USAID Tusome Pamoja

Midline Findings Report

FINAL

October 2018

This document was produced for review by the United States Agency for International Development (USAID). It was made possible by the support of the American people through the USAID. It was prepared by RTI International for the USAID|Tanzania Tusome Pamoja Program.

USAID Tusome Pamoja

Midline Findings Report

Contract Period: January 12, 2016 – January 11, 2021
Contract Number: AID-621-C-16-00003

Prepared for
USAID|Tanzania
United States Agency for International Development
Office of Acquisition and Assistance

ATTN: Laura Kikuli, Contracting Officer's Representative
lkikuli@usaid.gov

686 Old Bagamoyo Road, Msasani
P.O. Box 9130
Dar es Salaam, Tanzania

Prepared by
RTI International
3040 Cornwallis Road
Post Office Box 12194
Research Triangle Park, NC 27709-2194

Table of Contents

Page

Executive Summary 1
Key Findings 1

1 Introduction 2
1.1 USAID Tusome Pamoja Program Description and Key Indicators 2
1.2 USAID Tusome Pamoja Key Interventions 2

2 Midline Assessment Approach 3
2.1 Overview of the Assessment 3
2.2 Overview of Instruments 3
2.2.1 Midline Reading Assessment .... **Error! Bookmark not defined.**
2.2.2 Midline Mathematics Assessment ..... **Error! Bookmark not defined.**
2.3 Instrument Adaptation Process 5
2.4 Assessor Training 5
2.5 Data Collection 5

3 Sample 6
3.1 Population and List Frame 6
3.2 Sample and Methodology 6
3.3 Final Sample Counts and Estimates 7

4 Results and Findings 7
4.1 Reading Performance 7
4.1.1 Reading performance indicators 7
4.1.2 Reading performance categories 8
4.1.3 MRA results in detail 9
4.1.4 Discussion of MRA results 10
4.2 Mathematics Performance 10
4.2.1 Mathematics performance indicators 10
4.2.2 Detailed MMA results 11
4.2.3 Discussion of the MMA results 12

5 Conclusion 12

Annex A. Final Instruments 14

Annex B. Sample Methodology, Weighting, and Precision 24

Annex C. Inter-rater Reliability and Assessor Accuracy 28

List of Figures

Figure 1. Comparison of MMA Zero Scores by Subtask 12

List of Tables

Table 1. USAID Tusome Pamoja Baseline Indicator Values 1

Table 2. MRA Instrument Subtasks in Tanzania 4

Table 3. MMA Instrument Subtasks in Tanzania 4

Table 4. Sample Methodology Summary 6

Table 5. Sample Counts and Estimated Population Counts by Region 7

Table 6. Percentage of Children Meeting Reading Benchmarks 8

Table 7. Percentage of Children Meeting Reading Benchmarks by Gender 8

Table 8. Reading Categories by Grade 9

Table 9. Comparison of MRA Performance by Subtask 9

Table 10. Overview of Midline MRA Performance by Subtask and Gender 10

Table 11. Percentage of Children Meeting Mathematics Benchmark 11

Table 12. Percentage of Children Meeting Mathematics Benchmark by Gender 11

Table 13. Student Performance on MMA Subtasks 11

Table B.1. Schools Excluded from the List Frame Prior to Sampling 24

Table B.2 Population Counts of Schools and Pupils by Tusome Pamoja Region 24

Table B.3. Sample Methodology for the Tusome Pamoja Midline 2018 Assessment: Three-Stage Stratified Sample of Schools, Classrooms, and Pupils 25

Table B.4. Final Sample Counts of Assessments, Interviews, and Observations Collected for the Tusome Pamoja 2018 Midline Evaluation 26

Table B.5. Weighted Sample (Estimated Population) and Sample Counts by Each Stage Sampled (School, Classroom, Pupil) 26

Table B.6. Reading Fluency Means and Precision Estimate by Standard 27

List of Abbreviations

3Rs	reading, writing, and arithmetic
AAM	assessor accuracy measure
cwpm	correct words per minute
MMA	Midline Mathematics Assessment
MRA	Midline Reading Assessment
IRR	inter-rater reliability
MOEST	Ministry of Education, Science, and Technology
MOEVT	Ministry of Education and Vocational Training
ORF	oral reading fluency
USAID	United States Agency for International Development

Executive Summary

Key Findings

The purpose of this midline evaluation is to measure the impact of the US Agency for International Development (USAID) Tusome Pamoja Program by comparing midline results against the baseline values reported in 2016. According to the Program's Performance Monitoring and Evaluation Plan, there are three key reading and mathematics performance indicators to be reported for Standard 2 at midline:

- P-001: Percent of learners who demonstrate reading fluency and comprehension of grade-level text at the end of Standard 2 with US Government (USG) assistance (oral reading fluency [ORF] 45 correct words per minute [cwpm])
- P-002: Number and proportion (%) of learners who, after two years of schooling, can solve grade-level arithmetic problems. As measured based on the following subtask scores: Addition and Subtraction (Level 2): 80%; and Missing Number: 60%.
- P-008a: Number and proportion (%) of learners who, after two years of schooling, obtain zero scores on oral reading fluency (ORF: 0 cwpm)

Table 1 lists the indicator values for P-001, P-002, and P-008a. The first column shows the definition of the indicator, while the second and third columns show the percentage of children meeting the benchmarks at baseline and midline for each indicator.

Table 1. USAID Tusome Pamoja Baseline Indicator Values

	Percent of Children Meeting Benchmark	
	Baseline	Midline
P-001: Percent of learners who demonstrate reading fluency and comprehension of grade level text at the end of Standard 2 with USG assistance (ORF: 45 cwpm)	1.5% (± 0.5%)	3.5% (± 2.5%)
P-002: Number and proportion (%) of learners who, after two years of schooling, can solve grade-level arithmetic problems (Addition/Subtraction Level 2: 80%, and Missing Number: 60%)	1.4% (± 0.6%)	3.9% (± 2.4%)
P-008a: Number and proportion (%) of learners who, after two years of schooling, obtain zero scores on oral reading fluency (ORF: 0 cwpm)	31.0% (±2.9)	17.2% (±4.1)

Improvements were made for all three indicators. While the gains for P-001 and P-002 were modest, the results for P-008a show a largely significant and meaningful reduction in the percentage of students who were unable to read a single word at the end of Standard 2. This means that students are beginning to move out of the lowest categories of performance, but that it has been difficult to move a large proportion of the students into the highest performance category (which is to be expected based on the skewed distribution and the large proportion of students at the lower end).

1 Introduction

1.1 USAID Tusome Pamoja Program Description and Key Indicators

United States Agency for International Development (USAID)Tusome Pamoja is a five-year (2016–2021) USAID program designed to provide technical assistance and resource to the Government of the United Republic of Tanzania to strengthen the 3Rs (reading, writing, and arithmetic) performance of Tanzanian learners in Standards 1–4 and to pilot pre-primary interventions. The Program is being implemented in four regions in Mainland Tanzania—Iringa, Morogoro, Mtwara, and Ruvuma—and in Zanzibar.

USAID Tusome Pamoja's primary goal is to improve lifelong learning skills, defined as mastery of early grade reading, writing, and arithmetic as measured by the percentage of children who: (i) after two years of schooling, can read and comprehend grade-level text, encode simple sentences, and solve grade-level arithmetic problems (measured in all regions); and (ii) after four years of schooling, can read and comprehend grade-level text, respond to simple writing prompts, and solve grade-level arithmetic problems; and by (iii) the proportion (%) of children obtaining zero scores on oral reading fluency (ORF) at the end of Standard 2 primary or the equivalent. An additional objective of the Program is to develop, implement, and demonstrate the best approaches to strengthen the quality of education in the target regions for replication consideration in other, non-target regions. Within the target regions, there are estimated to be in excess of 1.4 million children enrolled, taught by more than 36,000 teachers in approximately 3,035 schools.

1.2 USAID Tusome Pamoja Key Interventions

The three core results, (i) improving reading instruction in the primary grades; (ii) improving the policy environment for reading; and, finally, (iii) increasing parental and community support for student reading, are broken down to intermediate-result activities or interventions. These multiple, intermediate results are designed to contribute to the successful achievement of each result area.

First, the Program will improve reading instruction in primary grades by providing quality instructional materials for 3Rs—reading, writing, and arithmetic—teaching, including student decodable texts, levelled classroom readers, read alouds, and teacher resource materials, as deemed necessary. Target teachers will be capacitated on effective use of classroom 3Rs teaching and learning materials through school-based professional development model, a community of learning approach, and continuous teacher professional development. Besides that, the school leadership (head teachers and ward education coordinators) will receive training and resource materials to support school-based teacher professional development, community engagement, school governance, planning, etc. Head teachers, ward education coordinators, quality assurance officers, district education officers, and district administrative officers will be trained in supervision and mentoring to provide overall support to teachers.

Second, the Program intends to improve the policy environment for reading by supporting the continuous collection, storage, and utilization of data at all levels. This entails the introduction of a school information system (paper-based and digital) and

school scorecards, strengthening the education management information system in Zanzibar, and designing and carrying out various studies. In addition to these things, USAID Tusome Pamoja will establish and support school-based mentoring, communities of learning for teachers and mentors, and regional systems for early grade reading assessments (using lot quality assurance sampling techniques) that will happen periodically.

Third, USAID Tusome Pamoja will increase parental and community support for pupils in lower grades to master the 3Rs. This will be achieved by establishing appropriate mechanisms to allow parents and community members to be involved in education matters. Such mechanisms include the development of parent–teacher partnerships, provision of school notice boards, capacity building of school management committees, and implementation of a social and behavior change communication campaign.

2 Midline Assessment Approach

2.1 Overview of the Assessment

To establish baseline values against which to measure the impact of USAID Tusome Pamoja, a baseline assessment was conducted between August 15 and September 2, 2016. The study targeted all Standard 2 children attending non-special needs education public schools within the five USAID Tusome Pamoja intervention regions and all Standard 4 children attending non-special needs education public schools in Zanzibar and Mtwara.

In order to measure the impact of USAID Tusome Pamoja halfway through the Program, a midline assessment was conducted in August and September 2018. The midline assessment was designed to focus explicitly on changes in reading and mathematics performance for Standard 2 students across all intervention regions (though Standard 4 students were also tested for internal monitoring purposes). A shortened Midline Reading Assessment (MRA) was administered to measure oral reading fluency and reading comprehension, while an Midline Mathematics Assessment (MMA) was used to measure a range of mathematics skills, as discussed below.

2.2 Overview of Instruments

2.2.1 Midline Reading Assessment

MRA measures the basic skills that a child must possess to eventually be able to read fluently and with comprehension—the ultimate goal of reading. The higher order skills (e.g., fluency and comprehension) build on lower order skills (e.g., letter sound knowledge, decoding), and the lower order skills have been shown to be predictive of later reading achievement. For the purposes of the midline, only the higher order skills were tested.

Table 2 summarizes the subtasks of the MRA used in this study; the actual instruments are provided in ***Annex A***.

Table 2. MRA Instrument Subtasks in Tanzania

Subtask	Skill	Description The child is asked to …
Oral Reading Fluency	Automatic and word recognition	… read aloud a grade-level short story printed on a page. (*Timed subtask*)
Reading Comprehension	Comprehension	… verbally respond to five oral questions (four literal and one inferential) that the assessor asks about the short story. (*Untimed subtask*)

2.2.2 Midline Mathematics Assessment

MMA gathers information about foundational mathematics competencies—those competencies that should typically be mastered during the very early grades, and without which children will struggle to succeed in mathematics or will potentially drop out. Subtasks selected for MMA are drawn from extensive research on early mathematics learning and assessment and were constructed by a panel of experts on mathematics education and cognition. The MMA protocol systematically assesses early numeracy skills, particularly those underlying number sense. These abilities and skills are key in the progression toward the ability to solve more advanced problems and the acquisition of more advanced mathematics skills.

Table 3 summarizes the subtasks of the MMA used in this study; the actual instruments are provided in ***Annex A***.

Table 3. MMA Instrument Subtasks in Tanzania

Subtask	Skill	Description The child is asked to …
Addition and Subtraction (Level 1 [basic facts])	This subtask requires knowledge of and confidence with basic addition and subtraction facts. It is expected that students should develop some level of automaticity and fluency with these facts because they need them throughout mathematics.	… mentally solve addition and subtraction problems, with sums and differences below 20. The problems ranged from those with only single digits to problems that involved the bridging of the 10. There were 10 items per addition and subtraction subtask. (*Timed subtask*)
Missing Number (number patterns)	This subtask requires the ability to discern and complete number patterns.	… determine the missing number in a pattern of four numbers, one of which is missing. Patterns used included counting forward and backward by ones, fives, tens, and twos. There were 10 items. (*Untimed subtask*)

Subtask	Skill	Description The child is asked to …
Addition and Subtraction (Level 2)[a]	This subtask requires the ability to use and apply the procedural addition and subtraction knowledge assessed in the Level 1 subtask to solve more complicated addition and subtraction problems.	… solve addition and subtraction problems that involve the knowledge and application of the basic addition and subtraction facts assessed in the Level 1 subtask. Students were allowed to use any strategy that they wanted, including the use of paper and pencil supplied by the administrator. The problems extended to the addition and subtraction of two-digit numbers involving bridging. There were five items per addition and subtraction subtask. (*Untimed subtask*)

[a] The Addition and Subtraction (Level 2) subtasks are more conceptual than the Addition and Subtraction (Level 1) subtasks because children must understand what they are doing when applying the Level 1 skills. Although the (Level 2) subtasks are not purely conceptual, because, with time, children will develop some automaticity with the items in these subtasks, they are more conceptual than the Level 1 subtasks, especially so for Standard 2 children.

2.3 Instrument Adaptation Process

Language and mathematics experts from the Tanzania Institute of Education and ministry officials from the Ministry of Education and Vocational Training; Ministry of Education, Science, and Technology; and the Ministry of Education, Science, and Technology, along with RTI technical staff and project specialists contributed to the development of the USAID Tusome Pamoja MRA and MMA instruments during a two-day workshop on June 16 and 17, 2016, in Dar es Salaam.

The instruments developed during the June 2016 meeting were largely adaptations of the instruments already developed for previous 2013 and 2016 national early grade reading and mathematics studies.

The instruments used for this midline evaluation were shortened versions (i.e., reduced number of subtasks) of the instruments used at baseline.

2.4 Assessor Training

Assessor training was held for 46 assessors during a five-day workshop held August 13th to August 17th, 2018, in Dar es Salaam. All assessors were trained to administer both the MRA and MMA instruments, as well as the brief accompanying questionnaires. Assessors participated in two assessor accuracy measures (AAM) that evaluated their ability to accurately score subtask items. Problematic items identified through the AAM exercise were addressed and improved upon during the training. Assessors who performed below the conventional threshold of 90% agreement received targeted assistance in order to improve their performance. The results of the AAM are presented in ***Annex C***.

2.5 Data Collection

Data collection took place August 23rd to August 31st, 2018, following the assessor training activities. A total of 41 assessors were deployed to the five regions in which USAID Tusome Pamoja operates. Assessors were divided into teams comprised of 2 to 3 assessors. General oversight and management of the data collection teams was provided by Tusome Pamoja monitoring and evaluation staff. Teams uploaded data

regularly. Quality checks and feedback on uploaded data were provided by RTI home office statisticians.

To supplement these data quality checks, the assessors in the field also performed daily inter-rater reliability (IRR) checks. For these checks, the assessors sampled one additional student. The student was given the MRA as normal with one assessor, while another assessor silently marked the students' responses as well on his or her own tablet. Once the student was dismissed, the assessors compared their marked responses and discussed any discrepancies, without being able to change their original responses. On average, assessors agreed 97% of the time. As such, there were no concerns about the quality of the data being collected. The results of the field IRR checks are presented in ***Annex C***.

3 Sample

3.1 Population and List Frame

The population of interest was all Standard 2 children attending non-special needs education public schools within the five USAID Tusome Pamoja intervention regions. The list of schools in the 2016 National Examinations Council of Tanzania Primary School Leaving Certificate Examination was used as the sampling frame for the schools in Mainland Tanzania, and the 2016 census list of schools collected by the Ministry of Education and Vocational Training was used for Zanzibar. See ***Annex B*** for details.

3.2 Sample and Methodology

The sample methodology calls for two three-stage samples of schools, classrooms, and children. ***Table 4*** below provides a summary of the sample methodology used.

Table 4. Sample Methodology Summary

Stage Number	Item Sampled	Stratified by	Probability of Selection
Stage 1	Schools (90)	**Region (5)**	Proportion proportional to enrollment*
Stage 2	Standard 2 Classrooms (90) Standard 4 Classrooms (90)	**Standard (Standards 2 and 4)**	Equal
Stage 3	Standard 2 Children (901) Standard 4 Children (894)	**Gender (Boys and Girls)**	Equal

*Proportional to Enrollment: For the Mainland data, enrollment was the total number of children who sat for the Primary School Leaving Certificate Examination in 2016. For the Zanzibar data, it was the Standard 2 enrollment for the 2016 census data.

The sample was derived to provide regionally representative estimates of children's performance. The sample was designed to be able to report the estimated mean ORF within each grade with a 95% confidence interval band of about ±3.0 words per minute.

In the study it was found that all confidence intervals for ORF across the grades were tighter than ±3.0, with bands ranging from ±1.9 words per minute to ±2.4.

3.3 Final Sample Counts and Estimates

A total of 1,795 Standard 2 and Standard 4 children in all five regions were assessed in 90 schools overall. From each of these schools, a Standard 2 and Standard 4 classroom was selected; the teacher of the selected classroom was interviewed; and the classroom inventory was conducted. The sampled counts and estimated population counts by region are presented in ***Table 5***.

Table 5. Sample Counts and Estimated Population Counts by Region

	Sample Counts	Estimated Population Counts
Schools	90	2,883
Standard 2 classrooms	90	5,046
Standard 4 classrooms	90	5,046
Standard 2 children	901	277,102
Standard 4 children	894	214,855

Unless explicitly stated, all reported estimates were calculated using the appropriate sample weights because the sample weights adjust for any under- or over-representation in the sample, making the estimates representative of the specified population.

4 Results and Findings

4.1 Reading Performance

Children's performance in reading was assessed using the MRA. The MRA results are presented in four parts: reading performance indicators, reading performance categories, detailed MRA results, and discussion of MRA results.

4.1.1 Reading performance indicators

The key purpose of this midline assessment was to measure the progress made by the USAID Tusome Pamoja Program since the baseline estimates in 2016. According to the Program's Performance Monitoring and Evaluation Plan, there are two key performance indicators to be reported for Standard 2 reading:

- P-001: Percent of learners who demonstrate reading fluency and comprehension of grade level text at the end of Standard 2 with USG assistance (ORF 45 cwpm)
- P-008a: Number and proportion (%) of learners who, after two years of schooling, obtain zero scores on oral reading fluency (ORF: 0 cwpm)

Accordingly, ***Table 6*** reports the percentage of students for both of these indicators. The three columns of the table display the baseline and midline estimates along with the target that was set at baseline.

Table 6. Percentage of Children Meeting Reading Benchmarks

	Baseline	Midline	Target
P-001: Percent of learners who demonstrate reading fluency and comprehension of grade level text at the end of Standard 2 with USG assistance (ORF: 45 cwpm)	1.5% (± 0.5%)	3.5% (± 2.5%)	6%
P-008a: Number and proportion (%) of learners who, after two years of schooling, obtain zero scores on oral reading fluency (ORF: 0 cwpm)	31.0% (±2.9)	17.2% (±4.1)	20%

While there was an increase in the percentage of students reading at benchmark (up to 3.5% from 1.5%), it was not large enough to meet the 6% target explicitly (though the target was within the estimate's confidence band). For non-readers, on the other hand, the percentage was reduced to 17.2%—which is below the 20% target.

The results for the two reading indicators, disaggregated by gender, are displayed in ***Table 7.*** This table shows that the margin between girls and boys remained similar from baseline to endline for indicator P-001 (approximately 1% difference) but that boys narrowed the gap in terms of zero scores. Specifically, boys reduced their percentage of zero scores by 16.5%, while girls reduced theirs by about 11%, leading to a closing of the gap between the two estimates.

Table 7. Percentage of Children Meeting Reading Benchmarks by Gender

	Baseline		Midline	
	Boys	Girls	Boys	Girls
P-001: Percent of learners who demonstrate reading fluency and comprehension of grade level text at the end of Standard 2 with USG assistance (ORF: 45 cwpm)	1.0% ±0.5%	2.0% ±0.8%	2.8% ±3.2%	4.1% ±3.3%
P-008a: Number and proportion (%) of learners who, after two years of schooling, obtain zero scores on oral reading fluency (ORF: 0 cwpm)	34.5% ±3.4%	27.6% ±3.5%	18% ±4.5%	16.4% ±5.7%

4.1.2 Reading performance categories

In order to further understand the shift in student performance from baseline to midline, the MRA results were combined into a single, four-category measure of readers with regard to their performance on the oral reading fluency and reading comprehension subtasks: Non-Readers, Beginning Readers, Progressing Readers, and Proficient Readers (***Table 8***). Proficient Readers were those who were able to correctly read 45 or more words of the story in one minute and with 80% or more comprehension. Conversely, Non-Readers were students who were unable to read a single word of the reading passage. Beginning Readers were correctly reading between 1 and 29 words per minute, and Progressing Readers were correctly reading at least 30 correct words per minute. ***Table 8*** shows that the large reduction in Non-Readers was spread throughout the ensuing three categories (with the largest change in Beginning Readers (~7%), followed by Progressing Readers (~4%), and finally Proficient Readers (~2%). This shows that significant progress is being made,

but that it will take more time to be able to move a larger proportion of students from the lowest categories to the highest.

Table 8. Reading Categories by Grade

Category	Type of Reader	Characteristic	Baseline	Midline
1	Non-Readers	Unable to read a single word of the story	31.0%	17.2%
2	Beginning Readers	Read between 1 and 29 words of the story correctly in one minute	57.3%	64.7%
3	Progressing Readers	Read at least 30 words of the story correctly in one minute	10.4%	14.8%
4	Proficient Readers	Read at least 45 words of the story correctly in one minute and with 80% or more comprehension	1.2%	3.4%

4.1.3 MRA results in detail

Detailed results from all of the MRA subtasks, comparing baseline and midline, are displayed in ***Table 9***. This table provides three measures for these subtasks: 1) percentage of children who were unable to correctly identify a single item; 2) average fluency scores (items per minute) for oral reading; 3) the average scores as measured by the percentage of correct items out of all the reading comprehension items. It is clear from this table the largest impacts are in terms of reducing zero scores for oral reading. However, reading comprehension zero scores also dropped about 9%, while mean scores (average percent correct) increased by approximately the same amount. This shows that performance in terms of oral reading fluency and reading comprehension has significantly improved since baseline.

Table 9. Comparison of MRA Performance by Subtask

MRA Subtask	Percentage of Students Who Scored Zero		Average Fluency Score		Average % Correct of Items	
	Baseline	Midline	Baseline	Midline	Baseline	Midline
Oral Reading Fluency	31% (±2.9)	17.2% (±4.1)	12.9 correct words per minute (±0.9)	16.4 correct words per minute (±1.9)	–	–
Reading Comprehension	43.2% (±3.1)	34.4% (±5.1)	–	–	24.6% (±1.7)	33.1% (±3.7)

The midline MRA results, disaggregated by gender, are displayed in ***Table 10***. It is clear from this table that girls slightly outperformed boys in oral reading and reading comprehension (with statistically significant differences in oral reading fluency and reading comprehension percent correct).

Table 10. Overview of Midline MRA Performance by Subtask and Gender

MRA Subtask	Percentage of Students Who Scored Zero		Average Fluency Score		Average % Correct of Items	
	Boys	Girls	Boys	Girls	Boys	Girls
Oral Reading	18.0% (±4.5)	16.4% (±5.7)	15 correct words per minute (±2.1)	17.8 correct words per minute (±2.2)	–	–
Reading Comprehension	38.8% (±7.4)	30.2% (±6.6)	–	–	30.7% (±4.8)	35.5% (±4.3)

4.1.4 Discussion of MRA results

Student performance on the MRA subtasks indicates that there was improvement from baseline to midline, but there is a continued need for even greater improvement in order to meet all targets by endline. While it is possible for students to learn to read a passage at the Standard 2 difficulty level with fluency and comprehension, the proportions of readers at this level remain low. This is largely a result of the starting ability level of students at baseline. In other words, with so few Progressing Readers it is difficult to significantly increase the proportion of Proficient Readers. At the other end of the spectrum, however, USAID Tusome Pamoja has shown a significant decrease in the proportion of Non-Readers (to a level that exceeds the midline target).

4.2 Mathematics Performance

Performance in mathematics was assessed using the MMA described in detail in Section 3.2. The MMA results are presented in three parts: mathematics performance indicators, detailed MMA results, and discussion of the MMA results.

4.2.1 Mathematics performance indicators

USAID Tusome Pamoja only has one reporting indicator for mathematics. In order to define this indicator, the following calculations were used:

- P-002: Number and proportion (%) of learners who, after two years of schooling can solve grade-level arithmetic problems:
  - Addition and Subtraction (Level 2): 80%; and
  - Missing Number: 60%

The results for indicator P-002 are displayed in ***Table 11.*** The table shows that there was growth from baseline (1.4%) to midline (4.3%), but that these improvements fell shy of the 6% target explicitly (though were once again within the estimate's confidence band). It is promising to note that some improvements were made, considering that USAID Tusome Pamoja has only just begun the mathematics portion of its intervention.

Table 11. Percentage of Children Meeting Mathematics Benchmark

	Baseline	Midline	Target
P-002: Number and proportion (%) of learners who, after two years of schooling, can solve grade-level arithmetic problems (Addition/Subtraction Level 2: 80%, and Missing Number: 60%)	1.4% (± 0.6%)	4.3% (± 2.4%)	6%

In ***Table 12*** these results are disaggregated by gender. The table suggests that there was no gap at baseline, but that slightly more boys were meeting the benchmark at midline than were girls. However, it should also be noted that this difference was not statistically significant, as it was well within the margin of error of the estimates.

Table 12. Percentage of Children Meeting Mathematics Benchmark by Gender

	Baseline		Midline	
	Boys	Girls	Boys	Girls
P-002: Number and proportion (%) of learners who, after two years of schooling, can solve grade-level arithmetic problems (Addition/Subtraction Level 2: 80%, and Missing Number: 60%)	1.4% ±0.8%	1.5% ±0.9%	5.1% ±3.7%	3.6% ±2.9%

4.2.2 Detailed MMA results

Complementing the mathematics performance indicator, this section provides an overview of the results for all MMA subtasks. ***Table 13*** summarizes baseline and midline performance on MMA. This shows that while there were small improvements in the rates of the most basic tasks (Addition and Subtraction Level 1), there were significantly larger improvements in Addition and Level 2 (for which the mean scores more than doubled).

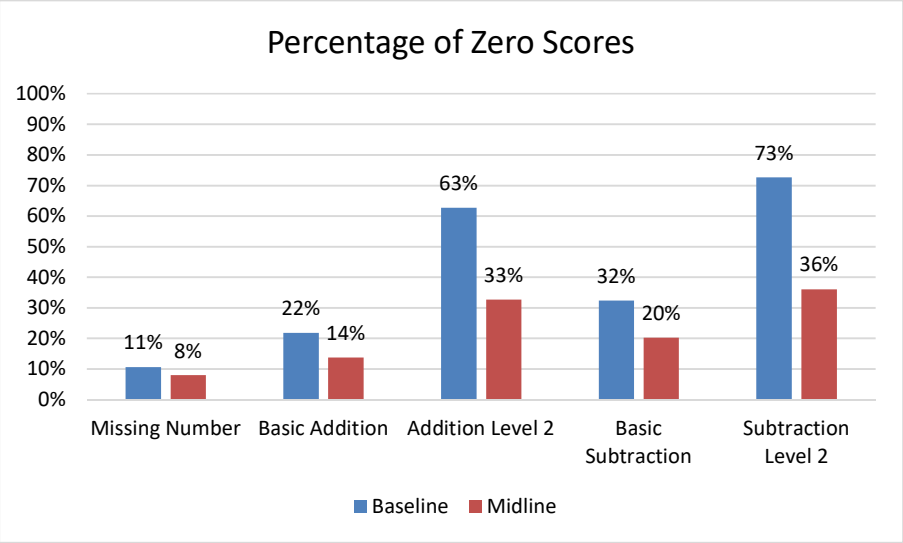
Table 13. Student Performance on MMA Subtasks

Subtask	Assessment	Mean Score
Addition (Level 1)	Baseline	6.6 per minute (±0.3)
	Midline	7.5 per minute (±0.7)
Subtraction (Level 1)	Baseline	4.9 per minute (±0.3)
	Midline	5.7 per minute (±0.7)
Missing number	Baseline	23.3% (±1.0)
	Midline	28.1% (±2.6)
Addition (Level 2)	Baseline	16.7% (±1.7)
	Midline	38.8% (±4.4)
Subtraction (Level 2)	Baseline	10.5% (±1.4)
	Midline	32.8% (±6.0)

While mean scores provide an important estimate for average performance, ***Figure 1*** summarizes the change in performance at the lower end of MMA subtasks (i.e.,

reductions in zero scores). It is clear from this figure that progress was made on all subtasks, but that the largest reductions in zero scores came in the subtasks of Addition Level 2 and Subtraction Level 2.

Figure 1. Comparison of MMA Zero Scores by Subtask



4.2.3 Discussion of the MMA results

There was improvement across all MMA subtasks from baseline to midline. In terms of the mathematics reporting indicator, the percentage of students at benchmark increased to 4.3%. While this was not equal to the target set at baseline, this level of improvement is promising given the fact that the 2019 school year will be the first with a comprehensive mathematics intervention across all regions. Additionally, the results show that students are making particularly good progress in the higher-order skills (such as Addition and Subtraction Level 2). Considering that doing mathematics (especially in the early grades) involves being able to read, write, compare, and perform basic operations with numbers, it seems that some of these conceptual skills may have benefited from the increased reading ability of students.

In terms of gender, boys performed slightly better than girls, but overlapping confidence intervals for all of these tasks makes it impossible to claim that differences in performance are statistically meaningful.

It is recommended that USAID Tusome Pamoja in-service teacher training in mathematics focus more on the multi-dimensional nature of mathematical proficiency than on more efficient teaching strategies that focus on mathematics as the memorization of facts, rules, and procedures.

5 Conclusion

The 2018 midline assessment was designed to accomplish one goal:

- Determine the level of impact from the USAID Tusome Pamoja Program across reading and mathematics indicators for Standard 2 by comparing program-level estimates from the 2016 baseline and the 2018 midline.

The results of the USAID Tusome Pamoja midline assessment show that gains were made across all indicators and subtasks. While the gains for the oral reading fluency and mathematics benchmark indicators were apparent, they were not large enough to meet the midline targets (though they were not statistically different from them based on the confidence intervals for the estimates). While some of the explanation for the lower than expected performance may be the result of increased strains on resources to due large increases in student enrollment (average Standard 2 enrollment in sampled school at baseline was 66 students, while it was 117 students at midline), it should also be noted that larger than expected gains were made at the lower end of the performance spectrum (i.e., in reducing zero scores). This means that students are beginning to move out of the lowest categories of performance, but that it has been difficult to move a large proportion of the students into the highest performance category (which is expected based on the skewed distribution and the large proportion of students at the lower end).

Reading results indicate that it is clear children need more practice to advance from laborious, word-by-word reading to reading in meaningful phrases; mathematics results show that there must be a shift in teaching from teaching for memorization of facts, rules, and formulas to teaching for understanding

Annex A. Final Instruments

Database ID: _____ **Namba ya utambuzi: _____**

Midline Mathematics Assessment – Tusome Pamoja Midline 2018

Maelekezo kwa ujumla

Fanya utangulizi/utambulisho rafiki na mwanafunzi kwa namna ya mchezo kwa mazungumzo mafupi kati yenu. ***(ona mfano hapo chini)***. . Mwanafunzi anatakiwa ahisi maswali anayoulizwa ni kama vile anacheza badala ya kujihisi anatahiniwa. Tumia muda huu kumuuliza ni lugha ipi atakuwa huru kuitumia. Soma taratibu kwa sauti na kwa ufasaha maelezo yaliyomo NDANI ya kisanduku tu.

Establish a playful and relaxed rapport with the child through a short conversation. The child should perceive the assessment almost as a game to be enjoyed rather than a test. Use this time to identify in what language the child is most comfortable communicating. Read aloud slowly and clearly ONLY the sections in boxes.

Habari ya asubuhi. Jina langu ni __ naishi _ . Napenda nikueleze kidogo habari zangu. [Eleza idadi ya watoto ulionao, umri wao, mchezo unaoupenda, kipindi cha redio au luninga unachopendelea, n.k.] Good morning. . *My name is ____ and I live in _____. I'd like to tell you a little bit about myself.*
[Number and ages of children; favourite sport, radio or television program, etc.]

1. **Unapenda kufanya nini unapotoka shuleni?** [Subiri jibu lake; kama mwanafunzi hataki kusema lolote, uliza swali la pili, lakini kama anaonekana angependa kuendelea kuongea basi endelea kupata ridhaa yake kwa maneno]
   What do you like to do when you are not in school?
   [Wait for response; if student is reluctant, ask question 2, but if they seem comfortable continue to verbal consent].
2. **Unapenda kucheza michezo ipi?** *What games do you like to play?*

A.**Tarehe ya tathmini** Date of assessment :	
B. **Jina la mtafiti/ /namba:** Assessor name/code :	
C. **JINA la shule na mahali ilipo:** NAME and location of school :	
D. **Namba maalum ya Shule:**	

F. **Darasa la mwanafunzi**: Student's grade level :	○ 1 = **Darasa la 1** Standard 1 ○ 2 = **Darasa la 2** Standard 2 ○ 3 = **Darasa la 3** Standard 3 ○ 4 = **Darasa la 4** Standard 4
G. **Umri la mwanafunzi**: Student's age:	**Umri/**Age: _____

Unique School code :	
E. **Namba maalum ya Mwanafunz:** Unique student code :	

H. **Jinsi ya Mwanafunzi:** Student's gender :	○ **msichana** /girl	○ **mvulana/** boy

Kazi 1: Namba inayokesekana: Zoezi kwa vitendo / Missing number: Practice	🕮 A1	🕐×
🗣 **Hapa pana baadhi ya namba 5,10,15. Namba gani itaingia hapa?** Here are some numbers. 5, 10, and 15. What number goes here? **5** **10** **15** **(20)** ✓🗣 **Vizuri ni 20. Tuendelee na swali jingine**. That's correct, 20. Let's do some more. ×🗣 **Namba 20 itakuwa hapa katika mfululizo huu. Tusome namba hizi kwa pamoja. [*Onesha kila namba*] . . . 5, 10, 15, 20. Tuendelee na swali jingine.** The number 20 goes here. Say the numbers with me. *[Point to each number]* 5, 10, 15, 20. 20 goes here. Let's do some more.		✋× ➲×

Kazi 1: Namba inayokosekana / Missing number	🕮 A2 & A3	🕐×
🗣 **Hapa kuna namba zaidi. [*Onesha katika sanduku*] ni namba gani inaingia hapa? [*Rudia kwa kila namba.*]** Here are some more numbers. [*Point to the box*] What number goes here? [Repeat for each item.] ✍ (✓) 1 = **Sahihi** / Correct (✓) 0 = **Jibu sio sahihi au halijajibiwa** / Incorrect or no response 1 — 5 6 7 **(8)** — 1 0 2 — 13 14 **(15)** 16 — 1 0 3 — 20 **(30)** 40 50 — 1 0 4 — **(300)** 400 500 600 — 1 0 5 — 2 4 6 **(8)** — 1 0 6 — 318 319 **(320)** 321 — 1 0 7 — 28 **(26)** 24 22 — 1 0 8 — 45 50 **(55)** 60 — 1 0 9 — 550 540 530 **(520)** — 1 0 10 — 9 14 **(19)** 24 — 1 0		✋ • **Endapo mwanafunzi amekosea mara 4 mfululizo. /** If the child makes 4 successive errors ➲ • **Endapo mwanafunzi hajajibu swali baada ya sekunde 5.** If the child doesn't respond after 5 seconds.

Kazi 2a:Kujumlisha: Hatua 1 / Addition: Level 1 | 🕮 B1 & B2 | 🕐 **Sekunde 60** seconds

🗣 **Hapa kuna maswali ya kujumlisha [*Onesha kwa mkono kutoka juu hadi chini*].**
Nitakwambia wakati wa kuanza na wa kumaliza. Toa majibu kwa kila swali na kama hujui jibu, endelea na swali linalofuata. Upo tayari?
Anzia hapa [*Onesha swali la kwanza*].
Here are some addition problems *[Glide hand from top to bottom]*. I will tell you when to start and when to stop. Say the answer for each problem. If you don't know an answer, move to the next problem. Are you ready?
Start here *[Point to first problem]*.

✍ **(/) Jibu sio sahihi au halijajibiwa /** Incorrect or no response
( ]) Baada ya swali la mwisho kusomwa/ After the last number read

4 + 1 = (5)	**12 + 3 = (15)**
2 + 4 = (6)	**13 + 4 = (17)**
6 + 2 = (8)	**11 + 6 = (17)**
4 + 5 = (9)	**6 + 5 = (11)**
3 + 3 = (6)	**10 + 4 = (14)**
8 + 1 = (9)	**9 + 7 = (16)**
7 + 3 = (10)	**6 + 6 = (12)**
2 + 7 = (9)	**5 + 6 = (11)**
5 + 5 = (10)	**8 + 7 = (15)**
1 + 9 = (10)	**8 + 10 = (18)**

✍ **Muda uliobaki (sekunde) /** Time left (seconds):

- **Endapo muda umekwisha (sekunde 60). /** If the time on the stopwatch runs out (60 seconds).

➲

- **Endapo mwanafunzi hajajibu swali baada ya sekunde 5. /** If the child doesn't respond to an item after 5 seconds.

Kazi 2b: Kujumlisha: Hatua 2 /Addition: Level 2	🕮 B3	🕐×
Karatasi na kalamu / Paper and pencil		
🗣 **Hapa kuna maswali ya kujumlisha [*Onesha kwa mkono kutoka juu hadi chini*].** **Nitakwambia wakati wa kuanza na wa kumaliza. Toa majibu kwa kila swali na kama hujui jibu, endelea na swali linalofuata. Upo tayari?** **Anzia hapa [*Onesha swali la kwanza*].** Here are some addition problems *[Glide hand from top to bottom]*. I will tell you when to start and when to stop. Say the answer for each problem. If you don't know an answer, move to the next problem. Are you ready? Start here *[Point to first problem]*.		✋ • **Endapo mtoto hajajibu swali lolote la hatua ya kwanza kwa usahihi. .** If the child did not answer any Level 1 question correctly. • **Endapo mwanafunzi amekosea mara 4 mfululizo. .** If the child makes 4 consecutive errors.
✍ **(✓) 1 = Sahihi /** Correct **(✓) 0 = Jibu sio sahihi au halijajibiwa /** Incorrect or no response **12 + 6 = (18)** **1** **0** **18 + 7 = (25)** **1** **0** **11 + 27 = (38)** **1** **0** **25 + 35 = (60)** **1** **0** **36 + 27 = (63)** **1** **0** **Mwanafunzi /** The pupil: ☐ **Ametumia vidole au kutali/** Used fingers or tallies ☐ **Ametumia karatasi na penseli /** Used paper & pencil ☐ **Amejibu maswali kwa kichwa /**Solved the problem(s) in his/her head		➲ • **Iwapo mwanafunzi atatumia njia isiyoridhisha (Mf; kuchora chora vimstari) muulize mwanafunzi kama anaweza kutumia njia nyingine? /** If the child uses an inefficient strategy (e.g., tick marks), ask the child "Do you know another way to solve the problem?" • **Endapo mwanafunzi ataendelea kutumia njia isiyosahihi au atasimama/ atakwama kwa sekunde tano katika swali hilo. .** If a child continues to use an inefficient strategy or stops on an item for 5 seconds.

Kazi 3A: Kutoa: Hatua ya 1 / Subtraction: Level 1	🕮 C1 & C2	🕐 **Sekunde 60**/60 seconds
Standards 1, 2, 3 and 4		

🗣 **Hapa kuna maswali ya kutoa [*Onesha kwa mkono kutoka juu hadi chini].***
Nitakwambia wakati wa kuanza na wa kumaliza. Toa majibu kwa kila swali na kama hujui jibu, endelea na swali linalofuata. Upo tayari?
Anzia hapa [*Onesha swali la kwanza*].
Here are some subtraction problems *[Glide hand from top to bottom]*.
I will tell you when to start and when to stop. Say the answer for each problem. If you don't know an answer, move to the next problem. Are you ready?
Start here *[Point to first problem]*.

✍ **(/) Jibu sio sahihi au halijajibiwa /** Incorrect or no response
( ]) Baada ya swali la mwisho kujibiwa / After the last number read

3 – 2 = (1)	**12 – 1 = (11)**
5 – 3 = (2)	**15 – 3 = (12)**
6 – 2 = (4)	**17 – 4 = (13)**
9 – 5 = (4)	**14 – 4 = (10)**
8 – 4 = (4)	**16 – 10 = (6)**
9 – 1 = (8)	**13 – 5 = (8)**
10 – 4 = (6)	**14 – 5 = (9)**
9 – 7 = (2)	**11 – 7 = (4)**
10 – 5 = (5)	**13 – 8 = (5)**
10 – 9 = (1)	**16 – 8 = (8)**

✍ **Muda uliobaki (sekunde) /** Time left (seconds):

✋
- **Endapo muda umekwisha (sekunde 60). .**
  If the time on the stopwatch runs out (60 seconds).

➲
- **Endapo mwanafunzi hajajibu swali baada ya sekunde 5. .**
  If the child doesn't respond to an item after 5 seconds.

Kazi 3b: Kutoa: Hatua ya 2 / Subtraction: Level 2	🕮 C3	🕐✖
Karatasi na kalamu /Paper and pencil		
🗣 **Hapa kuna maswali ya kutoa [*Onesha kwa mkono kutoka juu hadi chini].*** **Nitakwambia wakati wa kuanza na wa kumaliza. Toa majibu kwa kila swali na kama hujui jibu, endelea na swali linalofuata. Upo tayari?** **Anzia hapa [*Onesha swali la kwanza*].** Here are some subtraction problems *[Glide hand from top to bottom]*. I will tell you when to start and when to stop. Say the answer for each problem. If you don't know an answer, move to the next problem. Are you ready? Start here *[Point to first problem]*.		✋ • **Endapo mtoto hajajibu swali lolote la hatua ya kwanza kwa usahihi. .** If the child did not answer any Level 1 question correctly. • **Endapo mwanafunzi amefanya makosa manne (4) mfululizo. .** If the child makes 4 consecutive errors.
🖉 **(✓) 1 = Sahihi /** Correct **(✓) 0 = Jibu sio sahihi au halijajibiwa /** Incorrect or no response **18 – 4 = (14)** **1** **0** **25 – 7 = (18)** **1** **0** **36 – 12 = (24)** **1** **0** **40 – 17 = (23)** **1** **0** **46 – 29 = (17)** **1** **0** **Mwanafunzi /** The pupil: ☐ **Ametumia vidole au kutali/** Used fingers or tallies ☐ **Ametumia karatasi na penseli /** Used paper & pencil ☐ **Amejibu maswali kwa kichwa /**Solved the problem(s) in his/her head		➲ • **Iwapo mwanafunzi atatumia njia isiyoridhisha (Mf; kuchora chora vimstari) muulize mwanafunzi kama anaweza kutumia njia nyingine?** / If the pupil uses an inefficient strategy (e.g., tick marks), ask the child "Do you know another way to solve the problem?" • **Endapo mwanafunzi ataendelea kutumia njia isiyosahihi au atasimama/ atakwama kwa sekunde 5 katika swali hilo. .** If a pupil continues to use an inefficient strategy or stops on an item for 5 seconds.

Midline Reading Assessment – Tusome Pamoja Midline 2018

Maelekezo kwa ujumla

Fanya utangulizi/utambulisho rafiki na mwanafunzi kwa namna ya mchezo kwa mazungumzo mafupi kati yenu. ***(ona mfano hapo chini)***. Mwanafunzi anatakiwa ahisi maswali anayoulizwa ni kama vile anacheza badala ya kujihisi anatahiniwa. Tumia muda huu kumuuliza ni lugha ipi atakuwa huru kuitumia. Soma taratibu kwa sauti na kwa ufasaha maelezo yaliyomo NDANI ya kisanduku tu.

Establish a playful and relaxed rapport with the child through a short conversation. The child should perceive the assessment almost as a game to be enjoyed rather than a test. Use this time to identify in what language the child is most comfortable communicating. Read aloud slowly and clearly ONLY the sections in boxes.

Habari ya asubuhi. Jina langu ni __ naishi _ . Napenda nikueleze kidogo habari zangu. [Eleza idadi ya watoto ulionao, umri wao, mchezo unaoupenda, kipindi cha redio au luninga unachopendelea, n.k.] Good morning. *My name is ____ and I live in _____. I'd like to tell you a little bit about myself.*

[Number and ages of children; favourite sport, radio or television program, etc.]

1. **Unapenda kufanya nini unapotoka shuleni?** [Subiri jibu lake; kama mwanafunzi hataki kusema lolote, uliza swali la pili, lakini kama anaonekana angependa kuendelea kuongea basi endelea kupata ridhaa yake kwa maneno]

What do you like to do when you are not in school?

[Wait for response; if student is reluctant, ask question 2, but if they seem comfortable continue to verbal consent].

2. **Unapenda kucheza michezo ipi?** *What games do you like to play?*

A. **Tarehe ya tathmini** Date of assessment :	
B. **Jina la mtafiti/ /namba:** Assessor name/code :	
C. **JINA la shule na mahali ilipo:** NAME and location of school :	
D. **Namba maalum ya Shule:** Unique School code :	
E. **Namba maalum ya Mwanafunz:** Unique student code :	

F. **Darasa la mwanafunzi**: Student's grade level :	○ 1 = **Darasa la 1** Standard 1 ○ 2 = **Darasa la 2** Standard 2 ○ 3 = **Darasa la 3** Standard 3 ○ 4 = **Darasa la 4** Standard 4	
G. **Umri la mwanafunzi**: Student's age:	**Umri/**Age: _____	
H. **Jinsi ya Mwanafunzi:** Student's gender :	○ **msichana** /girl	○ **mvulana/** boy

<table>
<tr><th>Subtask 1a. ORAL READING PASSAGE</th><td>📖 A1
🕐 Sekunde 60
seconds</td><th>Subtask 1b. READING COMPREHENSION</th></tr>
<tr><td colspan="3">✋ 1. Iwapo mwanafunzi atashindwa kusoma japo neno moja kwa usahihi kabla ya neno lililomo katika kisanduku sema “Asante!”, sitisha zoezi hili, kisha weka alama kwenye kisanduku kilichopo hapa chini kisha endelea na zoezi linalofuata. Usimuulize maswali ya ufahamu. If the child does not provide a single correct word before the word in a box, say “Thank you!”, discontinue this subtask and check the box at the bottom. Do not ask any comprehension questions. 2. Iwapo mwanafunzi atasema “Sijui”, chukulia kama ni kosa. If a child says, “I don’t know,” mark as incorrect.</td></tr>
<tr><td>Muoneshe mwanafunzi karatasi yenye hadithi wakati unasoma maelekezo.
Show the child the sheet in the student stimulus booklet as you read the instructions.</td><td rowspan="3">➲ 3. Endapo mtoto hajajibu swali baada ya sekunde tatu.
If the child doesn’t respond to an item after 3 seconds.</td><td rowspan="2">Baada ya mwanafunzi kumaliza kusoma. ONDOA karatasi ya hadithi mbele yake.
Muulize mwanafunzi maswali yanayohusiana na hadithi aliyosoma. Mwanafunzi lazima asome hadithi ambayo inahusiana na maswali atakayoulizwa.
Iwapo mwanafunzi atashindwa kujibu swali baada ya sekunde kumi (10) weka alama ya ‘hakuna jibu’ na endelea kuuliza swali linalofuata. Usirudie kuuliza swali.
After the child is finished reading, REMOVE the passage from in front of the child.
Ask the child only the questions related to the text read. A child must read all the text that corresponds with a given question. If the child does not provide a response to a question after 10 seconds, mark “no response” and continue to the next question. Do not repeat the question.</td></tr>
<tr><td>🗣 Hii ni karatasi yenye hadithi fupi. Soma hadithi hii kwa sauti, haraka na kwa umakini: Ukimaliza kusoma nitakuuliza maswali kuhusu yale uliyosoma. Je umeelewa unachotakiwa kufanya. Nikisema “Anza” soma hadithi, haraka na kwa umakini kadri uwezavyo. Kama utaona neno usiloweza kusoma endelea kusoma neno linalofuata. Weka kidole chako kwenye neno la kwanza. Je, upo tayari? “Anza”.
Here is a short story. I want you to read it aloud, quickly but carefully. When you finish, I will ask you some questions about what you have read. When I say “Begin”, read the story as best as you can. If you come to a word you do not know, go on to the next word. Put your finger on the first word. Ready? Begin.
🕐 Anza kupima muda mara mwanafunzi anapoanza kusoma neno la kwanza.
Start the timer when the child reads the first word.
✋ Endapo muda umekwisha (sekunde 60). If the time on the stopwatch runs out (60 seconds).</td><td>🗣 Sasa nitaanza kuuliza maswali machache kuhusu hadithi uliyosoma. Jitahidi kujibu maswali vizuri kwa kadri uwezavyo. Now I will ask you a few questions about the story you just read. Try to answer the questions as well as you can.
✍ (✓) 1 = Sahihi / Correct
(✓) 0 = Isiyosahihi / Incorrect
(✓) . = Hakuna jibu / No response</td></tr>
<tr><td></td><td></td></tr>
</table>

✍ (/) **Weka alama ya mkwaju kwa maneno yote aliyokosea kuyasoma.** Mark any incorrect words with a slash. (Ø) **Iwapo uliweka alama ya kosa aliposahihisha kusoma neno alilokosea, zungushia neno kisha endelea.** Circle self-corrections if you already marked the word incorrect. ( ]) **Weka alama ya mabano katika neno la mwisho alilosoma.** Mark the final word read with a bracket.		One day Furaha went to the bush to pick fruits. On the way she saw an old woman sitting on the pathway. Near her was a basket of oranges. Some of them had spilled. The old woman was happy to see Furaha. She asked Furaha to remove a thorn that had pierced her leg. The old woman could not reach the thorn. Furaha took out the thorn slowly. The old woman thanked Furaha and gave her some oranges. 1. Where did Furaha go? [to the bush] 2. Who sat on the pathway? [an old woman] 3. What was in the basket? [oranges] 4. What did the old woman ask Furaha? [to remove a thorn] 5. Why did the old woman give Furaha oranges? [because she helped; she removed the thorn; she removed the thorn slowly]				
Piloted January 2016; Used February 2016 and August 2016			**Maswali [Majibu]** Questions [Answers]			
Siku moja Furaha alikwenda porini kuchuma matunda.	7		**Furaha alikwenda wapi?** [porini]	1	0	.
Njiani alimuona bibi amekaa barabarani.	12		**Nani aliketi barabarani?** [bibi]	1	0	.
Karibu yake lilikuwepo kapu la machungwa. Baadhi ya machungwa yalikuwa yamemwagika.	23		**Kwenye kapu kulikuwa na kitu gani?** [machungwa]	1	0	.
Bibi alifurahi kumuona Furaha. Bibi akamuomba Furaha amtoe mwiba mguuni.	33		**Bibi alimwomba Furaha afanye nini?** [amtoe mwiba]	1	0	.
Bibi hakuweza kuufikia ule mwiba. Furaha alimtoa mwiba polepole. Bibi alimshukuru Furaha na kumpa machungwa machache.	49		**Kwa nini Bibi alimpa Furaha machungwa?** [kwa sababu alimsaidia; alimtoa mwiba mguuni; alimtoa polepole]	1	0	.
✍ **Muda uliobaki (sekunde)** Time remaining (seconds)						
✍ **Sitisha zoezi kwa sababu mwanafunzi hana majibu kwenye mstari wa kwanza.** Exercise discontinued because the child had no correct answers in the first line.						

Annex B. Sample Methodology, Weighting, and Precision

This annex discusses the details of the sample, the population that it is meant to represent, and how the sample is properly representative of that population. It also discusses the precision estimates for the major outcome variables from which the sample size was derived.

Population of Interest and Sample Frame

The population of interest includes all primary government schools located in the five USAID Tusome Pamoja regions. More specifically, the population of interest was all Standard 2 and Standard 4 pupils enrolled in these 2,883 schools.

The sample frame used to draw the midline sample was created from two census lists of schools: one from the Mainland, the other from Zanzibar. The Mainland census data came from the 2016 Primary School Leaving Certificate Examination.[1]. The census data from Zanzibar came from the 2016 census list of schools collected by the Zanzibar Ministry of Education and Vocational Training. ***Table B.1*** provides the total number of schools in the list frame along with the total number of schools excluded for the given reason. Finally, it provides the total number of schools that make up the defined population. ***Table B.2*** provides the population of schools by region as well as the Standard 2 and Standard 4 enrollment. These updated pupil enrollment population figures by region and standard were based on the most up-to-date estimates from USAID Tusome Pamoja.

Table B.1. Schools Excluded from the List Frame Prior to Sampling

	# Schools	Percent
Total number of schools in the sample frame	16,141	
Not excluded [defined population]	**2,883**	**17.86**
Reason for exclusion		
Not a Tusome Pamoja Region	13,226	81.94
School was marked as non-government	32	0.2

Table B.2 Population Counts of Schools and Pupils by Tusome Pamoja Region

Region	Schools	Standard 2 Pupils	Standard 4 Pupils
Iringa	453	38,622	29,652
Morogoro	827	86,237	64,192
Mtwara	607	51,331	38,426
Ruvuma	721	58,057	44,466
Zanzibar	275	42,855	38,119
Total	**2,883**	**277,102**	**214,855**

[1] 2016 Primary School Leaving Certificate Examination : http://opendata.go.tz/dataset/primary-schools-leaving-certificate-examination-performance-ranking-with-location/resource/d4703a57-3fa1-4464-a459-683e3dfc5e04

Sample Methodology

The sample methodology was for a three-stage random sample of schools, classrooms, and pupils. Schools were separated by region, then sorted by district and number of pupils at the school. Schools were selected with probability proportional the number of pupils at the school. A total of 90 schools were selected (18 from each region), and for each selected school, 2 replacement schools were automatically selected. The replacement schools were the closest resemblance of the originally selected schools by district and enrollment. The replacement schools would replace the originally sampled school if the originally sampled school did not meet the requirements as defined by the population. Only 2 schools (both in Morogoro) needed to be replaced due to inaccessibility.

Once the assessment team arrived at the school, it randomly selected one Standard 2 classroom and one Standard 4 classroom. Within each of the selected Standard 2 and Standard 4 classrooms, 10 pupils were selected (5 boys and 5 girls). . ***Table B.3*** displays the sample methodology. ***Table B.4*** provides the final counts of completed school, classroom, and pupil assessments.

Table B.3. Sample Methodology for the Tusome Pamoja Midline 2018 Assessment. : Three-Stage Stratified Sample of Schools, Classrooms, and Pupils

Stage	Item Sampled (Expected Counts)	Stratification (Number of Strata)	Probability of Selection
1	Schools (90)	Region (5)	Proportional to enrollment*
2	Classroom (180)	Standard 2 /Standard 4 (2)	Equal
3	Pupils (1,795)	Boys/Girls (2)	Equal

*Proportional to Enrollment: For the Mainland data, enrollment was the total number of pupils who sat for the Primary School Leaving Certificate Examination in 2016. For the Zanzibar data, it was the Standard 2 enrollment for the 2016 census data.

Table B.4. Final Sample Counts of Assessments, Interviews, and Observations Collected for the Tusome Pamoja 2018 Midline Evaluation

Regions	Sampled Schools	Standard	Classrooms	Pupil Gender	Pupils Assessed
5 Tusome Pamoja Regions	90	Standard 2	90	Boys	448
				Girls	453
		Standard 4	90	Boys	446
				Girls	448
Total	**90**		**180**		**1,795**

Sample Weights

All sample weights were calculated as the inverse of the probability of selection at each stage of selection (school, classroom, and pupil). School weights were scaled to the known populations of schools by region. Classrooms were not scaled to the final population as that information was not available in the school census data. The Standard 2 and Standard 4 pupil weights were scaled to the population region and standard. We were unable to scale to the boy/girl pupil population because the Mainland sample frame did not provide enrollment by gender. ***Table B.5*** provides the weighted totals and the sample totals by region and standard.

Table B.5. Weighted Sample (Estimated Population) and Sample Counts by Each Stage Sampled (School, Classroom, Pupil)

Tusome Pamoja Region	*Weighted Sample (Estimated Population)*					*Sample*				
	School	Classroom		Pupils		Schools	Classroom		Pupils	
		S2	S4*	S2	S4		S2	S4	S2	S4
Iringa	453	688	688	38,622	29,652	18	18	18	180	180
Morogoro	827	1580	1580	86,237	64,192	18	18	18	180	177
Mtwara	607	1241	1241	51,331	38,426	18	18	18	180	177
Ruvuma	721	882	882	58,057	44,466	18	18	18	180	180
Zanzibar	275	655	655	42,855	38,119	18	18	18	181	180
Total	**2,883**	**5,046**	**5,046**	**277,102**	**214,855**	**90**	**90**	**90**	**901**	**894**

*S=standard; Standard 4 enrollment was not collected as part of the midline data collection. Therefore, Standard 2 values were used as proxies (thus leading to identical estimates for Standard 2 and Standard 4 in the weighted sample).

Precision Estimates

The Standard 2 and Standard 4 sample size calculations were derived using the oral reading fluency score and having a desired 95% confidence band of ±3.0 words per minute (wpm) by standard. With a predicted standard deviation of 25.0, and intraclass correlation of 0.25, the school sample size was estimated to be 90 schools across all five regions if we were to sample 10 pupils in each Standard 2 and

Standard 4. In other words, the number of schools was based on the ability to confidently estimate the average oral reading fluency score to within 3 correct wpm above or below the estimate. Larger sample sizes would have reduced this band, but the additional costs were not justified by the marginally increased precision. Additionally, it turns out that both Standard 2 and Standard 4 actually showed greater precision than that used for the sample calculation. . The Standard 2 estimate is very precise with ±1.9 wpm, and the Standard 4 precision estimate was ±2.4. ***Table B.6*** provides the estimates from the actual data.

Table B.6. Reading Fluency Means and Precision Estimate by Standard

Standard	n	Range (wpm)	Mean (wpm)	95% CI Band (wpm)	SD (wpm)	ICC	DEF
Std2	901	[0,76]	16.4	±1.9	14.2	0.245	2.2
Std4	893	[0,101]	36.7	±2.4	19.6	0.219	1.7
Overall	1794	[0,101]	25.3	±1.8	19.5	0.123	1.9

wpm: words per minute, CI: confidence interval, SD: Standard deviation, ICC: Intra class correlation, DEF: design effect

It should be noted that the midline sample size was never meant to be disaggregated beyond the Standard 2 and Standard 4 level. The midline data were never meant to be disaggregated by region or any other school, classroom, or pupil characteristic. For this reason, the midline sample was significantly smaller than the baseline sample—which was based on the need for disaggregation at the regional level.

Annex C. Inter-rater Reliability and Assessor Accuracy

The following table provides an overview of the percent agreement across all subtasks during the second assessor accuracy measure exercise.

Assessor Accuracy Measure Percent Agreement by Assessor								
Assessor	**Oral Reading**	**Reading Comprehension**	**Missing Number**	**Basic Addition**	**Addition Level 2**	**Basic Subtraction**	**Subtraction Level 2**	**Overall**
1	100%	100%	100%	100%	100%	100%	100%	100%
2	100%	100%	100%	100%	100%	100%	100%	100%
3	100%	100%	100%	100%	100%	100%	100%	100%
4	100%	100%	100%	100%	100%	100%	100%	100%
5	100%	100%	100%	100%	100%	100%	100%	100%
6	98%	100%	100%	100%	100%	100%	100%	100%
7	98%	100%	100%	100%	100%	100%	100%	100%
8	98%	100%	100%	100%	100%	100%	100%	100%
9	98%	100%	100%	100%	100%	100%	100%	100%
10	96%	100%	100%	100%	100%	100%	100%	99%
11	96%	100%	100%	100%	100%	100%	100%	99%
12	96%	100%	100%	100%	100%	100%	100%	99%
13	96%	100%	100%	100%	100%	100%	100%	99%
14	96%	100%	100%	100%	100%	100%	100%	99%
15	96%	100%	100%	100%	100%	100%	100%	99%
16	96%	100%	100%	100%	100%	100%	100%	99%
17	96%	100%	100%	100%	100%	100%	100%	99%
18	94%	100%	100%	100%	100%	100%	100%	99%
19	94%	100%	100%	100%	100%	100%	100%	99%
20	94%	100%	100%	100%	100%	100%	100%	99%
21	94%	100%	100%	100%	100%	100%	100%	99%
22	94%	100%	100%	100%	100%	100%	100%	99%
23	92%	100%	100%	100%	100%	100%	100%	99%
24	92%	100%	100%	100%	100%	100%	100%	99%
25	96%	100%	100%	95%	100%	100%	100%	99%
26	90%	100%	100%	100%	100%	100%	100%	99%
27	90%	100%	100%	100%	100%	100%	100%	99%
28	88%	100%	100%	100%	100%	100%	100%	98%
29	96%	100%	90%	100%	100%	100%	100%	98%
30	86%	100%	100%	100%	100%	100%	100%	98%
31	94%	100%	90%	100%	100%	100%	100%	98%
32	94%	100%	90%	100%	100%	100%	100%	98%

Assessor Accuracy Measure Percent Agreement by Assessor								
Assessor	**Oral Reading**	**Reading Comprehension**	**Missing Number**	**Basic Addition**	**Addition Level 2**	**Basic Subtraction**	**Subtraction Level 2**	**Overall**
33	88%	100%	100%	95%	100%	100%	100%	98%
34	92%	100%	90%	100%	100%	100%	100%	97%
35	90%	100%	90%	100%	100%	100%	100%	97%
36	80%	100%	100%	100%	100%	100%	100%	97%
37	98%	100%	100%	100%	80%	100%	100%	97%
38	96%	100%	100%	100%	80%	100%	100%	97%
39	90%	100%	80%	100%	100%	100%	100%	96%
40	96%	100%	90%	100%	80%	100%	100%	95%
41	82%	100%	90%	95%	100%	95%	100%	95%
42	98%	100%	80%	100%	100%	100%	80%	94%
43	88%	80%	90%	100%	100%	100%	100%	94%
44	96%	80%	80%	100%	80%	95%	100%	90%
45	90%	60%	70%	100%	80%	100%	80%	83%

The following table provides an overview of the percent agreement for oral reading and reading comprehension, obtained during the field inter-rater reliability activities (for each day of data collection). Each row in the table represents a pair of assessors who simultaneously scored a single student. These data were analyzed at the end of each day of data collection, and results were provided to the USAID Tusome Pamoja monitoring and evaluation team—who followed up with assessors in all cases where agreement was low. Overall, there was strong agreement throughout data collection, with an average of 97% agreement on oral reading and 98% agreement for reading comprehension.

Field Interrater Reliability Percent Agreement by Assessor			
ID	**Date**	**Percent Agreement for Oral Reading**	**Percent Agreement for Reading Comprehension**
RVGNSB	8/23/18	100%	100%
GGEQSG	8/23/18	86%	100%
TTRVQB	8/23/18	98%	100%
YQENNZ	8/23/18	100%	100%
BUQEGV	8/23/18	100%	100%
AZZXHW	8/23/18	94%	100%
VFFAFR	8/23/18	98%	100%
KEXNWE	8/23/18	100%	100%
YNKEGC	8/23/18	100%	100%
PVSQTQ	8/23/18	98%	100%
BCBSUR	8/23/18	100%	100%
CWQMQK	8/23/18	96%	100%
EZCTXA	8/24/18	94%	100%

Field Interrater Reliability Percent Agreement by Assessor			
ID	**Date**	**Percent Agreement for Oral Reading**	**Percent Agreement for Reading Comprehension**
EFFGXG	8/24/18	100%	0%
SPZRTH	8/24/18	100%	100%
ZBVBEG	8/24/18	96%	100%
KRNSUU	8/24/18	100%	100%
PUBSRM	8/24/18	94%	100%
KBWWRW	8/24/18	100%	100%
SZVBUE	8/24/18	94%	100%
HXQUQH	8/24/18	98%	100%
VKCBKV	8/24/18	98%	80%
NNBAYH	8/24/18	100%	100%
MHFWTP	8/24/18	100%	100%
BAPQBW	8/27/18	100%	80%
NQHNTH	8/27/18	96%	100%
XCWQUE	8/27/18	98%	100%
TXMSXG	8/27/18	84%	100%
NRVFBC	8/27/18	98%	100%
ENZSFB	8/27/18	98%	80%
QBSTQX	8/27/18	98%	100%
TEQFVR	8/27/18	100%	100%
VBHHWZ	8/27/18	98%	100%
RFZSRZ	8/27/18	100%	100%
UEZGVH	8/27/18	86%	100%
TKQWZG	8/27/18	86%	100%
RWHYKY	8/27/18	96%	100%
CGAMQR	8/27/18	100%	100%
VHNNWU	8/27/18	98%	100%
GTRFPU	8/27/18	86%	100%
SUUGAG	8/28/18	100%	100%
FVETTN	8/28/18	98%	100%
TTMBMM	8/28/18	100%	100%
NFVHPZ	8/28/18	100%	100%
VYCXCU	8/28/18	100%	100%
SUQRZK	8/28/18	92%	100%
WSYUYF	8/28/18	98%	100%
KQMEYS	8/28/18	98%	100%
CASCEC	8/28/18	98%	100%
EBFUGE	8/28/18	96%	100%
EBFQVX	8/28/18	98%	100%
KZFZUZ	8/28/18	96%	100%
HUFSGR	8/28/18	100%	100%

Field Interrater Reliability Percent Agreement by Assessor			
ID	**Date**	**Percent Agreement for Oral Reading**	**Percent Agreement for Reading Comprehension**
AZAKQM	8/28/18	98%	100%
CNMGZX	8/28/18	100%	100%
MWWHKQ	8/28/18	100%	100%
YFSQUN	8/29/18	100%	100%
HSMAWZ	8/29/18	100%	100%
ZHBNPT	8/29/18	100%	100%
RBMEYY	8/29/18	96%	100%
SWVQZF	8/29/18	100%	100%
QARAWY	8/29/18	100%	100%
HACHYT	8/29/18	94%	100%
TRPYRR	8/29/18	98%	100%
FBKFXF	8/29/18	100%	100%
HWCKEY	8/29/18	100%	100%
BWRPYQ	8/29/18	100%	100%
NYQKYH	8/30/18	100%	100%
FSKCUN	8/30/18	90%	100%
ZEQSVR	8/30/18	100%	100%
CRQGSQ	8/30/18	94%	100%
EUVAXQ	8/30/18	100%	100%
ECMXHM	8/30/18	100%	100%
XZYVFE	8/30/18	100%	100%
PRSTEY	8/30/18	94%	100%
GMCXCE	8/30/18	100%	100%
FCPKEH	8/30/18	98%	100%
RFFMMB	8/30/18	100%	100%
ZUBGMF	8/31/18	97%	98%
SRXHCN	8/31/18	100%	100%
NRRRWG	8/31/18	82%	100%
CTPZNX	8/31/18	100%	100%
YCGMPT	8/31/18	98%	100%
ZTRPPR	8/31/18	100%	100%
TMPFRK	8/31/18	98%	100%
PBKYZR	8/31/18	96%	100%
OVERALL		97%	98%