Education Officer observed:

In school a student might get education on health and he or she goes at home clean and says the teacher taught me this and that and when he or she changes he or she will automatically influence others to change.

The Sustainability Readiness Tool was administered to head teachers in all treatment and control schools at both the baseline and midline survey. At baseline, similar proportions of treatment (24 percent) and control schools (30 percent) reported teachers documenting health issues and health-related absences. At midline, 64 percent of treatment schools reported this practice compared to just 32 percent of Serengeti control schools.

The ongoing activity of WE groups in the communities may also contribute to sustained health outcomes through mutual support and information sharing related to maintaining a clean household environment. One group member observed:

We discuss environmental cleanliness, because some families don't have toilets and yet cleanliness is a very important thing. We encourage each one to have a toilet at home and maintain cleanliness to avoid diseases like cholera.

Social accountability may be as important as access to finance for improved latrines at home as community norms continue to shift. At the school level, PCI's role in facilitating dispensary connections may contribute to sustainable health-seeking behavior. One health teacher described an improved relationship with the local dispensary:

PCI has helped dispensaries which in turns helps us when we have shortage of health instruments we go to the dispensary and it is easier because we are connected already

Another health teacher echoed the reflections of multiple respondents that handwashing behaviors have noticeably increased.

It has improved like the education on washing hands maybe encouraging children to know the five places of washing hands when you come from using the toilet; before eating, when preparing food and when you want to eat food. What I see is that students have learned a lot.

Students regularly identified handwashing as a healthy behavior that was frequently practiced. In some cases, they cited a lack of soap at school.

Government Roles

The Mara Regional Education Officer said, "I believe that even when PCI are done with their job, the government will continue to sustain this project." However, at present, there is a lack of consensus on the roles and responsibilities of local, district, and regional government actors to sustain school feeding and other activities after 2021.

Many respondents identified a more central role for government officials to ensure that school feeding continues, but the roles they envisioned varied substantially. A number of interview subjects suggested a contribution mandate. An agricultural extension officer said, "the only way is for the government to make a statement requiring them to contribute or else nothing will be done."

When a Ward Education Officer in Musoma was asked about ways to expand community involvement in school feeding, he or she answered:

There are ways, but I think the government should manage that or impose a formal rule because some community members refuse to participate for no reason though they understand the necessity. Others refuse to participate because they know there are no any laws that force them to participate.

In fact, a remarkable number of qualitative respondents shared the view that a punitive mandate was the only way to mount a program without PCI support. Yet the same Education Officer went on to observe that some parents “tell their children to go to school where there is food rather than remain at home where there is starvation.” This tacit acknowledgment of household food insecurity stands in contrast with the assessment that parents refuse to participate, as stated above, “for no reason.”

A Village Chairperson who directly attributed the lack of local school feeding program contributions to poor weather conditions and household poverty explained, “we have given requests and if the government listens then it will make it mandatory for everyone to contribute.”

At the regional level, there was evidence of an unwillingness to acknowledge the lived experience of rural households. An interview with the Mara REO yielded the following exchange:

Interviewer: Some communities say that they are poor and can’t even provide food at home so how can they contribute food to school, what do you say about that?

REO: Actually, in reality they can manage, the only thing is that they have negative attitude that everything needs to be done by the government or PCI. It’s just a matter of educating them that they can give 20kgs of maize and 15kgs of beans which am sure they can manage and this way they can join and support the program.

At the local level, teachers saw a greater role for village leadership to spur community contributions. When asked if parents would contribute food after PCI withdrew support, one teacher stated:

When the parents have the motivation to contribute food it may continue but if the village government does not follow up it will surely fail.

But other teachers reported that village governments wouldn’t take the project seriously or prioritize it after PCI withdraws support.

Many other stakeholders felt that the role of government was not to mandate community contributions but rather to provide them itself, or at least to mobilize collective community resources for school food production. One District Education Officer reported:

I think the program should make progress and work with the ministry so as municipalities can have a budget to PCI programs to continue functioning and we stop only depending on parents to contribute. For us as a municipal to make a follow up on these activities we need a budget just like EQUIP. After PCI has reduced to contribute for food, they [parents] have also failed completely to do so. Therefore, the government should set a budget to enable the proceeding of the project.

A District Planning Officer from a treatment district was particularly thoughtful on the topic of feeding program sustainability:

DPO: We are setting aside a budget. We would like to support school feeding in our budget and this will be set aside from the following departments; agriculture department, health department and education department.

Interviewer: If I enquire further, how much will you put in the school feeding project according to your plan?

DPO: That is why I have told you that our resources are small for example; in the last budget we put 8,000,000/= TZS in this budget we plan to put 16,000,000/= if I'm not wrong, that is for food and we are targeting the small children but as we continue we find it is a good project then we will find a way of putting in as much money as we can.

He or she then elaborated on plans to develop school farms and common lands to grow food for schools:

Our biggest effort is channeled towards using the school farms and those in the village. So when the weather is conducive we will try our best to make sure that the school produces enough food from their farms because we have experience on what the students eat. I believe that if we will be able to harvest in school then we will be able to feed the students. That is the first strategy that we have.

We also have farms in our village which are not owned by anybody so we will also involve them and that is why I said that the village executive officers should be involved as part of this project and the village chairman should be involved in training them and as the co-coordinators we inform them to select a specific area where farming will be done for the school and by doing that I think we will get somewhere.

The organization also works with groups through women empowerment where we have formed our own groups where these groups rent out farms and cultivate. Our strategy is to also strengthen those groups and also link them with the village executive so that he can be aware of the groups and supervise them so that when they cultivate the land it should not be for the group but for the community in that when they harvest part of the harvest should go towards the school feeding program.

This Officer had a clear role in mind for PCI to support local ownership efforts:

PCI has a focal person even in the ministry of education; we are just implementers. In order for us to have the power to include it in the budget they have to talk to the ministry of education. Either the ministry of education can give that money as its own budget the way it brings capitation directly to the school or it can give us a directive that a certain amount is from the main government for the food for education project. The money that we have to budget for with the councilors is very minimal and actually cannot be enough to do everything like what PCI has done but we will try our best to get money even if it is subsidization in the village and school farms.

Like parents, government officials interviewed expressed a willingness to allocate resources to school feeding while also communicating that the feasibility of doing so was not always straightforward. Non-government actors shared the perspective that the public sector should play a greater role to sustain school feeding. A suggestion for local action was offered by a farmer in Bunda:

Interviewer: Do you think there is another person to be involved in the development of the food project at school apart from parents and other group members?

Respondent: Village leaders.

Interviewer: How?

Respondent: The village should have a farm for schools only to help those students with the food.

A WE group member in Butiama, among others, repeated this idea: “The government leaders could help people by allowing them to use open areas for farming.”

Another WE group member in Bunda offered a different version of a public mandate: “I think rules should be set that children must get food while in school. And the government should intervene in this.”

National stakeholders within the Government of Tanzania have asked PCI to lead the development of School Feeding Guidelines. Clarifying the roles within various levels of government, particularly identifying possible funding mechanisms, should be considered.

Section 8: Major Conclusions and Program Recommendations

Beneficiaries are satisfied with program delivery.

PCI's Food for Education program has undertaken a broad array of multisectoral activities across a large geographical area. From pedagogical interventions, to school feeding and SWASH activities, to WE and farmer groups, beneficiaries expressed appreciation for FFE program interventions. A wide variety of stakeholders report good program management, effective and communicative relationships with schools and government officials, and the alignment of program efforts with pressing community needs.

Literacy is difficult to influence in highly resource-scarce environments.

The program's highest impact goal is to improve the literacy of Mara children. A clear or consistent pattern of literacy gains is not detected in the midline evaluation. There are significantly greater improvements in meeting the national benchmark at the grade 2 level, but there are significant declines in grade 4. Learning outcomes are the result of numerous complex factors and improving one or two of them may not be enough if other key factors are not improving. Many contextual factors make literacy a difficult variable to influence over decades; a time horizon of two years to achieve impact was always extraordinarily ambitious. Resource constraints in Tanzanian schools, and the adverse outcomes they lead to, are well and widely documented. It is not unusual for classes to count over 100 enrolled students, and average class size has increased between baseline and midline. The caliber of literacy instruction is often inadequate, and teachers regularly transfer away from rural schools for a variety of reasons. Relatively low levels of literacy in the community, combined with an uneven valuation of education in general, undermine program efforts. In the baseline survey, 18 percent of parents reported that they could neither read nor write: children living in illiterate home environments cannot access support in their studies to reinforce classroom lessons.

WE and Farmer Groups and school gardens demonstrate positive, sustainable impacts.

The WE Groups and Farmer Groups are a sustainable success from the project. Current and former WE group members comprised over one-fifth of parent respondents. Focus group members had highly positive things to say about WE group impacts on their economic stability, and income. Some have also organized to make school improvements.

Twelve percent of parents reported being current farmer group members and nine percent reported past participation. Farmer group participants reported multiple benefits from joining, including improving yields and sustainable practices. Farmers discussed improved planting practices, crop rotation, fertilizer use, cover crops, and increased weeding as factors driving increased yields. They also reported that simultaneous participation in WE groups was complementary to their agricultural activities:

I have experienced a benefit that is beyond agriculture and it is through the savings and shares that we buy. During dividends I was able to spend my share on building my house.

Multiple respondents also remarked upon the sustainability of school gardens. When asked what he learned from PCI trainings, one agricultural extension officer observed:

There is so much, for instance how to prepare permanent gardens in schools. Permanent gardens differ with the ordinary gardens we knew that comprised of only vegetables; the former include fruit trees that even during dry season when the vegetables are extinct the trees bring a sign of life to the garden.

Evidence suggests these groups have been successful in making long-term improvements that will extend beyond the program.

Health, nutrition, and sanitation indicators are not negative, but could be more insightful. While there is little negative or positive change on the included indicators on healthy activities, it may be a result of the indicators themselves. Collecting more detailed or different data on health, nutrition, and sanitation may provide greater insights in the final evaluation.

In response to the Sustainability Readiness Tool at midline, just 8 percent of treatment schools reported that ward-level officials actively meet with parents to teach pro-education, health, SWASH, and nutrition messages and routinely monitor at the household level. This compares to 7 percent of sampled control schools reporting such parental engagement by ward officials. This low level of engagement suggests that health-focused behavior change may more quickly influence community practices when school-based health messaging is amplified by local government.

Government engagement is positive but public sector ownership is lacking.

Sustained government human and financial resource support of the program, as with education funding in general in Mara, is lacking. While government officials have positive things to say about the program and express a general willingness to support it, there is little clarity as to how ongoing provision of school food would sustain itself without the project. Government officials primarily look to communities to sustain the food necessary to continue the program, but given the communities are net importers and many lack sufficient foodstuffs in the home, it is unlikely to self-sustain.

Heavy emphasis on parental contributions risk making vulnerable households worse off.

In some cases, the solicitation of parental contributions may worsen the welfare of the most impoverished households. Mara Region is characterized by chronic food insecurity. While most acute during the annual lean season from October to January, many poor households face year-round food shortages. In the midline survey, 41 percent of treatment school parents reported that their children ate two meals per day. Twenty-three percent of those parents said that they had gone to bed hungry in the previous week, and 21 percent said their children had as well. A Serengeti parent lucidly described the strain of hunger throughout local communities:

Hunger even dissolves marriages, you find the woman running away from her husband just because she doesn't get anything to eat. It is easier for an adult, but a child will not understand staying hungry.

Many parents expressed appreciation for school feeding activities and a general *willingness* to contribute food but lacked the *ability* to do so (see Section 7). A pervasive trend throughout the qualitative interview was clear that encouragement to contribute has in many cases become coercion. When one Village Chairperson was asked why they planned to use coercion with parents, they responded:

Up till now it's through coercion because even paying the four hundred is an issue and it reaches a point that students are sent back home to collect the money. Sometimes I am forced to follow up on the parents because then we cannot pay the cooks. When the children are sent back then they bring the money but if we relent, the parents can stay for months without paying and we cannot pay the cooks and it is a big challenge... To say the truth if one is asked to bring two tins f[sic] four sacks people will end up taking each other to courts something we don't want an it will be a big challenge if the project will continue.

Most alarmingly, a focus group discussion in Bunda surfaced reports of conflict that had already occurred. A farmer reported that some people had been “attacked” at home and forced to contribute food, although this translation may have meant food forcibly taken. The discussion continued:

Then leadership has taken the decision to go inside of the house to take food he had prepared even for the past two years is being taken away. The second thing I see in other villages like Malambeka you know PCI is there but they kept talking to each other who will help our children and who did not accept that their children are going home without worry and nobody say today I am coming to arrest you.

People contribute because of the presence of the project that's why people contribute, but if the project will come to an end there will be a problem. Once PCI will leave there will be no supply of the food and children will not get food.

Participants of one adult focus group discussion also explained that students are the ones to bear adverse consequences when parents don't comply:

Respondent 1: It is not right to kick a student out of school for reasons of payment. He should be given a letter to summon the parent who would be told what the problem is.

Respondent 2: It is the parent who is supposed to pay for the food of the child because he is the one who earns while the child has no income, and if forced to gain some the child ends up dropping out of school for instance due to going fishing.

A Ward Education Officer was likely well-intentioned with a program design adjustment to increase attendance on days when meals are not served, but the outcome is certainly harmful to vulnerable students traveling long distances:

The feeding program has contributed a lot, good number of students is attending school. Some students attend classrooms only on Monday and Tuesday because there is meal, but on Wednesday, Thursday and Friday they do not attend. Recently we have decided to offer food by surprise in such a way they cannot know when food is available, since if the time table remain predictable only days when food is available we witness good attendance.

The interviews and discussions suggest the presence of community-led enforcement that, while well-intentioned, appear to have adverse effects on intended beneficiaries.

In sum, the project is satisfactorily delivering an array of school feeding and associated management interventions, alongside complementary programming in instructional quality, health, sanitation, agricultural productivity, and household financial services. Program implementation is reviewed positively by government officials, school leadership, parents, and students alike. Given a challenging operating environment replete with resource constraints, it is not wholly surprising that statistically significant impacts on attendance, attentiveness, instructional quality, literacy, and behavioral knowledge are not evident in quantitatively representative data. While community members express support for the continuation of school feeding, they face stark challenges in their current ability to furnish the requisite commodities to sustain the program beyond PCI.

Recommendations for Program Management

- Recalibrate expectations for literacy outcomes and impacts given the many prevailing challenges to quality education in Mara Region. Consider the introduction and emphasis of intermediate measures of progress in both the endline evaluation and also in related future programming to better trace causal pathways. In the endline evaluation, focus analysis on the averages and distributions of EGRA scores rather than benchmark cut-offs which can mask meaningful shifts. Consider also focusing attention on student engagement, access to materials, and other conditions which mediate literacy outcomes.
- Make efficient and evidence-based choices with program resources, including human resources, in the remaining two years of implementation. WE groups and farmer groups both emerged as successful and sustainable models. While WE groups have increased in number over time, there still appears to be ample scope for new groups before communities are saturated. PCI is best positioned to assess group demand and member potential.
- Several options exist to make WE groups more effective in dry months. First, cyclical payouts may be scheduled for the time of greatest need if that is not standard practice. Second, group members with diverse livelihoods are better able to cross-subsidize.²³ PCI may consider actively recruiting participants with off-farm income sources. Given the predominance of farming, it could also explore the introduction of off-farm income generating activities. This possibility is already under discussion among both PCI technical staff and group members. Third, PCI may consider working with groups to test approaches aimed at saving more collective resources for the lean season. For example, a portion of all savings or a portion of loan interest could accrue to a fund that can only be tapped during the lean season. Alternatively, seasonal loan caps or variance in month-specific interest rates could be explored to encourage larger investments during

²³ For example, see (Cassidy & Fafchamps, 2019)

less constrained periods. In some cases, PCI may also explore building connections between established WE groups undertaking business activities and formal financial institutions in the greater community.

- As the baseline evaluation previously recommended, pivot focus from a household food donation model to a community food production model, guided by the input of many evaluation respondents across both treatment and comparison communities. In addition to focusing efforts on school garden and demonstration plots, solicit input from parents, communities, and local leaders about other public or shared lands that could be utilized for school food production. When determining a long-term food production strategy, it cannot be forgotten that Mara Region is a net importer of food: No plan that focuses on local production – be it through school gardens or parental contribution – is likely to succeed alone: Importing food into the region, be it through markets or government provision, will be an integral part. School budgets to purchase food commodities should remain a medium-term goal.
 - Speak with beneficiaries to find out why Phase I and Phase II schools, which had agricultural demonstration plots during the baseline survey of Phase III, no longer do so. It may be that farmer groups have graduated to larger land areas in other locations, a positive advancement, but it may also be that collaborative efforts in a central location proved too challenging to sustain. In the latter case, the examination may yield fruitful lessons applicable to broader program close-out.
- Communicate to program staff, partners, and government at all levels that community food contributions are voluntary. The program was always designed to incorporate voluntary contributions, but this intent has not always been clearly communicated at the school and community level. Below-target levels of parental contributions should never be met with punitive responses. Engage partner schools to ensure that coercive tactics do not result in children withdrawing from school.
- Work closely with government actors at all levels to jointly devise non-coercive sustainability plans supported by public monetary and non-monetary resources (such as the provision of public lands for school food production).
- Explore lessons for program sustainability drawing from the experiences of Phase I and Phase II schools. On a host of metrics, strong gains for Phase III schools coincide with declines or stagnation of activity levels among earlier phase schools. If local ownership of activities has not taken root as anticipated in communities with active programming since 2010 or 2014, seek understanding of the underlying reasons.
- In its influence of national school feeding policy, PCI would add value to communities across Tanzania by advocating for clear and distinct budgetary allocations to cover commodity procurement, management, and preparation costs (including kitchen infrastructure) with limited or no transferability to other education sector expenditures. It should also focus efforts on disseminating both successes and challenges that have

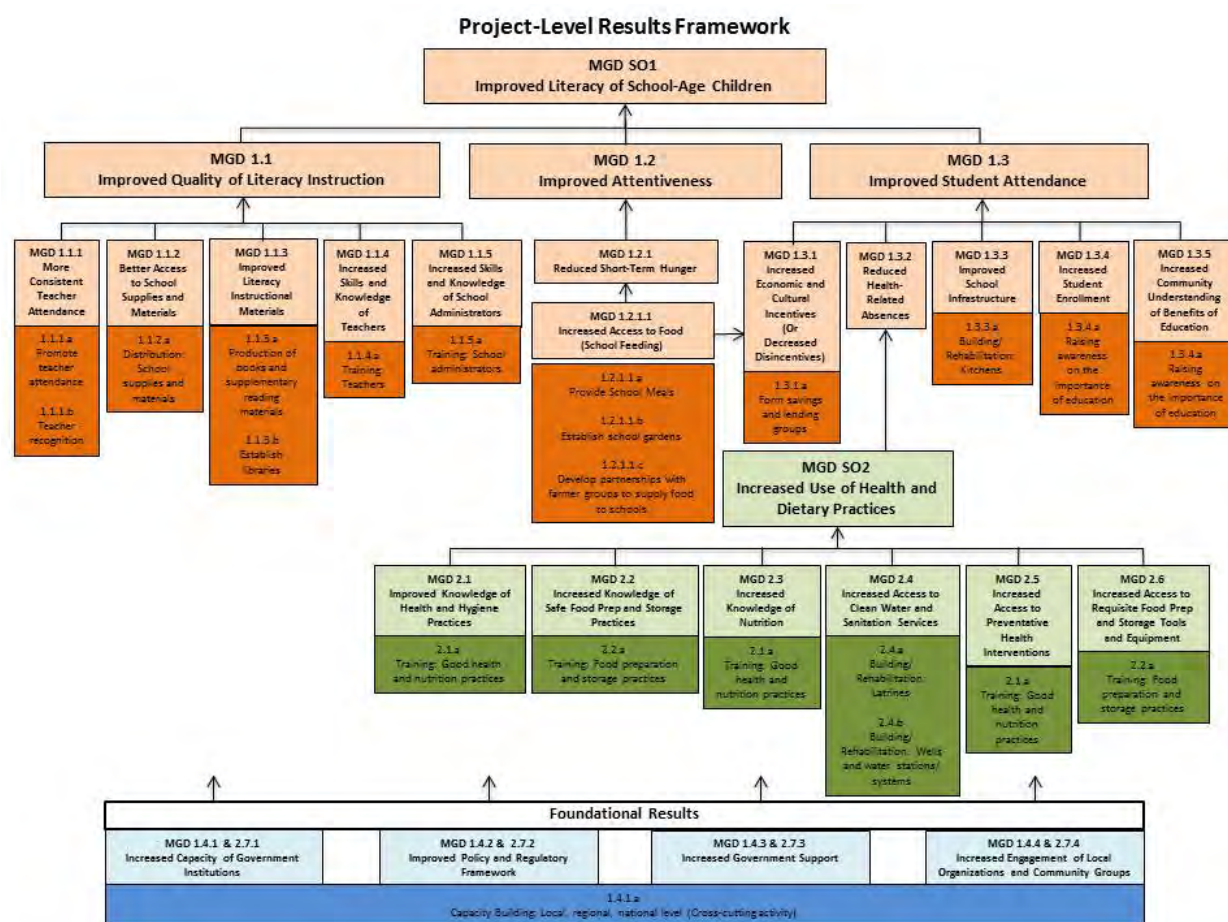
emerged in the examination of its own programming to support the Government of Tanzania's school feeding momentum while avoiding potential pitfalls. In turn, the Government should consider a phased roll-out of school feeding activities to hone best practices and minimize unintended negative outcomes.

Bibliography

- Adams, J., Bartran, J., Chartier, Y., & Sims, J. (2009). *Water, Sanitation and Hygiene Standards for Schools in Low-cost Settings*. Retrieved from https://www.who.int/water_sanitation_health/publications/wash_standards_school.pdf
- Cassidy, R., & Fafchamps, M. (2019). *Banker My Neighbour: Matching and Financial Intermediation in Savings Groups*. Retrieved from <https://web.stanford.edu/~fafchamp/cassidy.pdf>
- CDC. (2017). *CDC Notes from the Field: Ongoing Cholera Epidemic - Tanzania 2015-2016*. Retrieved from <https://www.cdc.gov/mmwr/volumes/66/wr/mm6606a5.htm>
- FEWS NET. (2013). *East Africa Seasonal Calendar: Tanzania*. Retrieved from <http://fews.net/east-africa/tanzania/seasonal-calendar/december-2013>
- FEWS NET. (2018). *Production and Trade Flow Map: Tanzania Maize*. Retrieved from <http://fews.net/sites/default/files/documents/reports/Tanzania%20Maize%20PTF%20Map%20August%202018.pdf>
- Government of Tanzania. (2012). *Population and Housing Census*.
- Government of Tanzania. (2016). *Tanzania Demographic and Health Survey and Malaria Indicator Survey 2015-16 Final Report*. Retrieved from <https://dhsprogram.com/pubs/pdf/FR321/FR321.pdf>
- Government of Tanzania. (2018). *Child Poverty in Tanzania*. Retrieved from <https://www.unicef.org/esa/sites/unicef.org/esa/files/2019-04/UNICEF-2018-Child-Poverty-in-Tanzania.pdf>
- Government of Tanzania. (2018). *Education Sector Development Plan: 2016/17 - 2020/21*. Retrieved from <https://www.globalpartnership.org/sites/default/files/2019-04-gpe-tanzania-esp.pdf>
- Government of Tanzania. (2018). *Global Initiative on Out-of-School Children*. Retrieved from <https://www.unicef.org/tanzania/media/621/file/Tanzania-2018-Global-Initiative-Out-of-School-Children-Mara-Region-Report.pdf>
- Population Health Methods. (2019). *Different-in-Difference Estimation*. Retrieved from <https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation>
- Sambhodi Limited. (2017). *Baseline Assessment Of The MCGovern-Dole Food for Education Programme*.
- Southern African Development Community. (2018). *SYNTHESIS REPORT on the State of Food and Nutrition Security and Vulnerability in Southern Africa*. Retrieved from 2016/17 Annual Agriculture Sample Survey Crop And Livestock Final Report
- Tanzania National Bureau of Statistics. (n.d.). *2016/17 Annual Agriculture Sample Survey Crop And Livestock Final Report*. Retrieved 2019, from <https://www.nbs.go.tz/index.php/en/census-surveys/agriculture-statistics/57-2016-17-annual-agriculture-sample-survey-crop-and-livestock-final-report>

- UN Children's Fund. (2019). *Bulletin: Cholera and AWD Outbreaks in Eastern and Southern Africa*. Retrieved from Relief Web: <https://reliefweb.int/report/mozambique/bulletin-cholera-and-awd-outbreaks-eastern-and-southern-africa-regional-update>
- UNICEF. (2018). *Education Budget Brief 2018, Tanzania*. Retrieved from <https://www.unicef.org/tanzania/media/1236/file/UNICEF-Tanzania-2018-Education-Budget-Brief.pdf>
- UNICEF. (2018). *Tanzania Mara Region Case Study Report*. Retrieved from <https://www.unicef.org/tanzania/media/621/file/Tanzania-2018-Global-Initiative-Out-of-School-Children-Mara-Region-Report.pdf>
- UNICEF. (2019). *UNICEF Education: What We Do*. Retrieved from <https://www.unicef.org/tanzania/what-we-do/education>
- USAID. (2018). *US President's Malaria Initiative: Fighting Malaria and Saving Lives Tanzania*. Retrieved from https://www.pmi.gov/docs/default-source/default-document-library/country-profiles/tanzania_profile.pdf?sfvrsn=20
- World Bank. (2017). *Implementation Completion and Results Report: Secondary Education Development Program II APL 1 Project*. Retrieved from <http://documents.worldbank.org/curated/en/336821499446341715/pdf/ICR00004211-06272017.pdf>
- World Bank. (2019). *World Bank Open Data*. Retrieved from data.worldbank.org
- World Food Program. (2019, April). *Southern Africa Region Monthly Food Price Update*. Retrieved from https://docs.wfp.org/api/documents/WFP-0000105207/download/?_ga=2.217533266.536034985.1566618310-1025721452.1566618310

Annex A: Results Framework



Annex B: Data Collection Tools List

Quantitative Data Collection Tools

Classroom Observation Tool

Grade 2 and 4 Interview Tool

Grade 6 and 7 Interview Tool

Head Teacher Interview Tool

Parent Interview Tool

School Sustainability Assessment

Qualitative Data Collection Tools

Agricultural Officer Interview Tool (KII)

District Focal Person Interview Tool (KII)

Education Officer Interview Tool (KII)

Literacy Teacher Interview Tool (KII)

Village Chairperson Interview Tool (KII)

Farmer Group Tool (FGD)

Head Teachers, Library, Garden, Store, and Health Teachers Tool (FGD)

Parent Group Tool (FGD)

Student (Grade 7) Group Tool (FGD)

WE Group Tool (FGD)

Annex C: Distribution of Qualitative Extracts

C.1: Program Relevance and Beneficiary Satisfaction

Code	Key Informant Interviews			Focus Group Discussions		
	Extract Count	Cases	% Cases	Extract Count	Cases	% Cases
Program Meets Community Needs	40	22	82%	19	13	35%
Coordination with Government or Other Programs	40	18	67%	7	5	14%
Beneficiary Satisfaction	40	21	78%	75	26	70%
Beneficiary Dissatisfaction	13	8	30%	35	11	30%
Comparison School Activities	NA	NA	NA	20	5	14%

C.2: Program Implementation and Effectiveness

Code	Key Informant Interviews			Focus Group Discussions		
	Extract Count	Cases	% Cases	Extract Count	Cases	% Cases
Student Attendance	36	15	56%	47	11	30%
Student Attention	9	7	26%	23	18	49%
Instruction/Teacher Quality	30	14	52%	27	19	51%
Literacy	9	9	33%	16	11	30%
Infrastructure, Facilities, Supplies	29	15	56%	48	23	62%
Parental Engagement	52	20	74%	43	23	62%
School Gardens	18	10	37%	28	14	38%
Farming Improvement	16	6	22%	15	8	22%
Nutrition	16	9	33%	16	13	35%
Health and Hygiene	16	8	30%	40	17	46%
WE Groups	3	3	11%	25	5	14%
Gender	26	11	41%	57	27	73%
Behavior Change	39	19	70%	27	13	35%
Policy Change	2	2	7%	0	0	0%

C.3: Program Sustainability

Code	Key Informant Interviews			Focus Group Discussions		
	Extract Count	Cases	% Cases	Extract Count	Cases	% Cases
Contribution Sustainability	102	25	93%	129	28	76%
Collaborative Relationships	9	7	26%	10	8	22%
Government Sustainability	71	22	82%	26	21	57%
Policy	29	13	48%	10	10	27%
School Management/Teaching Capacity	13	8	30%	16	11	30%
Education Outcomes	1	1	4%	2	2	5%
Farmer Group Value Add	17	6	22%	24	10	27%
WE Group Value Add	3	3	11%	31	6	16%
Challenges	107	25	93%	153	35	95%
Feedback/Suggestions for PCI	111	25	93%	121	29	78%

Annex D: Midline Data Collection Timeline

All quantitative and qualitative midline data was collected in May 2019. The following table outlines key dates throughout the preparation and fieldwork period along with the division of labor between the data collection consultant (IPSOS Tanzania) and the data analysis consultant (Trembley Consulting).

Activity	Task	Responsible	Date(s)
Government approvals	Consultant prepares and submits IRB materials	IPSOS Tanzania	Prior to April 30, 2019
	Consultant submits study information to NBS, COSTECH, PORALG for approval	IPSOS Tanzania	Prior to April 30
Data Review	Review relevant program materials	Trembley Consulting and IPSOS Tanzania	April 2019
Tools Development	Revise all quantitative and qualitative data collection tools	Trembley Consulting	April 26
Enumerator training	Enumerator training on all tools	IPSOS Tanzania	April 29 through May 4
	Pre-testing	IPSOS Tanzania	May 3
	Make changes to tools as needed	Trembley Consulting and IPSOS Tanzania	May 4
	Enumerators tested on EGRA to select top performers	IPSOS Tanzania	May 4
Data Collection	Quantitative and qualitative data collection conducted	IPSOS Tanzania	May 6 through May 24
	Submit final data collection deliverables and report to PCI	IPSOS Tanzania	July 29

Annex E: Methodology for Calculating Minimum Acceptable Diet (MAD)

The calculation of students having a minimum acceptable diet is based on subject responses of consumption across seven food group categories over a 24 hour recall period. Students are asked about each meal they ate in the period and what the meal was comprised of. They are found to have a minimum acceptable diet if they consumed at least four of seven distinct food groups and also ate three or more meals.

Food Group Categories for a Minimum Acceptable Diet

Grains, Roots, and Tubers (Maize, rice, wheat, sorghum, millet, bread, biscuits, cookies or any other foods made from maize, rice, wheat, sorghum, millet; white potatoes, white yam, white cassava or any other food made from white potatoes, white yam, white cassava)
Legumes and nuts (Beans, dried peas, lentils (Dengu), groundnuts, green beans (Choroko), njugu mawe, Kunde, seeds or foods made of these)
Dairy products (Milk, cheese(Samli), yogurt or other milk products)
Fish or Meat (Fish, sardines, furu, beef, pork, lamb (Nyama ya kondoo), goat, rabbit, chicken, duck, other birds, kumbi kumbi (small insects), senene (grasshopper))
Eggs (Eggs from chicken, duck or any other egg)
Vitamin A rich foods (Ripe mango, cantaloupe, ripe papaya, okra, dried peach, Ethiopian mustard and other leafy greens, 100% fruit juice from any of these, pumpkin, carrot, squash, sweet potato)
Other Fruits and Vegetables (broccoli, eggplant)

Annex F: Recommendations for Endline Survey and Evaluation

- Ensure that the evaluation process begins at least four months before data collection activities are to commence.
 - Reduce scope of instrumentation across both quantitative and qualitative field work.
 - Reduce sample sizes of qualitative field work.
 - Compile program staff input on instrumentation changes at least two months before data collection activities in order to finalize tools no later than six weeks prior to field work.
 - Allocate more training time for qualitative interviewers. This portion of data collection deviated significantly from the tools provided at midline. More time would also ensure quality translations into Swahili. Many parents, for example, were asked about the “losses” they incurred as a result of educating children when the translation should have been “costs.” In one group, a respondent corrected this and the conversation followed the intended direction. But in most focus groups, this led the conversation in an unintended and unhelpful direction.
 - Improve the quality of English language translations of qualitative discussions. Many were very challenging to make sense of.
 - In Serengeti comparison areas, ensure that both quantitative and qualitative respondents understand that they are not being considered for new program activities. In some focus groups, this was not clear to participants.
 - For EGRA score analysis, focus on students’ average scores rather than the proportion who surpass a benchmark cutoff according to national standards. Between baseline and midline, we observed large magnitude declines in benchmark attainment which corresponded to much smaller declines in average scores (meaning that many students are clustered right around the benchmark).
 - For household hunger questions, remove codes for seldom, often, etc. and just code count of days in the last week.
 - Link school feeding activity measurement to attention in class measurement
 - For pedagogical practices, better tailor classroom observation tool to specific practices promoted by the program. Seek out other ways to operationalize variables capturing behavior change.
 - Consider using item-based response analysis to have a better understanding of variation of improvement.
 - Change learning assessment indicators from number of students meeting national benchmarks to improvements in average scores.
 - Change attendance indicator from “students with attendance rate above 80 percent” to improvements in average attendance rates.
- Base indicator calculations on the descriptions herein and in the Stata code provided.

- Revamp questions on school facilities to include those added at midline as well as independent inspection to meet the USDA MGD definitions of improved water sources and sanitation facilities.

Annex G: Original Terms of Reference

Terms of Reference for Food for Education III Midterm Evaluation Data Analysis & Reporting

1. Introduction

Since 2010, PCI has been implementing United States Department of Agriculture (USDA) funded McGovern-Dole Food for Education (FFE) programs in Tanzania. Phase I was implemented from September 2010 to December 2014, and phase II was implemented between September 2013 – December 2016. PCI is currently implementing phase III from October 2016 – December 2021. The project aims to deliver a high quality, sustainable program designed to improve literacy of school-aged children in three districts in the Mara Region: Bunda, Butiama, and Musoma. PCI works closely with national, regional, district, and village governments to implement a high-quality education program in a total of 231 schools, which includes 156 schools that participated in phases I and/or II, and 75 new schools. A logical framework outlining the project's theory of change is included in Appendix A.

Tanzania is midway through the five-year phase III implementation period, and PCI is soliciting a consultant to conduct data analysis and reporting for the mid-term evaluation of the program. PCI is concurrently hiring a local consultant (hereafter referred to as “data collection consultant”) with significant experience in Tanzania to lead and execute the data collection component of the evaluation with support from PCI). The consultant hired to support the analysis and reporting will be hereafter referred to as “evaluation consultant.”

2. Program Description

2.1. Key Activities and Strategies

PCI's FFE program in Tanzania are focused on achieving desired results under two strategic objectives and three key intermediate results, specified in the Results Framework presented in Annex 14.3.

- Improved literacy of school-age children through improved quality of literacy instruction, improved student attentiveness and improved student attendance
- Improve literacy instruction quality by promoting teacher attendance, providing teacher recognition and school supplies, establishing mini-libraries and classroom reading corners, distributing reading materials and training teachers and school administrators;
- Improve attentiveness and reduce short term hunger by providing school meals and increasing farmer agricultural productivity and production, in turn contributing to locally-owned and sustained school feeding programs;
- Improve student attendance by forming savings and lending groups that provide financial stability to keep children in school, providing training on better health and nutrition practices to decrease health related absences, increasing community understanding on the benefits of education, and in Tanzania, promoting parent engagement of children's education through Education Cascade Groups (ECGs);
- Increase knowledge of safe food preparation and storage practices by training relevant stakeholders, and increase access to food preparation and storage tools and equipment through the distribution of food and preparation equipment;

- Increase access to clean water and sanitation services through rehabilitation or construction of latrines and water systems that are maintained through active school WASH infrastructure committees;
- Increase access to preventative health interventions through capacity strengthening and training;
- Increase community understanding of the benefits of education through Education Cascade Groups; and
- Increase local food contributions to schools through improved agricultural productivity in farmers groups, as well as improved food storage and handling practices.

2.2. Beneficiaries

In collaboration with Musoma, Bunda and Butiama Districts, the project was initially implemented in 103 and 134 government primary schools for phase I and II, respectively, in the Mara region of Tanzania. Due to increased enrollment rates, Musoma District Council has since split 22 schools into A and B, hence making a total of 156 schools by the end of FFE phase II. During the third phase, (FFE III), there are a total of 231 schools, which include all FFE I and FFE II schools. Currently, FFE III activities directly benefit 168,160 primary school children and 3,266 adults (including 2,313 teachers, 341 volunteer student aides (VSAs), 357 parent cooks, and 255 security guards) who participate in the program at the school level. Household and community-level interventions such as Women Empowered (WE) and supporting farmer groups directly benefit 7,249 and 2,175 adults, respectively. Additionally, a focus on local capacity strengthening directly benefit 145 Government of Tanzania (GoT) officials and approximately 231 school administration members. Overall, an estimated total of 442,743 beneficiaries will benefit from program activities, including 287,938 direct and 154,805 indirect beneficiaries over the life of the project.

2.3. Intervention Communities

Tanzania's FFE III project operates in three districts in the Mara Region: Bunda, Butiama, and Musoma. Communities in the project areas are primarily rural and most commonly rely on farming and fishing as key livelihood activities. The baseline evaluation also collected data from the neighboring Serengeti district as a comparison group, which has similar demographic characteristics to project districts.

3. Evaluation Scope and Objectives

The mid-term findings will improve PCI's understanding of its performance and help the team devise an appropriate plan to make mid-term adjustments and improvements on program activities and strategies to ensure optimal outcomes of desired results. The midterm evaluation will be conducted by an external evaluation team and has the following purposes:

- Assess the relevance of the project strategy and approach, and the validity of assumptions made during project design;
- Measure progress from baseline, including the effectiveness, efficiency and timeliness of interventions in achieving established targets, identify enablers and constraints to progress, and provide recommendations for mid-course corrections to strengthen performance;
- Assess sustainability efforts to date;

- Document lessons learned, challenges and unanticipated effects;
- Provide recommendations for areas of focus for the final evaluation, including review and strengthen data collection systems and metrics.

The scope of the midterm evaluation will include the entire project, and in collaboration with PCI, project partners, beneficiaries, USDA, GoT/GoG Ministries of Education, and other stakeholders. PCI will solicit the input of project staff, partners and USDA on the purpose and scope of the midterm evaluation during the planning stages and adjust the TOR and evaluation activities accordingly.

3.1. Evaluation Questions and Learning Objectives

The consultant will review existing baseline data collection tools and results in collaboration with PCI staff to collect data that will allow PCI respond to the following learning questions.

Program Design, Implementation and Relevance

- To what extent is the results framework and program approach valid for the local context?
- Which program strategies implemented have been more effective and relevant in contributing to program results? Which were the least effective and least relevant?
- Does the project design, structure, logic or management need to be modified to facilitate the achievement of desired outcomes and impact?
- How has implementation and accomplishment of program activities been carried out at the technical /programmatic, organizational and contractual levels?
- To what degree have the needs of the beneficiary population been included in the design of the program?
- Where similar activities are occurring within the same project area, how well does the project coordinate and complement existing programs to ensure efficient and effective use of resources and avoid duplication?
- To what extent did PCI address the baseline evaluation recommendations and lessons learned from the FFE phase II programs in the respective countries?

Sustainability

- How well have the projects developed and maintained collaborative relationships with appropriate government agencies, local partners, USDA and other key stakeholders?
- What efforts has PCI made to transition school feeding activities to local actors (local government, private sector, federal government agencies, parents)? How successful have those efforts been?
- What progress has the project made to ensure the sustainability of adequate human, financial, infrastructure, capital and other resources?
- To what extent have partners, governments, and local communities internalized and implemented efforts to sustain school feeding activities, such as providing in-kind support for school feeding activities?
- What specific local, regional, governmental, and school stakeholder institutional capacities have been developed to foster host-country engagement and accountability. What additional capacities should be developed for which stakeholders?
- What elements or considerations should be incorporated into public policy in the local governments to sustain a school feeding program?

- Does the participation of mothers and/or other family members in the WE groups contribute to the sustainability of the program and program outcomes?

Outcomes and Impact

- Has the program achieved the expected outcomes for literacy, teaching quality, student attention and attendance, and nutrition?
- To what degree have health, hygiene and nutrition behaviors and practices promoted by the project been adopted by the targeted beneficiaries?
- What evidence exists to demonstrate the relationship between project activities and outcomes and impact related to literacy, health and nutrition?
- Has participation in WE groups led to increased school involvement and improved student outcomes (particularly for school attendance)?
- Has the project met targets for key performance indicators? If targets are not being achieved, why not?
- What is the level of satisfaction of the direct beneficiaries or participants?

Outcomes and Impacts Related to Literacy:

- What changes in reading comprehension have happened since the baseline evaluation?
- Is there a correlation between literacy outcomes and characteristics of the student, teacher and school? For example, what is the relationship between the amount of literacy instruction in the classroom, quality of literacy instruction, class size, sex of the teacher, hunger level and nutrition of the student, and access to libraries and reading corners?
- Have teachers improved the way they implement literacy instruction in the classroom as a result of participating in teacher training on literacy? For example, are there changes in classroom organization, literacy instruction content, teaching methods, or use of different teaching materials?

Gender

- To what extent are project activities reaching female and male beneficiaries equally?
- What differential outcomes and impacts, if any, are project activities having on women and girls? Men and boys?
- To what extent is the project actively addressing gender-specific barriers and constraints to education and community engagement?
- Has household decision making changed among households with WE group members, particularly related to decisions around their children's education, and importance of girls attending school?

4. Evaluation Methodology

The study will utilize a quasi-experimental design using mixed-methods and a random sample of schools and individuals. The quantitative data collected through representative surveys will provide or confirm mid-term values for all required project indicators (see section *Required Indicators* below) as well as any additional data points deemed relevant by PCI and the consultant. Qualitative methods will involve a broader selection of stakeholders and provide additional information to understand the program's operating context, triangulate quantitative findings, strengthen credibility and validity and provide a greater level of detail and contextualization of findings.

The midterm evaluation includes a control group comprising a sample of schools and respondents not participating in the PCI FFE projects in order to provide a counterfactual for the evaluation. Selection of the control group should ensure comparability between control and treatment and should be consistent with baseline sample selection.

4.1. Sampling Strategy

PCI will prepare the quantitative and qualitative sampling strategy in advance of implementing the evaluation and will provide a detailed sample sizes per strata to the consultant during the start-up phase. The consultant will review, validate, and make any changes to the sampling strategy as needed. The sample frame is all schools participating in the FFE program, and their related students, parents, teachers and school administrators, as well as all schools in the control communities. Sample sizes of students, parents and school administrators for quantitative indicators will allow the detection of statistically significant changes between baseline and mid-term, as well as the detection of statistically significant differences between control and treatment schools at an 80% sample power and 95% confidence level. It is expected that the evaluation consultant will review and understand the sampling strategy.

4.2. Data Collection Methods

Quantitative data collection activities will include literacy tests, student, teacher, administrator and parent surveys, and school and in-class observation checklists. Tanzania uses the Early Grade Reading Assessment (EGRA) to assess literacy among 2nd and 4th graders. In addition to literacy tests, surveys will be conducted with students in both intervention and comparison areas that will focus on key performance indicators related to receipt of school meals, health and hygiene practices, nutrition and food security and attentiveness in the classroom.

Qualitative data will be collected from key respondent groups and stakeholders to triangulate and contextualize the quantitative and secondary data. The beneficiary groups who should be interviewed and/or engaged during the baseline study, and proposed data collection methodologies, are listed in Table 1 below.

All data collection instruments will be based on instruments used during the baseline evaluations to ensure comparability of data over time. The consultant will review and revise baseline tools for use in the MTE as part of the start-up process, with a focus on improving the instruments and using consistent data collection tools in both countries, where possible.

PCI and the data collection consultant will manage all data collection activities, and both will be in communication with the evaluation consultant throughout.

Table 1. Proposed primary data sources and data collection methods

Beneficiary Group	Proposed Data Collection Methodologies	Data Collected
Parents	Quantitative survey with parents	- Midterm results for key performance indicators - Participation and engagement in school activities; knowledge and practices related to health, hygiene, nutrition, and sanitation

		Ability to do small-scale agriculture to contribute to school feeding in general
WE Group Members	Focus group discussions	- Changes in household decision making around education - Involvement of WE group members in schools - Changes in child's performance in school
Teachers	- Representative school-level survey for quantitative indicators - Observational validation of infrastructure and other secondary data - Key informant interviews (KIIs) and/or FGDs	- Quantitative indicator data - School characteristics (physical infrastructure, demographics, school absences, etc.) - Knowledge and practices around literacy instruction, and education in general - Existing knowledge and practices, such as teaching methodologies, health, hygiene, nutrition, and sanitation
School Directors or Head Teachers	- Quantitative survey - KII	- School characteristics (physical infrastructure, demographics, school absences, etc.) - Quantitative indicators (school enrollment, attendance levels, attention levels, teacher capacity and performance, school gardens)
Students grades 4-6	- Quantitative survey for key performance indicators - Key informant interviews (KIIs) - Classroom observation	- Knowledge and practices related to literacy, health, hygiene, nutrition, dietary consumption, and sanitation habits outside of school - Classroom attention and staying "on task" (grade 2)
Students for literacy tests only	EGRA in Tanzania 5.	- Percent 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 - Percent of students who score above national benchmarks in non-word reading - Percent of students who score above national benchmarks in oral fluency reading
Government officials, including MOE representatives	KIIs for qualitative data and to triangulate and contextualize quantitative / secondary data	- Coordination and communication between different stakeholders, e.g. community members, schools, and municipal governments and agencies - Adoption of policies and processes to help ensure the sustainability of school feeding programs

Local government representatives, including school administrators at the village, ward/municipality and district/region levels	KIs for qualitative data and to triangulate and contextualize quantitative / secondary data	- Coordination and communication between different stakeholders, e.g. community members, schools, and municipal governments and agencies - Potential and progress towards sustainability - Social participation/community engagement in school activities
Local community organizations, such as farmer groups	FGDs for qualitative data and to triangulate and contextualize quantitative / secondary data	- Coordination and communication between different stakeholders, e.g. community members, schools, and municipal governments and agencies - Participation/community engagement in school activities

5.1. Required Indicators

The consultant will be required to analyze and report on the following key performance indicators using primary data collected during this mid-term evaluation. Indicator IDs in the below table reference whether the indicator is custom or a standard indicator based on the 2016 FFP/MGD Indicator Guidance. Data for additional indicators collected by PCI through routine monitoring will be provided to the consultant for a complete picture of midterm project performance.

Table 2. Required Indicators

Indicator ID	Indicator	Disaggregation
MGD #26	Percent 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	Sex, grade
Custom	Percent of students who score above national benchmarks in non-word reading	sex
custom	Percent of students who score above national benchmarks in oral fluency reading	sex
custom	Percent of students who are off task at the beginning, middle, and end of a classroom lesson	sex
Custom	Percent of students who indicate they are “full” during the day	sex
Custom	Percent of students who were absent from school due to illness in the past six months	sex
custom	Percent of school-age children receiving a minimum acceptable diet	sex
Custom	Percent of students who can identify at least three good health/hygiene practices	
Custom	Percent of mothers (or care providers) in targeted communities who can identify at least three important health/ hygiene practices	

Custom	Percent of students that can identify at least three important nutritional/ dietary recommendations	sex
Custom	Percent of parents that can identify at least three important preventative health practices	
Custom	Percent of parents in target communities who can name at least three benefits of children completing primary school	
MGD #1	Number and Percentage ²⁴ of students regularly (80%) attending supported schools	sex
MGD #22	Number and Percentage of schools using an improved water source	
MGD #23	Number and Percentage ²⁵ of schools with improved sanitation facilities	
MGD #3	Number of school administrators in targeted schools who demonstrate use of new techniques or tools as a result of USDA assistance	
MGD #4	Number of teachers/educators/teaching assistants in target schools who demonstrate use of new and quality teaching techniques or tools	

5.2. Data Analysis

The data collection consultant (with PCI's support) will conduct quantitative and qualitative data collection. Full, cleaned datasets and translated qualitative data will be given to the evaluation consultant immediately following data collection (May 2019).

The evaluation consultant will conduct quantitative & qualitative data analysis. Mid-term results should be calculated based on each indicator's definition and presented in comparison to baseline results and to the control results for each country. Appropriate tests of statistical significance should be used to compare baseline to mid-term results for the KPIs listed above. The consultant is also expected to analyze and present key results not related to the above KPIs, as relevant to answering the evaluation learning questions.

Qualitative data should be analyzed using appropriate techniques, such as content analysis, and must go beyond simply pulling out key quotes. PCI will provide all quantitative data to the consultant in the form of focus group transcripts and interview notes.

The evaluation consultant will present an analysis plan to PCI during the start-up process that explains the data analysis approach, methodologies and tools. The evaluation consultant will also be engaged prior to data collection to ensure tools and systems are designed in a way that ensures efficient and effective data analysis.

6. Key Tasks of the Consultancy

²⁴ Percentage based on total number of students enrolled in intervention schools in the current school year

²⁵ Percentage based on total number of intervention schools

6.1. Evaluation Consultant Responsibilities

- Review of project documents to understand the project design, local context, progress to date and baseline evaluation methodology and results
- Review and propose revisions of all data collection tools in collaboration with PCI and data collection consultant.
- Support the in-country enumerator training, in partnership with data collection consultant
- Ensure data quality by checking incoming data for completeness and validity and communicating any issues to the data collection team/consultant
- Review, code and clean quantitative datasets before analysis. Data coding and cleaning will be enabled by the use of mobile data collection platforms, but the consultant will be responsible for ensuring that all final data meet PCI data standards.
- Analyze quantitative and qualitative findings for a comprehensive final report
- Produce mid-term evaluation report for PCI's submission to USDA
- Provide a summary presentation in Powerpoint and highlights brief to PCI

6.2. PCI Responsibilities

- Provide all project documentation to consultants
- Prepare all introductory letters for regional and district government officials to introduce the study
- Coordinate on-boarding meetings for the data collection consultant with PCI and with the evaluation consultant
- Review methodology, timeline, and analysis plan from evaluation consultant
- Develop implementation management dashboard and set up automated email reports on survey completion
- Support data collection consultant to coordinate the enumerator training, including dates, location and accommodations
- Provide in-country logistics for consultant teams
- Review and approve evaluation deliverables from both evaluation consultants

5.3 Data Collection Consultant Responsibilities

- Prepare for data collection and enumerator training by reviewing project documentation and participating in several meetings with the PCI team
- Prepare and submit study information to all relevant Government of Tanzania partners, including but not limited to NBS, COSTECH, and PORLAG for approval
- Finalize data collection instruments
- Prepare and submit IRB materials for local Tanzania approval, as determined necessary
- Hire and manage team of enumerators and supervisors
- Co-lead 5-day enumerator training, including 1 day of field testing
- Conduct quality assurance as per the above list scope of work
-
- Work with Evaluation Consultant to build surveys into mobile data collection platform
- Collect an estimated 3000 surveys in 100 schools
- Collect qualitative data from an estimated 70 participants in Bunda, Butiama, and Musoma, and translate all transcripts into English
- Submit clean quantitative data and transcribed qualitative data to the Evaluation consultant

- Prepare a final report detailing data collection processes, lessons learned, and recommendations for the endline evaluation

7. Deliverables

1. Inception report that includes implementation timeline, analysis plan and tests of statistical significance
2. Revisions to the data collection tools, as appropriate/needed, in consultation with PCI
3. Evaluation report outline
4. Presentation of initial results to PCI, prior to writing evaluation report
5. Draft Midterm evaluation report
6. Final versions of midterm evaluation report, approved by PCI. Reports will be in line with PCI guide for evaluation report writing and editing, as well as PCI evaluation report minimum standards
7. Finalized, clean analytical dataset. Quantitative datasets must adhere to PCI data standards, to be provided to consultant
8. Data dictionaries or codebooks for the analytical dataset
9. Syntax and coding files used to assemble or analyze the data
10. A table of contents describing each of the files
11. Summary Power Point presentation and highlights brief of the key findings

8. Estimated Timelines

Data collection will be conducted in May 2018. All school data must be collected prior to exams, which begin 27 May. It is expected data collection itself will take a max of 19 calendar days. Enumerator training should occur immediately prior to data collection, which is scheduled for 29 April – 2 May. Table 3 below outlines the agreed upon timeline, including responsibilities of the data collection consultant, evaluation consultant, and PCI.

Table 3. Timeline for Midterm Evaluation

Activity	Task	Responsible	Start Date	End Date	Comments
Government approvals	Consultant prepares and submits IRB materials	Data collection consultant	10-Apr	30-Apr	
	Consultant submits study information to NBS, COSTECH, PORALG for approval	Data collection consultant	10-Apr	30-Apr	
Inception Report	Submit inception report to PCI	Evaluation consultant	11-Apr	12-Apr	
Data Review	Review relevant program materials	Data collection / evaluation consultant	10-Apr	30-Apr	
Hire enumerators	Hire full team of enumerators	Data collection consultant	22-Apr	26-Apr	Share list & CVs with PCI

Training preparation	Finalize plans for training, including training manuals and presentations	Data collection / evaluation consultant/PCI	24-Apr	25-Apr	Meet at Ipsos offices in Dar
Enumerator training	Enumerator training on all tools (including EGRA)	Data collection consultant	29-Apr	2-May	May 1 is a holiday. Evaluation consultant to support.
	Pre-testing	Data collection consultant	3-May	3-May	Evaluation consultant to support
	Make changes to tools as needed	Data collection consultant	4-May	4-May	In partnership with evaluation consultant
	Enumerators tested on EGRA to select top performers	Data collection consultant	4-May	4-May	
Data Collection	Quantitative and qualitative data collection conducted	Data collection consultant	6-May	24-May	Schools have exams 27 May - 7 June so no data collection can happen during this time. Schools close 7 June.
	Submit final data collection deliverables and report to PCI	Data collection consultant	27-May	14-June	
Data Analysis	Quantitative and Qualitative Data Analysis	Evaluation consultant	15-June	12-Aug	
	Send preliminary findings to PCI	Evaluation consultant		12-Aug	
Write Report	Send first draft of report to PCI	Evaluation consultant		26-Aug	
	PCI reviews and provides feedback to evaluation consultant	PCI	26-Aug	6-Sept	We expect there may be multiple rounds of revision
	Evaluation consultant produces final report	Evaluation consultant		25-Sept	

9. Qualifications

- Consultant or team members must have an academic background in social / education and /or health sciences or related field; minimum of Masters Degree
- Extensive experience in leading project evaluations using a range of quantitative and qualitative data collection and analysis methods for quasi-experimental designs
- Experience with administering literacy assessments such as the EGRA and analyzing the collected data with concise results and presentation.
- Experience in administering gender analysis tools and analyzing the collected data

- Experience in leading evaluations in the areas of international development, specifically education programs and public health; experience in evaluating school feeding programs strongly preferred
- Experience in the evaluation of projects funded by U.S. government/USDA strongly preferred
- Knowledge of Tanzanian context, particularly around national legislation relating to education, health and nutrition, preferred
- Demonstrable capacity to deliver high quality outputs within the proposed timeframe

10. Corrupt Practices

PCI requires that applicants observe the highest standard of ethics during the procurement and execution of such contracts. In pursuance of this policy, PCI agrees with the Tanzanian Government in:

1. Defining for the purposes of this provision, the terms set forth below as follows:
 - A. "*Corrupt practice*" means the offering, giving, receiving or soliciting, directly or indirectly, of anything of value to influence the action of a public official in the procurement process or in contract execution;
 - B. "*Fraudulent practice*" means a misrepresentation or omission of facts in order to influence a procurement process or the execution of a contract;
 - C. "*Collusive practice*" means a scheme or arrangement between two or more applicants, with or without the knowledge of PCI, designed to establish prices at artificial, non-competitive levels; and
 - D. "*Coercive practice*" means harming or threatening to harm, directly or indirectly, persons or their property to influence their participation in a procurement process, or affect the execution of a contract.
2. Rejecting a recommendation for award if it determines that the applicant recommended for award has, directly or through an agent, engaged in corrupt, fraudulent, collusive or coercive practices in competing for the contract in question; and
3. Debarring an applicant from participation in procurement for a specified period of time if it at any time determines that the firm has engaged in corrupt, fraudulent, collusive or coercive practices in competing for, or in executing, a contract.

Annex H. Inception Report

Background

Project Background

PCI is implementing five-year project funded under the USDA's McGovern-Dole International Food for Education and Child Nutrition Program, which is in its third phase. The project is designed to improve the literacy of school-age children through improved literacy instruction, student attentiveness and student attendance. Secondly, the project intends to increase the use of positive health and dietary practices. The third phase is scheduled to run October 2016 to September 2021.

Key Activities and Strategies²⁶

According to PCI's project documentation, the program in Tanzania is focused on achieving desired results:

- Improved literacy of school-age children through improved quality of literacy instruction, improved student attentiveness and improved student attendance
- Improve literacy instruction quality by promoting teacher attendance, providing teacher recognition and school supplies, establishing mini-libraries and classroom reading corners, distributing reading materials and training teachers and school administrators;
- Improve attentiveness and reduce short term hunger by providing school meals and increasing farmer agricultural productivity and production, in turn contributing to locally-owned and sustained school feeding programs;
- Improve student attendance by forming savings and lending groups that provide financial stability to keep children in school, providing training on better health and nutrition practices to decrease health related absences, increasing community understanding on the benefits of education, and in Tanzania, promoting parent engagement of children's education through Education Cascade Groups (ECGs);
- Increase knowledge of safe food preparation and storage practices by training relevant stakeholders, and increase access to food preparation and storage tools and equipment through the distribution of food and preparation equipment;
- Increase access to clean water and sanitation services through rehabilitation or construction of latrines and water systems that are maintained through active school WASH infrastructure committees;
- Increase access to preventative health interventions through capacity strengthening and training;
- Increase community understanding of the benefits of education through Education Cascade Groups; and
- Increase local food contributions to schools through improved agricultural productivity in farmers groups, as well as improved food storage and handling practices.

²⁶ Thorough descriptions of the project, its activities, and its goals are described in detail in appropriate documents.

Evaluation Background

The purpose of the mid-term evaluation is to evaluate the progress of the project so far, whether interventions are being delivered as planned, in accordance with the project budget and work plan. The Mid-term evaluation will seek to respond to the following evaluation questions:

1. Assess the relevance of the project strategy and approach, and the validity of assumptions made during project design;
2. Measure progress from baseline, including the effectiveness, efficiency and timeliness of interventions in achieving established targets, identify enablers and constraints to progress, and provide recommendations for mid-course corrections to strengthen performance;
3. Assess sustainability efforts to date;
4. Document lessons learned, challenges and unanticipated effects;
5. Provide recommendations for areas of focus for the final evaluation, including review and strengthen data collection systems and metrics.

The evaluation will follow a quasi-experimental design methodology, with intervention schools and control schools, that allows for rigorous impact and attribution analysis and accounts for any confounding factors. The data collection instruments will include a combination of qualitative and quantitative tools. This will include a desk review of all the project documentation, an evaluation of the reading level (EGRA) student, parents and teacher questionnaire, focus groups with parents and teachers, KKI with staff, stakeholders and partners as well as field visit and observation.

Evaluation Design

Study Objective

The evaluation will focus on three key topics: relevance, sustainability, and effectiveness. Special attention will be paid to literacy outcomes and gendered experiences during the evaluation. As the program's highest objective is improved literacy among primary school students, the role of activities in driving literacy gains will be a central focus of the analysis.

Many of the project's activities have highly gendered elements. In addition, education itself has different barriers, benefits, and processes to boys and girls. For these reasons, the analysis will investigate throughout how the processes and outcomes vary by gender.

Evaluation Criteria

Relevance

- To what extent is the results framework and program approach valid for the local context?
- To what degree have the needs of the beneficiary population been included in the design of the program?

- Where similar activities are occurring within the same project area, how well does the project coordinate and complement existing programs to ensure efficient and effective use of resources and avoid duplication?
- Which program strategies implemented have been more effective and relevant in contributing to program results? Which were the least effective and least relevant?
- How has implementation and accomplishment of program activities been carried out at the technical /programmatic, organizational and contractual levels?
- To what extent did PCI address the baseline evaluation recommendations and lessons learned from the FFE phase II program in Tanzania?
- Does the project design, structure, logic or management need to be modified to facilitate the achievement of desired outcomes and impact?

Sustainability

- How well has the project developed and maintained collaborative relationships with appropriate government agencies, local partners, USDA, and other key stakeholders?
- What efforts has PCI made to transition school feeding activities to local actors (local government, private sector, federal government agencies, parents)? How successful have those efforts been?
- Does the participation of mothers and/or other family members in the WE groups contribute to the sustainability of the program and program outcomes?
- What progress has the project made to ensure the sustainability of adequate human, financial, infrastructure, capital and other resources?
- To what extent have partners, government representatives, and local communities internalized and implemented efforts to sustain school feeding activities, such as providing in-kind support for school feeding activities?
- What specific local, regional, governmental, and school stakeholder institutional capacities have been developed to foster host-country engagement and accountability. What additional capacities should be developed for which stakeholders?
- What elements or considerations should be incorporated into public policy in the local governments to sustain a school feeding program?

Effectiveness

- Has the program achieved the expected outcomes for literacy, teaching quality, student attention and attendance, and nutrition?
- To what degree have health, hygiene and nutrition behaviors and practices promoted by the project been adopted by the targeted beneficiaries?
- What evidence exists to demonstrate the relationship between project activities and outcomes and impact related to literacy, health, and nutrition?
- Has participation in WE groups led to increased school involvement and improved student outcomes (particularly for school attendance)?
- Has the project met targets for key performance indicators? If targets are not being achieved, why not?
- What is the level of satisfaction of the direct beneficiaries or participants?

Special Topics of Interest

Literacy

What changes in reading comprehension have happened since the baseline evaluation?

Is there a correlation between literacy outcomes and characteristics of the student, teacher and school? For example, what is the relationship between the amount of literacy instruction in the classroom,

quality of literacy instruction, class size, sex of the teacher, hunger level and nutrition of the student, and access to libraries and reading corners?

Have teachers improved the way they implement literacy instruction in the classroom as a result of participating in teacher training on literacy? For example, are there changes in classroom organization, literacy instruction content, teaching methods, or use of different teaching materials?

Gender

To what extent are project activities reaching female and male beneficiaries equally?

What differential outcomes and impacts, if any, are project activities having on women and girls? Men and boys?

To what extent is the project actively addressing gender-specific barriers and constraints to education and community engagement?

Has household decision making changed among households with WE group members, particularly related to decisions around their children's education, and importance of girls attending school?

Evaluation Approach

Our team will utilize a mixed-methods (quantitative and qualitative) approach to the mid-term evaluation and will assess both performance and impact indicators. We will assess the overall program strategic objectives and performance indicators listed in the Performance Monitoring Plan, measuring how much progress has been made against the baseline targets.

Throughout evaluation design and data collection preparation periods, qualitative and quantitative data collection methods, sampling decisions, and development of tools have been conducted simultaneously to ensure the quantitative and qualitative data complement each other, act as forms of verification of each other's accuracy, and that their combined information will meet all aspects necessary to answer all research questions.

Difference in Difference Analysis

The overall approach will use two methods to compare the data. For the key indicators of interest, the evaluation will focus on using a difference-in-difference method to determine the effects of the project. The evaluation will compare improvement in treatment schools since baseline to improvement in comparison schools since the baseline.

Subgroup Analysis

Some detailed analysis, such as variation by school or demographic subgroup, cannot be completed as a difference-in-difference analysis due to limitations of the data provided from baseline. In these cases, the analysis will focus on comparing differences between the treatment and control groups at midline.

Power Analysis

The overall objective of the FFEIII project is to improve literacy of school-age children through an integrated approach to address root causes of poor literacy outcomes for students. This section defines the

1. Variable of interest for determining minimum sample size
2. Statistical parameters used in testing

3. Minimum detectable sample required by tests
4. Minimum sample sizes used during data collection

Variable of Interest

The learning indicators are:

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

1. Percentage of students who score above national benchmarks in oral fluency reading
2. Percentage of students who score above national benchmarks in non-word reading
3. 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

These three indicators correspond to three sections (known as subtasks) of the Early Grade Reading Assessment:

1. Non-word (or devised word) reading: the number of invented words a student can read in 60 seconds
2. Oral reading fluency: the number of correct words in a passage that can be read aloud in 60 seconds
3. A reading comprehension quiz after reading a short story.

For simplicity, minimum sample size calculations focus on oral reading fluency, but the sample sizes for all three remain the same. The primary way it will be measured is as the improvement in average oral reading fluency scores since baseline, comparing treatment and control students, using difference-in-difference modeling.

Statistical Parameters

Based on most common statistical practices, we recommend using a minimum statistical power of 0.8 and a significance level of 0.05. Based on the externally provided statistical recommendations, we include adding a 10 percent buffer to the sample to account for invalid data. We recommend assuming an intra-cluster correlation (correlation of student literacy within schools) at 20 percent. Provided the decision to re-use the same schools in the baseline, we calculated minimum sample sizes based on 50 treatment and 50 control schools. To estimate the standard deviation of EGRA scores, we averaged the standard deviation of Grade II scores and Grade IV scores at baseline.

Table 1. Parameters for Power Analysis

Statistical Parameters	
Parameter	Value
Power (β)	.80
Significance Level (α)	.05
Sample Buffer	.10
Intra-Cluster Correlation	.10
Ratio of Treatment:Control	1:1
Logistical Parameters	
No. Treatment Schools	50
No. Control Schools	50
Phase I & II to Phase III School Ratio	2:1
Estimate of 1 Standard Deviation	21.3 words per minute

Minimum Detectable Size

The most important decision in deciding on a sample size is what the minimum detectable difference between the two groups (treatment and control, or midline and baseline). Evaluations often measure this in terms of a standard deviation (SD) improvement – for example, a minimum detectable effect of 1 standard deviation would mean that tests would find a statistically significant difference only if the average student score at midline was equal to a student in the top 16th percentile of the baseline group. If the average score at midline were only equal to a student in the top 20th percentile at baseline, it would not be detectable in a statistically significant way.

Based on discussions with PCI, it was agreed to design the sample sizes using a minimum detectable difference of 0.28 standard deviations, which should allow the detection of an improvement of reading 7 correct words per minute or more.

Calculating minimum samples

Determining the minimum sample size requires calculating (i) the minimum sample size for a simple random sample; (ii) accounting for attrition or invalid data; and (iii) accounting for cluster sampling; and (iv) multiplying the minimum sample size by the smallest subgroup of interest for testing.

Simple Random Sample. By conducting a power analysis in Stata, it was determined that if the study uses a simple random sample, a minimum of 330 observations would be required to detect a 0.33 SD difference. Assuming 10 percent of data collected would be invalid or attrite, the minimum sample sizes increase to 367 observations.

Cluster Analysis. For a minimum detectable improvement of 7 words per minute using cluster analysis, the minimum sample must be increased to 572 observations. A simple random sample would mean randomly selecting from all students in all project and control schools. Such random selection would require visiting many more schools. Instead, cluster sampling involves first selecting a group of schools, and randomly sampling a number of students from each school. Because students at the same school

are likely to have similar learning results, the sample must increase to must account for intra-cluster correlation.

The following table shows the minimum sample size given sampling in 100 schools by different minimum detectable differences in the average number of words a student correctly reads in a minute. Increasing the minimum detectable effect size is the synonymous with reducing the specificity of the test: a minimum effect size of 9 words per minute means that the evaluation cannot reliably detect an improvement of only 7 words per minute. As can be seen, the minimum sample sizes increase exponentially as the tests become more precise.

Table 2. Minimum Sample Sizes by Reading Improvement

Improved Words per Minute	Sample Minimum	Design Effect	Cluster Size	Test + Attrition	Simple Random Sample Min.
5	1,558	2.2	14.3	713.3	642
6	892	1.8	9.9	496.7	447
7	572	1.6	7.3	366.7	330
8	394	1.4	5.6	280.0	252
9	286	1.3	4.4	220.0	198
10	221	1.2	3.6	180.0	162
11	176	1.2	3.0	150.0	135

Disaggregation by Grade and Phase

The above minimum sample sizes are based strictly on tests that make use of the entire sample. For any subgroup analysis to meet the statistical power described above, that subgroup must meet the minimum size. Requiring testing differences at the Grade II and Grade IV level independently means doubling the sample size. Requiring testing differences by grade and phase III vs. Phase I & II, schools doubles the total sample size again. This increases the minimum sample size for the EGRA from 572 to 2,288.

Grade. To measure using the above minimum detectable sizes for each grade separately, the minimum sample sizes above would need to be used for each grade. For this reason, the minimum sample size must not be a minimum of 572 students total but 572 Grade II students plus 572 Grade IV students.

Grade and Phase. Testing whether students in Phase III schools have significantly different scores than those from Phase I and II requires the minimum number of observations in treatment schools to meet the minimum observation requirement of 572 treatment students in each grade.

This exercise determines a minimum sample size for EGRA data collection at 2,288 observations. It was agreed with PCI that other indicators required less precision to be able to detect change, as learning is a more complex thing to assess. It was agreed to use a minimum of 200 students in Grades VI and VII for their surveys, and a minimum of 640 parents (one of each gender in each grade in the study).

Sampling Plan

School Selection

To ensure comparability, the same schools are being used in the midline evaluation as were used in the baseline evaluation. This has the key benefit of any stochastic errors or time-invariant sampling bias during the baseline will be unchanged in the midline. Given that variables of interest will be represented in terms of improvement since the baseline, this will cancel out any static effects of the sample selection. We recommend continuing this plan and using the same schools in the endline evaluation as well.

Student Selection

At each school, 12 Grade II, 12 Grade IV, 2 Grade VI, and 2 Grade VII students are randomly selected from the enrollment lists of that grade. Half of the students from each grade will be male, and half will be female. They are randomly selected by an advance team, which provides consent forms and assigns unique student identification numbers.

Focus Groups

A total of 37 Focus Group Discussions (FGD) will be convened across program and comparison areas in May 2019. The table below provides details on the composition of these groups. Most groups are expected to have between seven and 10 participants. The teacher group discussions are expected to have up to five participants. Discussions should run between 60 and 90 minutes and will be audio recorded. Based on the input and priorities of PCI technical leads, the decision was made to focus FGD data collection efforts on program participants and their communities in order to maximize evaluation insights for intervention sustainability.

Table 3. Focus Group Discussions

Group Type	FFE Program Areas	Comparison Areas	Selection Criteria
Parents	6	4	3 high performing, 3 low; 2 high and 2 low in Serengeti
WE Groups	5		3 high performing, 2 low
Farmer Groups	6		3 high performing, 3 low
Grade 7 Boys	3	1	
Grade 7 Girls	3	1	

Header Teachers with Store, School Garden, Library, and Health Teachers	5	3	
Subtotal:	28	9	
Total FGDs:	37		

Key Informant Interviews

Key Informant Interviews (KII) will be conducted with 27 strategic FFE project stakeholders. Please see Table 4 for a breakdown of KII respondents by category. Interviews are expected to run from 45 to 70 minutes and will be audio recorded. The KII data collection efforts, in alignment with the FGD design, focus predominantly on the experiences of program participants and their communities. This approach aims to maximize the collection of qualitative data that illuminates and informs program sustainability efforts in the last two years of FFE project implementation.

Table 4. Key Informant Interview Subjects

	FFE Program Areas	Comparison Areas	Selection Criteria
Literacy Teacher	6		4 high performing teachers, 2 low
Village Chairperson	3		Convenience
Agricultural Extension Officer (Village)	2		1 high performing officer, 1 low
Agricultural Extension Officer (Ward)	2		1 high performing officer, 1 low
Ward Education Officer	3		Convenience
FFE Focal Person - District Planning Officers	4		Full set
District Education Officer	4	1	Full set
Subtotal:	24	1	
Regional Agricultural Officer	1		Single position
Regional Education Officer	1		Single position
Total KIIs:	27		

Data Collection Support

Research Tools

All midline quantitative and qualitative data collection tools were drafted and finalized between April 19th and May 2nd. These tools were informed by baseline tools, a revised evaluation scope that elevated different learning priorities related to program quality and sustainability, and input from project technical and monitoring and evaluation team leads.

Four new quantitative tools (for parents, head teachers, grade II/IV students, and grade VI/VII students), three revised quantitative tools (classroom observation, attendance record, and school sustainability assessment), and 10 new qualitative tools were finalized during the week of quantitative and qualitative enumerator training.

Quality Assurance

During Training

One of the consultants, Dr. Hamel, attended segments of both quantitative and qualitative enumerator trainings between Monday, April 19th and Saturday, May 4th. She reviewed the quantitative training schedule for the week and provided modest input. She debriefed with the IPSOS quantitative training leads on a daily basis to keep apprised of progress and challenges.

The qualitative training schedule was less clearly presented and discussed. Dr. Hamel attended the full qualitative training from Thursday, May 2nd through Saturday, May 4th. This included one pilot Focus Group Discussion session during the morning of Friday, May 3rd. Dr. Hamel raised concerns about training adequacy which were addressed by IPSOS after the first complete day of data collection.

Data Collection

Dr. Hamel was closely involved with all aspects of the data collection planning process during the week of enumerator training. This included engagement on topics such as parental consent procedures, selection bias by teachers, absenteeism, substitute and resampling, payment for participation, advance team directives, enumerator quality control, and various logistical contingencies. She joined both quantitative and qualitative teams for pilot testing on Friday May 3rd and participated in the qualitative research debrief of that pilot. On Monday, May 6th, she did not observe data collection but participated in both quantitative and qualitative day-end debrief meetings which facilitated the swift correction of discrepancies in sampling approach. She again joined both teams on Tuesday, May 7th, visiting multiple schools, conversing with head teachers about program experiences and evaluation expectations, and observing KIIs with the Regional Education Officer and the Regional Agriculture Officer in Musoma. At the daily debrief meeting, the decision was made to reduce the length of four qualitative tools and to prioritize additional training time for the qualitative team. Ongoing correspondence with the IPSOS team indicates that data collection is proceeding smoothly.

In addition, Mr. Trembley carefully monitored data collection progress remotely. Within the first four days of data collection, the IPSOS team provided all of the data uploaded by enumerators at that point. Mr. Trembley reviewed the data and began to conduct analysis immediately to search for any unforeseen problems. Using quantitative techniques, Mr. Trembley identified a limited number of questions with response patterns that were not expected on the literacy examination and informed IPSOS and PCI of the issues. He also helped identify problems with student identification number assignment that IPSOS promptly developed remedies for, including additional feedback to advance teams and enumerators.

Analytical Plan

Mixed Methods Analysis

During the analytical phase, once the quantitative and qualitative data are received, they will each be prepared for thorough review through coding and verification of accuracy and quality. Initial quantitative findings will be used to identify areas of interest for the qualitative analysis, and the qualitative analysis will highlight interesting questions for a greater quantitative analysis. Through rounds of reviewing the information simultaneously, followed by iterations of periods of deeper analysis into each, we will thoroughly triangulate answers to the research questions

Quantitative Methods

Sample Analysis

The first step in the analysis is to determine whether the samples are representative and comparable. Tests for equality will be used to compare demographic differences between:

1. Baseline and midline groups
2. Midline Treatment and midline control groups
3. Midline treatment and the entire beneficiary group

If it is determined that any of these groups are not comparable, statistical compensations and due consideration will be used through the rest of the analysis.

EGRA Analysis Plan

Preparation of the learning assessment data requires first ensuring the quality of the data by eliminating invalid observations and compensating for limitations of the data collection. This includes identifying outlier response patterns (such as questions that for some reason students were unlikely to get correct), accounting for enumerators who tested harder or easier than the average, or other artifacts of the data that would require compensation. Statistical weights and compensations may be used to improve the accuracy of the analysis. Tests for consistent and coherent data will be conducted.

Next, scores will be calculated for each activity, subtask and learning indicator of interest. Finally, the analysis will report the findings by each subgroup of interest. These include the learning indicators described in the Statement of Work and Performance Monitoring Plan.

Finally, learning indicators will be reported by subgroups of interest. These include by grade, phase of implementation at school, gender, and treatment and control group. Tests will be conducted to test for

significant differences in learning indicators by treatment group by each grade, within the grade and phase, and by gender both between and within the treatment and control groups.

Key Indicator Analysis

Non-learning indicators, including those on knowledge, opinions, and practices, will also be cleaned to ensure responses are coherent and logical. Indicators will be reported on by treatment and control, differences between Phase III treatment schools and treatment schools from earlier phases, and by gender.

In addition to the key reporting completed above, a more thorough analysis will be informed by findings from the focus groups, key information interviews, and listed research questions. These will involve theories of process, variation in implementation and issues within schools and regions, and challenges faced by different types of beneficiaries.

Qualitative Methods

Qualitative data, in the form of Focus Group Discussions (37) and Key Information Interviews (27), will be transcribed in their original Swahili and translated into English by IPSOS. The English language transcripts will then be coded and analyzed thematically using either the Dedoose or QDA Miner software package. Analysis will be structurally informed by the program's Results Framework (theory of change) alongside stated midline evaluation objectives, with a strong emphasis on school feeding program sustainability after 2021. The focus will draw from grounded theory, an inductive and observation-led approach to inferring logical frameworks and causal pathways as articulated by research subjects. Tools were redesigned to allow for the robust exploration of subject experiences as differentiated by gender.

Team Structure and Composition

The team includes two experts: Reid Hamel, a demographer with extensive experience working in food security, and Andrew Trembley, a quantitative analyst who specializes on evaluations in the education sector. Both team members have balanced careers between conducting research in their sectors and in the design and in the monitoring & evaluation of international development projects.

Dr. Reid Hamel

Dr. Hamel is an experienced research of food security investments worldwide, with an emphasis on agricultural intensification and sustainability; good governance; risk management, financial services, and social protection; smallholder livelihoods; and nutrition. She has published numerous studies on smallholder agriculture, good governance, social protection, and nutrition. She recently designed and taught a new course on global food security and population growth at Middlebury College.

Dr. Hamel completed her Ph.D. at the University of California, Berkeley, where her dissertation quantified the public-sector cost savings to social safety net programs associated with the incorporation of voluntary family planning services. She was previously associate director of the Department of

Hunger and Livelihoods at Save the Children, where she led research design and evaluation strategy across a broad Food for Peace portfolio. She was the lead author on a Save the Children village savings and lending group impact evaluation in 2014. She has also worked with the Economic Development Program at the Asia Foundation, with the Caucasus Research Resource Centers, and as a Peace Corps volunteer in Azerbaijan.

Andrew Trembley

Mr. Andrew Trembley is an education economist with extensive experience conducting statistical analyses for education evaluations. Six of his eight years of experience were spent evaluating education projects for the World Bank and increasing data capacities for government education statistics offices in Latin America & Anglophone Africa.

Between 2014 and 2016, Andrew lead the quantitative analysis for five interventions supported by an \$80 million World Bank loan to the Government of Guatemala. He designed a panel data analysis of two girls' education programs and worked with the government's education statistics office to link their school census and human resource databases for the first time. By using advanced techniques to identify key subpopulations, Andrew used sex and ethnicity disaggregated data to identify three key regions that were the source of the nation's failure to improve girls' -- information previously unknown to government. In addition, barring reliable government population estimates, Andrew adapted the government's official statistical methods to birth and death records and longitudinal surveys to create the most statistically rigorous population modeling available for the country, which were eventually adopted for ministry of education reporting. In addition, Andrew conducted the cost-benefit analysis to determine the value-for-money of each intervention under the project.

Between 2013 and 2016, Andrew supported the design and implementation of results monitoring on several education projects. He worked closely with ministry of education officials and statisticians to design statistically rigorous data systems and their eventual use for impact evaluations. He has collaborated with governments, donor partners, and contracting implementers in several countries in Latin America and sub-Saharan Africa to implement modern data collection techniques and technology and develop distributed reporting systems.

For five years (2011-2016), Andrew was part of a team of researchers who developed a quantitatively rigorous methodology to assess teacher preparation systems worldwide. He was instrumental to the project's conception, methodological design, and ongoing analysis. He designed the survey materials and training and specified the standardized quantitative analysis methodology, and authored over 30 national and regional reports analyzing teacher preparation and support methods abroad

Andrew Trembley leads the evaluation end ensure the technical requirements of the evaluation and its analysis are met. With extensive experience in evaluation design and execution, he will ensure the research methods meets the highest standards of research quality. He will revise the data collection tools, clean, collate, and conduct the statistical analysis of the data.

Reid Hamel supports the preparation, training, and initial supervision of data collection. She will also lead the writing of the report. She will identify key themes, highlight insights and actionable items to relate to PCI.

Deliverables

The following are the deliverables as agreed upon between PCI and the evaluators.

1. Revisions to the data collection tools, as appropriate/needed, in consultation with PCI²⁷
2. Inception report that includes implementation timeline, analysis plan and tests of statistical significance²⁸
3. Evaluation report outline
4. Presentation of initial results to PCI, prior to writing evaluation report
5. Draft Midterm evaluation report
6. Final versions of midterm evaluation report, approved by PCI. Reports will be in line with PCI guide for evaluation report writing and editing, as well as PCI evaluation report minimum standards
7. Finalized, clean analytical dataset. Quantitative datasets must adhere to PCI data standards, to be provided to consultant
8. Data dictionaries or codebooks for the analytical dataset
9. Syntax and coding files used to assemble or analyze the data
10. A table of contents describing each of the files
11. Summary Power Point presentation and highlights brief of the key findings

Timeline

Data collection will be conducted in May 2019. All school data must be collected prior to exams, which begin 27 May. It is expected data collection itself will take a maximum of 19 calendar days. Enumerator training should occur immediately prior to data collection, which was completed during 29 April – 3 May. Table 5 below outlines the originally agreed upon timeline, including responsibilities of the data collection consultant, evaluation consultant, and PCI.

Table 5. Activity Timeline

Activity	Task	Responsible	Start Date	End Date	Comments
Government approvals	Consultant prepares and submits IRB materials	Data collection consultant	10-Apr	30-Apr	
	Consultant submits study information to	Data collection consultant	10-Apr	30-Apr	

²⁷ Already completed.

²⁸ This document.

	NBS, COSTECH, PORALG for approval				
Inception Report	Submit inception report to PCI	Evaluation consultant	11-Apr	12-Apr	
Data Review	Review relevant program materials	Data collection / evaluation consultant	10-Apr	30-Apr	
Hire enumerators	Hire full team of enumerators	Data collection consultant	22-Apr	26-Apr	Share list & CVs with PCI
Training preparation	Finalize plans for training, including training manuals and presentations	Data collection / evaluation consultant/PCI	24-Apr	25-Apr	Meet at Ipsos offices in Dar
Enumerator training	Enumerator training on all tools (including EGRA)	Data collection consultant	29-Apr	2-May	May 1 is a holiday. Evaluation consultant to support.
	Pre-testing	Data collection consultant	3-May	3-May	Evaluation consultant to support
	Make changes to tools as needed	Data collection consultant	4-May	4-May	In partnership with evaluation consultant
	Enumerators tested on EGRA to select top performers	Data collection consultant	4-May	4-May	
Data Collection	Quantitative and qualitative data collection conducted	Data collection consultant	6-May	24-May	Schools have exams 27 May - 7 June so no data collection can happen during this time. Schools close 7 June.
	Submit final data collection deliverables and report to PCI	Data collection consultant	27-May	14-June	
Data Analysis	Quantitative and Qualitative Data Analysis	Evaluation consultant	15-June	12-Aug	
	Send preliminary findings to PCI	Evaluation consultant		12-Aug	

Write Report	Send first draft of report to PCI	Evaluation consultant		26-Aug	
	PCI reviews and provides feedback to evaluation consultant	PCI	26-Aug	6-Sept	We expect there may be multiple rounds of revision
	Evaluation consultant produces final report	Evaluation consultant		25-Sept	