



FINAL

*TZ21 Endline Evaluation Report:
Evaluation conducted in Zanzibar,
Mtwara, and Lindi*

21st Century Basic Education Program (TZ21)

With support from USAID



For Creative Associates International, Inc.
By School-to-School International

February 2015

Contents

Contents.....	1
Acknowledgements	5
Acronyms	6
Executive Summary	7
Introduction	11
Purpose of endline	11
Methodology	11
Research questions.....	12
Sample.....	12
Instrument development	13
EGRA development	13
Administration.....	16
Training of Data Collectors	16
Data Collection.....	16
Data Entry and Cleaning.....	17
Data Analysis	17
Reporting, presentations, and validation of findings	17
Findings	18
Question 1: What changes can be observed in the teaching and learning of reading in TZ21 project schools in Mtwara and Zanzibar from baseline to endline?	18
EGRA results, midline to endline, Zanzibar.....	18
EGRA results, midline to endline, Mtwara.....	23
EGRA results, baseline to endline	27
Teaching and learning in TZ21 schools: Zanzibar.....	30
Teaching and learning in TZ21 schools: Mtwara.....	33
Question 2: What differences can be observed in the teaching and learning of reading in TZ21 schools in Mtwara as compared to schools in Lindi, a non-TZ21 region, from baseline to endline?.....	35
EGRA results	35
Results from classroom observations	40
Question 3: To what extent are students, teachers, head teacher, schools and communities implementing practices or activities to support reading?	41
Practices & activities supporting reading, Zanzibar	41
Practices & activities supporting reading, Mtwara: Head Teachers' views	42
Question 4: Which factors at the student, classroom, and school levels are associated with strong reading performance?	42
Discussion.....	45
Recommendations.....	47

Annex A: EGRA passages	49
Annex B: Test and item statistics, EGRA	52
Annex C: Results of T-tests of differences between time points, EGRA Standard 2, Timed Reading Passage (60 seconds)	53
Annex D: Minimum Performance Criteria	56
Annex E: Practices supporting reading	58
Practices & activities supporting reading, Zanzibar: Teachers' views	59
Practices & activities supporting reading, Zanzibar: SMCs' views	61
Practices & activities supporting reading, Mtwara: Teachers' views	66
Practices & activities supporting reading, Mtwara: SMCs' views	68
Annex F: Additional discussion of findings: student, teacher, and Head Teacher characteristics	72
Teacher characteristics and performance, Zanzibar	72
Head Teacher characteristics and performance, Zanzibar	73
Teacher characteristics and performance, Mtwara	74
Head Teacher characteristics and performance, Mtwara	75

List of Tables

Table 1: Number of participants included in the midline	13
Table 2: EGRA Tasks by Test Administration	15
Table 3: Task statistics: Percent correct by task, Zanzibar.....	19
Table 4: reading summary score by gender	20
Table 5: Percent correct scores by task and gender, highlighting differences in achievement by gender.....	21
Table 6: Number and percent of students by reading performance level, Zanzibar.....	21
Table 7: Percent of students by timed reading performance level and comprehension level- midline	22
Table 8: Percent of students by reading performance level and comprehension level- endline	22
Table 9: Task statistics: Percent correct by task, Mtwara	24
Table 10: EGRA summary score by Gender, Mtwara	25
Table 11: Percent correct scores by task and gender, highlighting differences by gender, Mtwara.....	25
Table 12: Number and percent of students by reading performance level (CWPM), Mtwara	26
Table 13: Percent of students by reading performance level and comprehension level - midline, Mtwara.....	27
Table 14: Percent of students by reading performance level and comprehension level - endline, Mtwara	27
Table 15: Percent correct by EGRA task, Mtwara vs. Lindi	36
Table 16: Overall EGRA Scores by Gender, Mtwara vs. Lindi	37
Table 17: Midline to Endline EGRA Gains by Task and Gender – Mtwara vs. Lindi.....	38
Table 18: Gains in fluency from midline to endline – Mtwara vs. Lindi.....	38
Table 19: Zero Scores from midline to endline – Mtwara vs. Lindi.....	39
Table 20: Use of reading strategies in Mtwara and Lindi: Changes in frequency of observations from midline to endline	40
Table 21: EGRA performance and student, teacher, and Head Teacher characteristics	44

Table 22: Task statistics: Means and item-total correlations.....	52
Table 23: Task statistics: p-values and item-totals, Mtwara.....	52
Table 24: Student behaviors observed in the classroom, midline to endline, Mtwara vs. Lindi	56
Table 25: Teacher behaviors observed in the classroom, midline to endline, Mtwara vs. Lindi	57
Table 26: Learning materials used in classroom observation lesson, midline to endline, Mtwara vs. Lindi	57
Table 27: How Head Teachers help teachers teach better - Head Teacher responses from most frequent to least frequent, Zanzibar	58
Table 28: Steps Head Teachers take to help their teachers improve reading instruction - Head Teacher responses from most frequent to least frequent, Zanzibar	59
Table 29: How teachers know when students are having trouble with reading, teacher responses from most frequent to least frequent.....	60
Table 30: How teachers help students who are having trouble learn to read - teacher responses from most frequent to least frequent, Zanzibar	61
Table 31: Topics discussed at teacher meetings – teacher responses from most frequent to least frequent, Zanzibar	61
Table 32: SMC main functions – Member responses from most frequent to least frequent, Zanzibar	62
Table 33: Issues raised at most recent SMC meeting – Member responses from most frequent to least frequent, Zanzibar	62
Table 34: SMC roles in promoting quality in school - Member responses from most frequent to least frequent, Zanzibar	64
Table 35: How Head Teachers help teachers teach better – Head Teacher responses from most frequent to least frequent, Mtwara	65
Table 35: Steps Head Teachers take to help their teachers improve reading instruction - Head Teacher responses from most frequent to least frequent, Mtwara.....	66
Table 37: How teachers know students are having trouble learning to read - Teacher responses from most frequent to least frequent, Mtwara	67
Table 38: How teachers help students who are having trouble learn to read - Teacher responses from most frequent to least frequent, Mtwara	67
Table 39: Topics discussed at teacher meetings - Teacher responses from most frequent to least frequent, Mtwara	68
Table 40: SMC main functions - Member responses from most frequent to least frequent, Mtwara.....	69
Table 41: Issues raised at most recent SMC meeting - Member responses from most frequent to least frequent, Mtwara.....	70
Table 42: SMC roles in promoting quality in school - Member responses from most frequent to least frequent, Mtwara	71
Table 43: EGRA summary scores correlated with student characteristics, Zanzibar	72
Table 44: EGRA summary scores correlated with teacher characteristics, Zanzibar	73
Table 45: EGRA summary scores correlated with student characteristics, Mtwara	74
Table 46: EGRA summary scores correlated with teacher characteristics, Mtwara	75
Table 47: EGRA summary scores correlated to Head Teacher characteristics, Mtwara	76

List of Figures

Figure 1: TZ21 midline sampling plan, number of schools.....	12
Figure 2: Student’s ages, Standard 2, Zanzibar.....	18
Figure 3: Comparison of overall EGRA scores, midline to endline, Zanzibar	18
Figure 4. Student's age, Mtwara	23

Figure 5: Comparison of overall EGRA scores, midline to endline, Mtwara	23
Figure 6: EGRA mean scores by task, Mtwara.....	24
Figure 7: Number and percent of students by reading performance level (CWPM), Mtwara	26
Figure 8: Average correct words read per minute, timed reading task, baseline to endline .	28
Figure 9: Percentage of zero scores from baseline to endline, Zanzibar.....	29
Figure 10: Percentage of zero scores from baseline to endline, Mtwara.....	30
Figure 11: Percentage of classrooms in which reading materials were displayed	32
Figure 12: Percentage of classrooms in which half or more than half of students showed learning materials	34
Figure 13: Percentage of classrooms in which reading materials were observed	34
Figure 14: Midline to Endline Gains by Task - Mtwara vs. Lindi	36
Figure 15: Use of reading strategies in Mtwara and Lindi: Changes in frequency of observations from midline to endline	40
Figure 16: Percentage of teachers using reading techniques and methods – Mtwara vs. Lindi	41

Acknowledgements

TZ21 would like to thank the following people for their assistance in this midline evaluation effort:

Creative Associates/USA: Karen Tietjen, Project Advisor, who played a central role in the design of this evaluation, and who provided guidance on instrument development and helpful comments on drafts of this report. Thanks also to Sakil Malik, Project Director, and to Semra Seifu, Program Manager for their administrative support.

TZ21 staff: A special thanks goes to Creative Associates-TZ21 in country staff: Renuka Pillay, Chief of Party; Revocatus Balashuha Grants and IP's Coordinator; Daudi Kwebwa, M&E Advisor; Felix Mbogella, Deputy Chief of Party (Mtwara); Charles Nonga, Deputy Chief of Party (Zanzibar); the TZ21 field teams in Mtwara and Zanzibar. All provided helpful input and ongoing support to the field team for the organization of each activity related to this assessment.

Evaluation field team: Many thanks to Alice Michelazzi, Assessment Manager, for the expert support she provided in the revision of the instruments, organization of the work, and ongoing supervision and quality assurance she provided for instrument development and piloting, data collection, and data entry. Thanks also to the support staff for ensuring the smooth organization of data collection: Joseph Mattogoro, Regional Assessment Coordinator; Nassir Nassour, Regional Assessment Coordinator, Ambakisye Jonas, Data Entry Supervisor. We also would like to thank the enumerators and the data entry operators in Lindi, Mtwara, Zanzibar and Dar Es Salaam for a well-accomplished data collection and data entry exercise.

If the work represented in this report is of outstanding quality, it is due to the efforts of the colleagues mentioned above; if any shortcomings can be found, they are attributable to STS staff, who nevertheless made every effort possible to accurately reflect the contributions of the TZ21 project in this evaluation.

Acronyms

CWPM	Correct Words Per Minute
EGRA	Early Grade Reading Assessment
EGMA	Early Grade Mathematics Assessment
EMIS	Education Management Information Systems
ESRF	Economic and Social Research Foundation
MOEVT	Ministry of Education and Vocational Training
SMC	School Management Committee
TZ21	21st Century Basic Education Program
USAID	United States Agency for International Development

Executive Summary

The 21st Century Basic Education Program (TZ21) is an education project funded by USAID whose aim is to improve children's reading ability in Standards 1 and 2 in Mtwara and Zanzibar, and to strengthen systems that support literacy development.

In order to measure the effectiveness of its interventions, TZ21 conducted a series of evaluations, beginning with a baseline in 2012, a midline at the end of 2013 and an endline conducted at the end of 2014. This report presents the results of the endline while comparing results from all three evaluations.

The endline evaluation was designed by Creative Associates International, Inc., the prime contractor for TZ21 in collaboration its subcontractors, along with TZ21 subcontractor School-to-School International, USAID, the MOEVT/Tanzania, and the MOE/Zanzibar for the midline evaluation. The evaluation was designed to answer two questions:

- What changes can be observed in the teaching and learning of reading in TZ21 project schools in Mtwara and Zanzibar from baseline to endline?
- What differences can be observed in the teaching and learning of reading in TZ21 schools in Mtwara as compared to schools in Lindi, a non-TZ21 region, from baseline to endline?

Lindi, a region adjacent to Mtwara, was included as a counterfactual for TZ21 schools in Mtwara from midline on; Lindi was not included in the baseline. No comparison schools were used for Zanzibar since no adjacent or comparable regions exist to serve as counterfactuals.

Sample

The endline evaluation used the same 3-stage cluster design as the one established for the midline evaluation for Zanzibar and Mtwara. In the first stage, a subset of schools was randomly selected from the random sample used for the TZ21 baseline. In the second stage, Standard 2 (S2) classes were randomly selected from within those schools if there was more than one S2 classroom per school. In the third stage, S2 students were randomly selected from within the selected classes. Comparison schools in Lindi were selected purposively using a matched pair design to link them to comparable schools in Mtwara. In total, 120 schools were included in the endline: 40 in Zanzibar, 40 in Mtwara, and 40 in Lindi. In each school, 10 students were to be assessed and interviews, their teachers were observed and interviewed, the Head Teacher and SMC were interviewed, and an inventory of the school conducted. The data were collected late September/early October 2014 and entered in Tanzania. Analyses and report writing were carried out in October and November in North America.

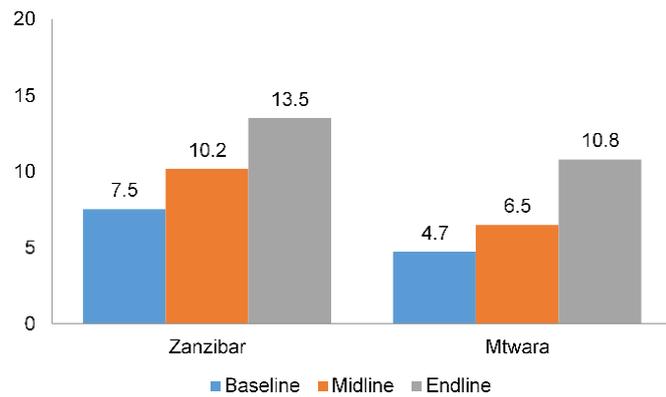
Key findings

EGRA assessment was of high quality: The strength of the EGRA as a measure of student reading ability was extremely high, with strong reliability estimates (Cronbach Alpha), strong item discrimination statistics (item-total correlations), and acceptable difficulty ranges (p-values) – results can be found in Annex B.

Overall EGRA scores

increased: Students’ EGRA scores increased in both Zanzibar and Mtwara over the life of TZ21. Two measures were used to calculate gains. First, the number of Correct Words Per Minute (CWPM) was calculated from baseline to endline. CWPM shows how many words of a short text each student can read in one minute. Students’ CWPM increased from baseline to endline from 7.5 to 13.5 in Zanzibar and from 4.7 to 10.8 in Mtwara. Both gains were statistically significant (see figure right).

Average correct words read per minute, timed reading task, baseline to endline



A second measure was used to capture gains by students on all nine EGRA tasks combined. For this measure, overall EGRA mean scores, or a composite of all nine tasks, was calculated. Again, mean scores increased in both Zanzibar and Mtwara from midline to endline, and the differences were statistically significant.

Comprehension scores increased:

Significant gains were made from midline to endline in Zanzibar on six tasks and in Mtwara on eight (the actual scores are reported in the body of this report) (see table right). Both places saw significant gains in letter name recognition, invented word decoding, and timed and untimed passage reading. Importantly, both places also saw significant gains in timed reading comprehension – a foundational skill for children to be able to “read to learn” in later years.

Gain from midline to endline by EGRA task
X indicates significant difference

Task	Zanzibar	Mtwara
1 Phonemic awareness		x
2* Letter name recognition	x	x
3 Letter sound knowledge	x	
4 Invented word decoding	x	x
5a Passage reading	x	x
5b Passage comprehension	x	x
6a* Passage reading	x	x
6b* Passage comprehension		x
7* Oral passage comprehension		x

Zero scores decreased, though remained high: The proportion of students who received zero scores – scores given when students are not able to answer a single question on a given task – declined substantially from baseline to endline in both Zanzibar and Mtwara. The difference was significant on all tasks in Mtwara. Nevertheless, zero scores remain high: at endline, 49% of students in Zanzibar and 56% of students in Mtwara were unable to read a single word on the timed reading passage.

Girls in Zanzibar performed better than boys, while boys in Mtwara performed better than girls. In Zanzibar, girls performed significantly better than boys on EGRA overall at endline, as was the case at midline. Girls in Zanzibar also saw more significant gains from midline to endline than boys, both on overall EGRA scores and on three tasks: letter name recognition, letter sound knowledge, and timed reading comprehension. Boys in Zanzibar saw significant gains from midline to endline on two tasks: letter name recognition and letter sound knowledge. In

Mtwara, no significant difference was found between boys' and girls' performance at endline, though boys showed significantly greater gains from midline to endline on four EGRA tasks: invented word decoding, timed and untimed reading, and timed reading comprehension. These results suggest that boys in Zanzibar, and boys and girls in Mtwara, need additional help on phonological knowledge (phonemic awareness, letter sound recognition), and that boys in Zanzibar and girls in Mtwara need additional help developing their fluency and comprehension skills.

Mtwara and Lindi: Reading gains were comparable in both regions, but Mtwara created more new readers. Students' EGRA scores in both Mtwara and Lindi improved from midline to endline, and in each region, the differences on seven of the nine tasks were significant. On one task, oral passage comprehension, students in Mtwara made significant gains, but students in Lindi did not. Though gains were comparable in Mtwara and Lindi, Mtwara appears to have created more new readers: from midline to endline, the proportion of students in Mtwara with zero scores decreased significantly on all EGRA tasks, while Lindi the proportion of students in Lindi did not.

Teachers' use of improved instructional strategies increased, especially in Zanzibar. In Zanzibar, teachers' use of strategies such as beginning the lesson with prior knowledge or things taught increased in Zanzibar from 20% at midline to 58% at endline. And in Zanzibar, teachers were also more frequently observed asking students to create or write their own texts (20% midline, 75% endline) and asking students to answer questions about the meaning of text or to draw pictures to show they understand the text (20% midline, 50% at endline) – all key strategies in building students' comprehension skills. The practice of teachers reading out loud remained frequent in Mtwara and increased in Zanzibar. And the frequency of students playing learning games, skits or songs also increased in both Zanzibar and Mtwara. Teachers in both places reduced their practice of criticizing, scolding or punishing their students – in Zanzibar, by 65% (78% at midline, 13% at endline) and in Mtwara, by 22% (30% at midline, 8% at endline). However, some declines in the use of target strategies were also observed – see warning signs below.

Materials were more available and used more frequently: The proportion of children with textbooks or printed materials increased – in Zanzibar, from 57% at midline to 90% at endline and in Mtwara, from 10% at midline to 80% at endline. And more reading materials were found in classrooms at endline in Mtwara, including small books (23% at midline, 50% at endline) and big books (23% at midline, 40% at endline). Displays of student work were more frequently observed in Zanzibar (5% at midline, 38% at endline) and materials provided by TZ21 were observed more frequently in Zanzibar, from 13% at midline to 53% at endline. .

In Mtwara, students with more frequent reading classes had better learning outcomes. At midline, most teachers in Mtwara reported that they were teaching reading every day for 16-30 minutes. By endline, most were teaching less frequently – 3-4 to four times per week - but for more than 30 minutes per day. Children in Mtwara performed significantly better on EGRA at endline, suggesting that an increase in class time leads to improved reading outcomes. The same pattern, however, was not found in Zanzibar.

Significant correlations were found between reading performance and student characteristics. Students with the highest scores on the EGRA in Zanzibar and Mtwara typically had attended a nursery or pre-school, had textbooks in English and Kiswahili in school, had their exercise books checked at home by their parents,

had books or reading materials at home, had a father who can read and write, and had parents who read them stories at home. These findings suggest that the provision of textbooks at school and the establishment of a culture of reading at home with a focus on parental support for reading, are strong predictors of success in learning to read.

Significant correlations were found between reading performance and some teacher, Head Teacher, and school characteristics, but the same correlations rarely occurred in both places. For example, significant correlations were found in Zanzibar, but not in Mtwara, when children with higher EGRA scores had teachers who arranged letter or word tiles and guided students to associate words with letters. Conversely, significant correlations were found in Mtwara, but not in Zanzibar, when children with higher EGRA scores had teachers who read out loud to their students, used a textbooks, and guided students to read printed materials or books.

Recommendations:

1. Continue to provide support to teachers to assist them with the transition from more traditional, whole-class methods to ones that enable learners to work independently in groups, pairs, and individually.
2. Continue to develop students' reading comprehension skills: Assist teachers with the transition from teaching foundational skills to strengthening the reading of connected text and improved comprehension.
3. Continue to provide teachers with regular school-based support focused on early grade reading, including follow-up at school level. Reinforce mechanisms for mentoring & coaching as well as support for literacy acquisition in the home and community. These include WEC observations and support as well as training provided by experts from the MOEVT and education partners.
4. Provide teachers with support in strategies to help struggling students (especially ones with zero scores), including formative and summative assessment strategies, remediation strategies (see next recommendation), individualized instruction, and identification of early warning signs that students are struggling.
5. Provide teachers with explicit support in remediation strategies: Establish a remedial program to inform teachers, Head Teachers, education administrators, and parents of “early warning signs” of children who are struggling to become early grade readers. Help teachers learn to diagnose reading problems, and provide them with tools for addressing these problems, with special adaptations for use in large classes, where a large percentage of children are likely to need support.
6. Ensure the timely distribution and appropriate use of instructional materials. Promote the distribution of textbooks and supplementary materials to Tanzanian primary schools, and systematically monitor their use in the classroom. Establish standards for the use of materials instruction to ensure their regular and effective use.

Introduction

The 21st Century Basic Education Program (TZ21) (hereafter called “TZ21”) is an education project funded by USAID whose aim is to improve children’s reading ability in Standards 1 and 2 in Mtwara and Zanzibar, and to strengthen systems that support literacy development (see box to the right). TZ21 is managed by Creative Associates International and is implemented in collaboration with several partners, including School-to-School International (STS), which is responsible for the design and implementation of the TZ21 baseline, midline, and endline evaluations. This report presents the results of the endline evaluation. It includes data gathered from a sample of schools where TZ21 is implementing its program in Mtwara and Zanzibar, as well as a sample of schools in Lindi serving as comparison schools for project schools in Mtwara.

Purpose of endline

The purpose of the endline evaluation was to measure progress in TZ21 schools over the course of project implementation. TZ21 was initiated in 2010 and a baseline evaluation was conducted in February 2012 under management of the World Bank to assess the status of teaching and learning in project schools at the start of their participation in project activities. For that baseline, four data collection instruments were used:

- An Early Grade Reading Assessment (EGRA) for Standards 1 through 4
- Early Grade Mathematics Assessment (EGMA) for Standards 1 through 4
- A Head Teacher questionnaire, and
- A teacher questionnaire.

Data collection for the midline evaluation was conducted in September-October 2013. To construct the instruments for use in the midline evaluation, two principal sources were used: the instruments used by the World Bank in the baseline, and instruments already developed and in use by Creative in other countries, especially the classroom observation and SMC interview forms.

The endline evaluation uses data from both baseline and midline evaluations to measure change over time within project schools; data are also used to compare changes from midline to endline between project schools in Mtwara and comparison schools in Lindi. Findings are discussed to inform recommendations on future interventions targeted towards the improvement of early reading in Tanzania.

Methodology

Areas of support provided by TZ21

- Reading Program
- Reading Across the Curriculum
- Resource Utilization
- Mentoring & Coaching
- Whole School Development
- Reading Corners/Libraries
- ICT Basic Literacy
- e-Content
- Reading Events
- Research (baseline & midline assessment)
- EMIS

Key activities accomplished in the reading program to date

- Curriculum and text reviews
- Materials development
- Pupils' designed materials
- Consensus and advocacy seminars
- Multiple trainings
- School-based reading corners
- Mentoring & coaching

Research questions

Principal questions: The evaluation was designed to answer two questions:

- What changes can be observed in the teaching and learning of reading in TZ21 project schools in Mtwara and Zanzibar from baseline to endline?
- What differences can be observed in the teaching and learning of reading in TZ21 schools in Mtwara as compared to schools in Lindi, a non-TZ21 region, from baseline to endline?

Though Lindi, a region adjacent to Mtwara, did not participate in the World Bank baseline, the midline and endline evaluations included these region as a point of comparison for TZ21 Mtwara schools. For Zanzibar, it was not possible to take a comparable set of comparison schools since there are no adjacent or comparable regions to Zanzibar.

Supporting questions: For all three places¹, three supporting questions were used to guide the design of the endline:

- To what extent are students, teachers, head teacher, schools and communities implementing practices or activities to support reading?
- Which factors at the student, classroom, and school levels are associated with strong reading performance?
- What can be learned about the success of TZ21 to be applied to other USAID projects or initiatives in Tanzania?

Sample

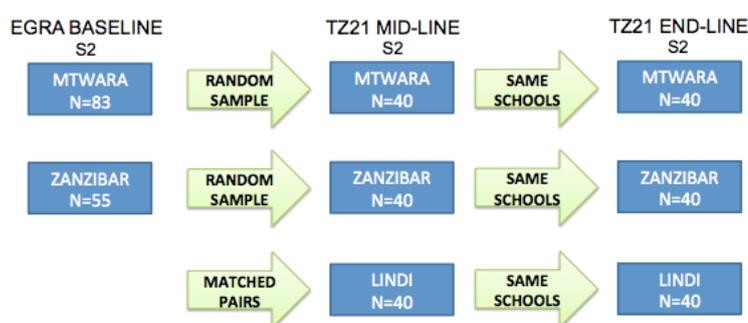
The endline evaluation used the same 3-stage cluster design as the one established for the midline evaluation for Zanzibar and Mtwara. In the first stage, a subset of schools was randomly selected from the random sample used for the TZ21 baseline. In the second stage, Standard 2 (S2) classes were randomly selected from within those schools if there was more than one S2 classroom per school. In the third stage, S2 students were randomly selected from within the selected classes. To select the students, enumerators used the interval method of random selection by counting the number of students on the class list, then dividing by ten.

Comparison schools in Lindi were selected purposively using a propensity scoring matched pair design to link them to comparable schools in Mtwara. Only schools in districts bordering Mtwara were included in the sample.

In total, 120 schools participated in the endline, 40 from each place: Zanzibar, Mtwara, and Lindi. A total of 1,196 out of 1,200 students participated, slightly lower than the targeted number due to one class in Lindi in which only six students were present.

The number of schools and students selected provide the ability to generalize results to

Figure 1: TZ21 midline sampling plan, number of schools



¹ Because Mtwara and Lindi are regions of Tanzania and Zanzibar is its own Republic (and not a region of Tanzania), we use the term “places” in this report instead of “regions.”

all public primary school students in Mtwara and Zanzibar with a 95% confidence interval and up to a 5% margin of error, with the possibility of detecting an effect size of 0.17. This sample size provided a high level of power, with the possibility of detecting small effects, meaning that correlations with a value higher than 0.17 are statistically significant. On the other hand, school results, such as hours of instruction, are generalizable to all public primary schools in Mtwara, and to all public primary schools in Zanzibar, with a 95% confidence interval, a margin of error 15.5%, with the possibility of detecting an effect size of 0.42. School-level analyses are thus able to detect medium effects, meaning that correlations with a value higher than 0.42 are statistically significant.

Because schools in Lindi were selected purposively, results are only applicable to the schools in the evaluation and cannot be generalized to all schools in Lindi, or the district from which they were selected.

The total numbers of participants involved in this evaluation are as follows:

Table 1: Number of participants included in the endline

Participant	Lindi		Mtwara		Zanzibar	
	Targeted	Actual	Targeted	Actual	Targeted	Actual
Head Teachers or other school administrators	40	40	40	40	40	40
S2 Teachers interviewed	40	40	40	40	40	40
Students (EGRA & questionnaire)	400	396	400	400	400	400
Teachers observed	40	40	40	40	40	40
SMC members interviewed	200	192	200	191	200	190
Total	720	708	720	711	720	710

Instrument development

Instruments used at the endline were the same as those used at the midline, with some minor edits, except for the EGRA (see *EGRA development* below). This was done to ensure that the endline EGRA test was of comparable difficulty to the baseline and midline EGRA tests. As with the midline, all instruments were bilingual, in English and Kiswahili, adapted for Mtwara and Zanzibar. (For a full description of instrument development, see *TZ1 Midline Evaluation Report: Evaluation conducted in Zanzibar, Mtwara, and Lindi*.)

The six instruments implemented at endline were as follows:

1. Head Teacher interview form
2. School data form
3. EGRA & student questionnaire, Standard 2
4. Classroom observation form
5. Teacher interview form
6. School Management Committee interview form

EGRA development

Parallel items from the midline EGRA instruments were developed for use on the endline in order to avoid the “test re-test” effect, in which children remember items

from the pretest and, subsequently, score higher on the post-test. For two of the endline EGRA tasks, common items were pulled from the baseline EGRA instrument. These items were not expected to be recalled from their administration two years prior.

Results of the baseline and midline versions of the EGRA were compared and results were found to be comparable, with overall mean scores on the two assessments diverging by less than one percentage point. Although the specification of the midline EGRA and the endline EGRA are the same (i.e. number of tasks and types of tasks), their content is different. Anchor items, or a subset of tasks that were identical between administrations, were used to put the overall EGRA score from the endline EGRA on the same scale of value as the midline EGRA. (Anchor items are highlighted in Table 2 below.) Thus, any difference in difficulty between the two assessments are taken into account, and differences in the overall scores between midline and endline represent real differences in the reading competency of students.²

As with the previous administrations of the EGRA, the endline EGRA consisted of two parts: a literacy assessment and a student context interview. An overview of the tasks comprising the EGRA at each time point is provided in Table 2.

² The equating method used for this procedure is based on the observed scores for anchor items and overall EGRA scores. Differences in the observed scores for anchor items represent real differences in the level of reading competency amongst students. A linear analysis was applied to re-scale overall EGRA scores on the endline EGRA based on the differences observed between anchor items from midline to endline. This produced an “equated” overall EGRA score used for endline analyses. It should be noted that direct comparisons between midline and endline anchor items and the overall EGRA scores presented in this report can be directly compared. However, direct comparisons between non-anchor tasks should be taken with caution because their level of difficulty is not exactly the same between test administrations.

Table 2. EGRA Tasks by Test Administration

Baseline		Midline		Endline	
Task #	Description	Task #	Description	Task #	Description
1	Letter-sound knowledge	1	Phonemic awareness	1	Phonemic awareness
2	Invented word decoding	2*	Letter name recognition	2*	Letter name recognition
3a	Oral passage reading (timed)	3	Letter sound knowledge	3	Letter sound knowledge
3b	Reading comprehension	4	Invented word decoding	4	Invented word decoding
4a	Oral passage reading (untimed)	5a	Passage reading	5a	Passage reading
4b	Reading comprehension	5b	Passage comprehension	5b	Passage comprehension
5	Student context interview	6a*	Passage reading	6a*	Passage reading
		6b*	Passage comprehension	6b*	Passage comprehension
		7*	Oral passage comprehension	7*	Oral passage comprehension

*Anchor items – Identical items on both midline EGRA and endline EGRA

As can be seen, for both the TZ21 midline and endline, tasks from the baseline EGRA were retained, but new content was developed – for example, different words, sounds, and stories. Two tasks were also added: letter name recognition, phonemic awareness, and oral passage comprehension. The EGRA assessments were conducted with students individually over a 15-20 minute period during which the enumerator first put the child at ease with friendly conversation, then led the student through a series of diagnostic tasks. The seven tasks comprising the endline EGRA assessment tested the following reading performance skills:

1. Phonemic awareness, in which students demonstrated their ability to distinguish sounds by listening to three words, two of which started with the same sound and the third with a different sound. The task was to state the word that began with a different sound. This is an oral task.

2. Letter name recognition, in which students were asked to read letters from a sheet and name the letters.

3. Letter sound knowledge, in which students were asked not to give the name of the letter, but the sound it makes.

4. Invented word decoding, in which students were asked to read invented words out loud (e.g., “zabu”) – a test of their ability to decode.

5a. Passage reading (timed), in which students read a story to the enumerator within one minute.³

³ Task 5a was a “timed reading” task. This means that students were asked to read a passage. They then received two scores. The first was the *number of words* read correctly within the minute, the second was *rate* at which the student had read (also called their fluency score). To calculate a fluency score, the student’s reading speed is calculated over a one-minute period. For example, if a student read 30 words in one minute, her rate was 30 words per minute; if she read 30 words in 30 seconds, her rate was 60 words per minute. For task 6a, the

5b. Passage comprehension, in which students were asked literal and inferential questions based on the timed passage. With literal questions, the answer is provided explicitly in the passage; with inferential questions, the answer does not appear explicitly in the text and must be inferred based on information provided in the text.

6a. Passage reading (untimed), in which the student read the passage for up to three minutes.

6b. Passage comprehension, in which the student was asked questions based on the untimed passage.

7. Oral passage comprehension, in which the enumerator read a passage while the student listened without reading. The enumerator then asked several comprehension questions.

(Annex A presents the three passages and comprehension questions.)

These tasks were followed by a student context interview, in which the student was asked a series of questions to identify contextual factors such as home language or literacy levels of parents. These questions were asked in order to identify variables that might be associated with stronger or weaker reading performance.

Administration

Training of Data Collectors

Enumerator training consisted of two 3-day workshops: one in Zanzibar for enumerators recruited there, and one in Mtwara for enumerators from Mtwara and Lindi. Almost all enumerators had served as enumerators in the TZ21 midline. In each workshop, enumerators were re-introduced to the six endline instruments cited above, then practiced using the instruments in neighboring schools. Each enumerator had one opportunity to practice using each instrument.

Data Collection

All data for this evaluation were collected over a four-week period starting during September and October 2014, when data were collected in Zanzibar (September 15-19), Mtwara (September 25-October 6), and Lindi (October 7-17). Data collection in Zanzibar spanned a shorter time period due to schools' double shifts and shorter distances among different schools, allowing enumerators to observe two schools per day at each school site. Teams of three enumerators visited each school for one day to administer all six forms. Enumerators typically worked in pairs to conduct the EGRA assessments while the third enumerator conducted the interviews. Classroom observations were conducted by two enumerators, who discussed their observations after each observations and adjusted their responses in light of these discussions in order to increase reliability. Conflicting observations were eliminated from the analysis. Qualitative data were collected and recorded in Kiswahili, then translated into English for analysis.

“untimed reading” task, students were asked to read a second passage during which they were timed, but this time they were given up to 180 seconds to read the passage.

Each team was assigned a Local Assessment Coordinator (LAC) who visited schools during data collection to assure quality. Each LAC used a supervisor's checklist to ensure protocols were followed, including detailed questions following the implementation of the EGRA assessment. Each enumerator team had a Team Leader who reported in person to his/her LAC each day after data collection was completed. Discussions focused on the number of forms completed each day and issues arising. The LACs then provided their reports on a daily basis to the Assessment Manager.

At the end of each school visit, the LACs collected the forms from each school, placed them in designated envelopes, and packed them in boxes which in turn were sealed, labeled, and secured. When the data collection exercise was complete, the boxes were shipped to the TZ21 office in Dar es Salaam.

One problem encountered in the data collection phase occurred in Lindi where one school had only six students in Standard 2. In this instance, the targeted numbers of students was not reached for the EGRA assessment. In Zanzibar it was necessary to visit one school twice as it was not possible to meet the relevant respondents on the first visit. Enumerator teams received substantial encouragement and support from Regional and District Education Officers as well as school personnel and SMC members.

Data Entry and Cleaning

STS developed the data entry templates, selected and trained the data entry operators, and supervised the data entry exercise. All data were entered in EpiInfo 7 software and stored in a secure folder on a virtual shared drive (Dropbox), to which the STS Assessment Manager exported each data file in Excel and organized a common data base which was then checked by a team of technical experts at STS.

Data Analysis

Once the data entry exercise was completed, STS conducted its own internal data cleaning. Then, descriptive tables were generated for all items in the interview questionnaires and observation forms and shared with Creative for review. Inferential analyses were also run, including correlations and tests of significance of difference (T-tests and chi-square) to identify statistically significant differences between variables. Questions that called for "other" responses were recorded in Kiswahili, then translated into English by the Assessment Manager for analysis. (For a description of how correlations and timed reading passage results were calculated, see Annex B.)

Reporting, presentations, and validation of findings

Three reports were prepared based on this endline: a 10-page summary report for Mtwara, a 10-page report for Zanzibar, and this final report containing findings from Mtwara, Zanzibar, and Lindi. Three PowerPoint presentations were also prepared and presented: one for the Mtwara results, presented in Mtwara and Dar es Salaam in December 2014; one for the Zanzibar results presented in Zanzibar in January 2015; and one for USAID/Tanzania presented to the mission in January 2015.

All of these presentations were used as opportunities to validate the findings presented in this report, and changes were made according to observations made by participants in those presentations. The two biggest changes included the elimination of the analyses pertaining to frequency and duration of reading lessons, due to the confusion on the part of some teachers concerning the notion of a “reading lesson” (see Question 3: Practices & activities supporting reading, Zanzibar) and the addition of the analysis about the differing correlations vis-à-vis reading outcomes between Mtwara and Zanzibar (see Discussion: Differences between Mtwara and Zanzibar at endline).

Findings

This section is organized around each of the key questions presented in the Methodology section above.

Question 1: What changes can be observed in the teaching and learning of reading in TZ21 project schools in Mtwara and Zanzibar from baseline to endline?

As with the midline evaluation, the status of teaching and learning in TZ21 project schools was measured using a combination of EGRA results and information about teaching and learning as captured in Head Teacher interviews, teacher interviews, and observations of reading lessons in Standard 2 classrooms.

EGRA results, midline to endline, Zanzibar

The reading skills of a total of 400 students were assessed in Zanzibar using the endline EGRA assessment. Of these, 207 were girls, 192 were boys, and 2 students did not have a reported gender. Students’ reported ages were just slightly higher than those reported at the midline assessment, ranging from six to 13 years old, with the majority reporting 9 years of age (see Figure 2):

At endline, the overall mean EGRA score in Zanzibar was 29.7% – a composite score averaging the scores of all tasks on the test. This was a statistically significant increase of about 5

percentage points over the overall EGRA score observed at midline ($p < 0.001$). The following histograms illustrate the change from midline to endline, with the trend curve moving slightly to the right, a positive direction:

Figure 2: Student’s ages, Standard 2, Zanzibar

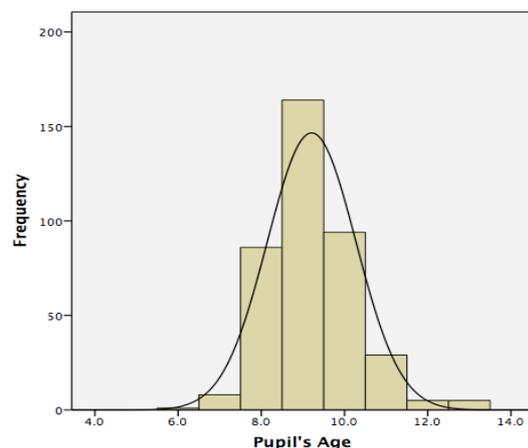
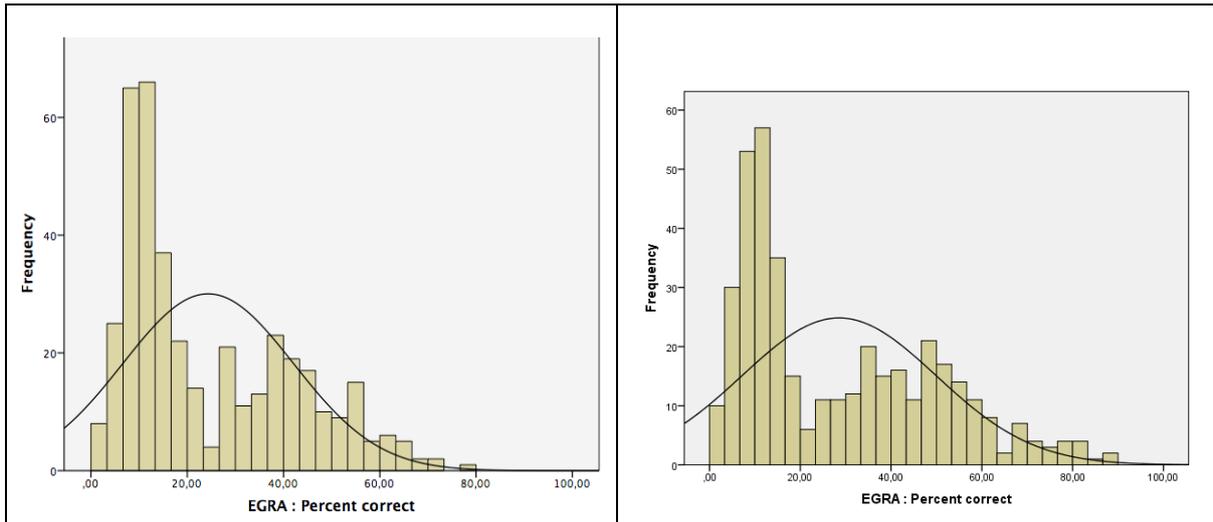


Figure 3: Comparison of overall EGRA scores, midline to endline, Zanzibar

Distribution of EGRA scores, midline	Distribution of EGRA scores, endline
--------------------------------------	--------------------------------------



As Table 4 and Figure 3 show, as they did in the midline, students again scored highest at endline on the oral passage comprehension task and second highest on the untimed reading task. Students scored lowest on the timed comprehension task, invented word decoding, and letter sound recognition. Gains from midline to endline were observed to be significant for the letter name recognition, letter sound recognition, invented word decoding, timed passage reading, timed passage comprehension, and untimed passage reading tasks. Other observed gains or decreases were not found to be statistically significant.

Table 3: Task statistics: Percent correct by task, Zanzibar

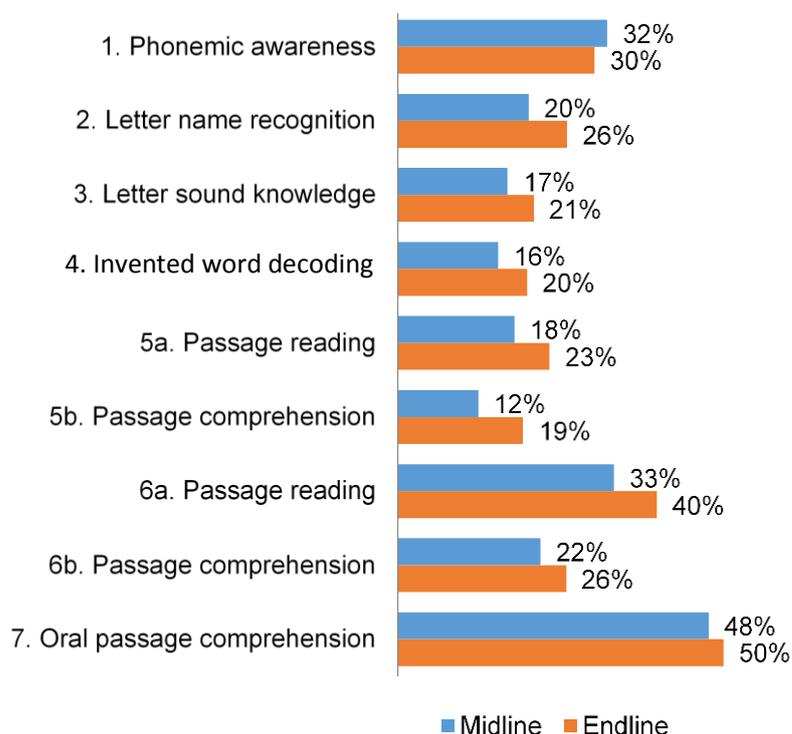
	Midline	Endline	Gain
1. Phonemic awareness	32.4%	30.4%	-2.0%
2. Letter name recognition	20.2%	26.2%	5.9%***
3. Letter sound knowledge	17.0%	21.0%	4.1%***
4. Invented word decoding	15.5%	20.0%	4.5%*
5a. Passage reading (timed)	18.0%	23.5%	5.4%**
5b. Passage comprehension (timed)	12.5%	19.4%	6.9%***
6a. Passage reading (untimed)	33.4%	40.1%	6.6%*
6b. Passage comprehension (untimed)	22.1%	26.1%	4.0%
7. Oral passage comprehension	48.1%	50.4%	2.3%
Overall EGRA Score	24.4%	29.7%	5.4%***

*Significant at $p < 0.05$

**Significant at $p < 0.01$

***Significant at $p < 0.001$

Figure 1: Midline and endline EGRA mean scores by task



As seen at the midline assessment, girls performed better than boys in terms of their overall reading score. The difference of about 6 percentage points at endline was observed to be statistically significant ($p < 0.01$).

Table 4: reading summary score by gender

Groups	N	Midline		N	Endline	
		Mean (%)	SD		Mean (%)	SD
Boys	192	22.0	17.12	181	26.1	22.37
Girls	207	26.4*	17.94	213	32.5**	24.39

*Significant at $p < 0.05$

**Significant at $p < 0.01$

The following table shows that by the time of the endline assessment, on average, girls outperformed boys on the majority of tasks, excluding phonemic awareness, the untimed reading comprehension, and the oral passage comprehension tasks (where mean scores were found to be the same between boys and girls). Although significant differences observed between boys and girls appeared to shift across tasks between midline and endline assessments, girls outperformed boys on all tasks except the oral passage comprehension task, where boys outperformed girls at midline, but where no difference between boys' and girls' performance was observed at endline.

Significant gains in the percentage correct achieved by girls from midline to endline were observed with respect to letter name recognition ($p < 0.001$), letter sound knowledge ($p < 0.05$), the timed reading comprehension tasks ($p < 0.05$). Gains in the

overall EGRA score for girls were also found to be significant ($p < 0.01$). For boys, significant gains were observed with respect to the letter name recognition ($p < 0.01$) and letter sound knowledge ($p < 0.01$) tasks. Although girls outperformed boys across several tasks at endline (as shown in Table 6), tests of differences in gains between girls and boys were not found to be significant, suggesting that the strides made by both boys and girls from midline to endline were equivalent. Together, these findings suggest girls' overall reading ability continues to be higher than that of boys' overall reading ability at the time of endline despite gains made between midline and endline assessments.

Table 5: Percent correct scores by task and gender, highlighting differences in achievement by gender

	<u>Midline</u>		<u>Endline</u>		<u>Gain, midline to endline</u>	
	Girls	Boys	Girls	Boys	Girls	Boys
1. Phonemic awareness	32.6%	32.2%	30.6%	30.2%	-2.0%	-2.0%
2. Letter name recognition	22.7%	17.5%	29.2%***	22.4%	6.6%	4.9%
3. Letter sound knowledge	19.8%	13.9%	23.1%**	18.3%	3.4%	4.4%
4. Invented word decoding	17.2%*	13.5%	23.3%**	15.6%	6.1%	2.1%
5a. Passage reading (timed)	20.6%*	15.0%	27.8%***	17.9%	7.2%	2.9%
5b. Passage comprehension (timed)	13.9%	10.8%	23.1%***	14.6%	9.2%	3.8%
6a. Passage reading (untimed)	39.6%***	26.5%	45.3%**	33.4%	5.7%	6.9%
6b. Passage comprehension (untimed)	25.8%**	17.7%	28.6%	22.5%	2.8%	4.8%
7. Oral passage comprehension	45.7%	50.4%*	48.1%	52.7%	2.4%	2.3%
Overall EGRA Score	26.4%	22.0%	32.5%	26.1%	6.0%	4.2%

*Significant difference between girls and boys at $p < 0.05$

**Significant between girls and boys at $p < 0.01$

***Significant between girls and boys at $p < 0.001$

Students' fluency can also be understood according to performance standards. That is, if we divide the total student population into groups by performance level, we can track the progress of each group over time. The following table presents students in four groups: non-readers (see "zero scores" below), and readers who can read 1-16 CWPM (early readers), 17-44 CWPM (emergent readers), and 45 or more words per minute (fluent readers). These categories were established by another USAID-funded reading project in Kenya, where EGRA was also administered in Kiswahili, and so are used here.

As can be seen, the greatest proportion of students in Zanzibar fall into the "non-reader" category at both midline and endline, indicating that nearly half of all readers participating in the endline EGRA could not correctly read a single word on the timed reading passage at midline or endline. However, those who could read at midline improved at endline: over one-quarter of students (29%) fall into the "emergent reader" performance category at the time of the endline assessment—a slight improvement in the percentage of students meeting these criteria at midline. The percentage of fluent readers appears to have remained fairly steady from midline to endline.

Table 6: Number and percent of students by reading performance level, Zanzibar

Category (Performance level)	CWPM	Midline		Endline	
		#	%	#	%
Non-Reader	0	191	48%	196	49%
Early reader	1 to 16	93	23%	64	16%
Emergent reader	17 to 44	90	23%	117	29%
Fluent reader	45 and above	26	7%	23	6%

Tables 8 and 9 show the distribution of students across both reading performance levels and the average percentage correct on the timed reading comprehension task at midline and endline, respectively. These findings suggest that, on average, as a student's performance level increased, so too did their comprehension score. As with any data, there are exceptions to this average trend, such that some emergent, and even fluent readers, were seen to score very low on comprehension. In these instances, even though students may be able to read quickly, they do not necessarily understand the content that they are reading.

In general, however, increases in students' comprehension scores across the early, emergent and fluent performance levels are observed between midline and endline assessments. This seems particularly true for fluent readers, most (96%) of whom obtained 80% correct or higher on the comprehension task at endline. Non-readers, of course, could not correctly answer any of the comprehension questions because they were only asked questions if they were able to read at least some of the text.

Table 7: Percent of students by timed reading performance level and comprehension level- midline

Category (Performance level)	CWPM	# of students	% of Students by Comprehension Level						Total
			0%	20%	40%	60%	80%	100%	
Non-Reader	0	191	100%	0%	0%	0%	0%	0%	100%
Early reader	1 to 16	93	26%	72%	2%	0%	0%	0%	100%
Emergent reader	17 to 44	90	3%	59%	24%	10%	3%	0%	100%
Fluent reader	45 and above	26	31%	15%	31%	23%	0%	0%	100%

Table 8: Percent of students by reading performance level and comprehension level- endline

Category (Performance level)	CWPM	# of students	% of Students by Comprehension Level						Total
			0%	20%	40%	60%	80%	100%	
Non-Reader	0	196	100%	0%	0%	0%	0%	0%	100%
Early reader	1 to 16	64	23%	58%	19%	0%	0%	0%	100%
Emergent reader	17 to 44	117	6%	25%	43%	21%	3%	2%	100%
Fluent reader	45 and above	23	0%	0%	4%	0%	70%	26%	100%

EGRA results, midline to endline, Mtwara

A total of 400 students participated in the EGRA assessment in Mtwara. This consisted of 181 boys, 214 girls, and 5 students whose gender was not recorded. Of those students reporting their age (n=348), students ranged from 5 to 12 years old, with the majority reporting 9 years of age (see Figure 5).

Students' overall mean score on the EGRA assessment in Mtwara at the time of the endline assessment was 25.1% – a composite score for all tasks on the test. This was a statistically significant increase of nearly 7 percentage points since the midline assessment ($p < 0.001$). The following histograms illustrate the change from midline to endline, with the trend curve moving slightly to the right, a positive direction:

Figure 4. Student's age, Mtwara

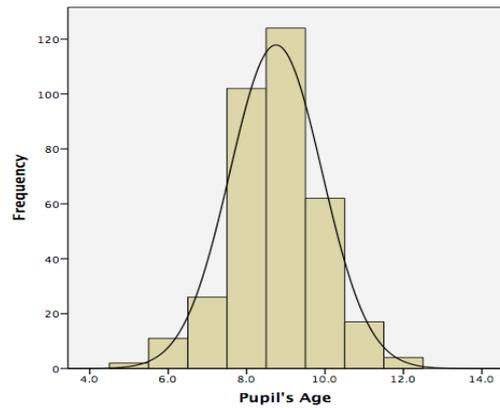
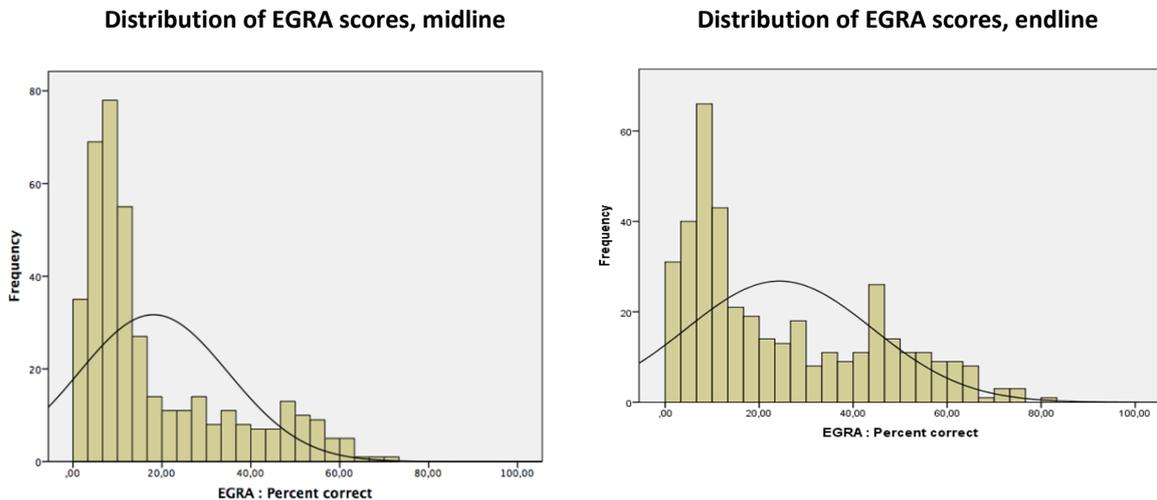


Figure 5: Comparison of overall EGRA scores, midline to endline, Mtwara



As with the midline assessment, students scored highest on the oral passage comprehension. Significant gains were observed across all tasks except for letter sound knowledge, and students scored lowest again on the same two tasks: timed passage comprehension and invented word decoding tasks.

Table 9: Task statistics: Percent correct by task, Mtwara

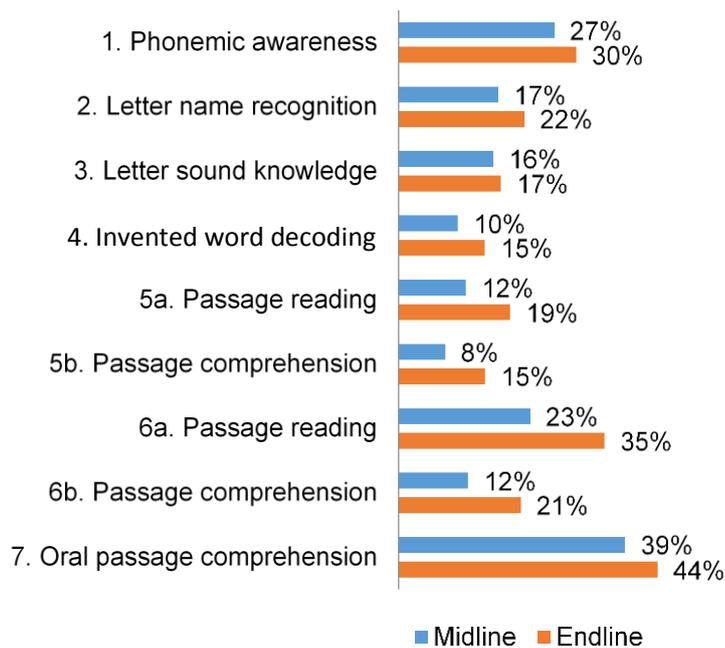
Task	Midline	Endline	Gain
1. Phonemic awareness	26.7%	30.5%	3.7%*
2. Letter name recognition	17.1%	21.6%	4.5%***
3. Letter sound knowledge	16.2%	17.5%	1.3%
4. Invented word decoding	10.1%	14.8%	4.6%**
5a. Passage reading (timed)	11.5%	19.1%	7.6%***
5b. Passage comprehension (timed)	8.0%	14.8%	6.8%***
6a. Passage reading (untimed)	22.6%	35.3%	12.7%***
6b. Passage comprehension (untimed)	11.9%	20.9%	9.1%***
7. Oral passage comprehension	38.8%	44.4%	5.6%**
Overall EGRA Score	18.1%	25.1%	6.9%***

*Significant difference between midline and endline at $p < 0.05$

**Significant difference between midline and endline at $p < 0.01$

***Significant difference between midline and endline at $p < 0.001$

Figure 6: EGRA mean scores by task, Mtwara



Boys achieved a higher overall EGRA score than girls at the endline assessment. However, as with the midline analysis, the difference between girls' and boys' scores at endline was not statistically significant.

Table 10: EGRA summary score by Gender, Mtwara

Groups	N	Midline		N	Endline	
		Mean	SD		Mean	SD
Boys	181	17.23	15.96	179	27.55	23.31
Girls	214	19.07	17.56	203	23.16	20.78

The following table shows that boys made significant strides from midline to endline, evidencing significant gains in four tasks: invented word decoding, timed and untimed reading, and timed reading comprehension.

Although not highlighted in Table 12, it is of note that girls did show significant gains from midline to endline in letter name recognition ($p < 0.05$), and the timed ($p < 0.05$) and untimed reading comprehension ($p < 0.01$) tasks. Although the differences between boys and girls were not observed in terms of their endline overall EGRA scores, both boys and girls were observed to make significant gains in their overall EGRA scores from midline to endline, and these gains were greater for boys than for girls.

Table 11: Percent correct scores by task and gender, highlighting differences by gender, Mtwara

	Midline		Endline		Gain	
	Girls	Boys	Girls	Boys	Girls	Boys
1. Phonemic awareness	26.4%	27.0%	28.4%	32.0%	2.1%	4.9%
2. Letter name recognition	18.2%	15.9%	21.3%	21.9%	3.2%	5.9%
3. Letter sound knowledge	17.5%	14.9%	17.4%	17.5%	-0.1%	2.6%
4. Invented word decoding	11.7%	8.6%	12.5%	17.4%*	0.8%	8.8%**
5a. Passage reading (timed)	13.2%*	9.8%	16.7%	22.2%*	3.4%	12.4%*
5b. Passage comprehension (timed)	9.0%	7.1%	12.7%	17.4%*	3.7%	10.4%*
6a. Passage reading (untimed)	25.2%***	20.2%	32.4%	39.3%	7.2%	19.1%*
6b. Passage comprehension (untimed)	13.8%*	9.9%	21.2%	21.6%	7.4%	11.7%
7. Oral passage comprehension	36.8%	41.5%*	40.7%	49.7%***	3.9%	8.2%
Overall EGRA Score	19.1%	17.2%	23.2%	27.5%	4.1%	10.3%*

*Significant difference between girls and boys at $p < 0.05$

**Significant difference between girls and boys at $p < 0.01$

***Significant difference between girls and boys at $p < 0.001$

Figure 7 and Table 12 present the number and percentage of students distributed across four reader performance levels at both midline and endline for the timed reading task. While the majority of students fall within the category of “non-reader” (zero words read), this number declined from midline (67%) to endline (56%), with these students moving into the early reader and emergent reader categories. The proportion of fluent readers stayed the same.

Figure 7: Number and percent of students by reading performance level (CWPM), Mtwara

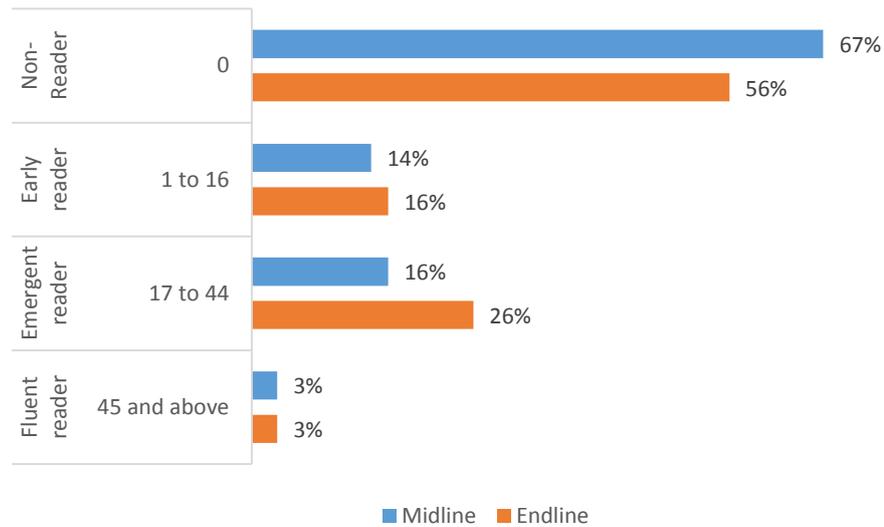


Table 12: Number and percent of students by reading performance level (CWPM), Mtwara

Category (Performance level)	CWPM	Midline		Endline	
		#	%	#	%
Non-Reader	0	266	67%	222	56%
Early reader	1 to 16	57	14%	62	16%
Emergent reader	17 to 44	64	16%	103	26%
Fluent reader	45 and above	13	3%	13	3%

Tables 13 and 14 present how students in each performance category, as measured by CWPM, performed on the comprehension test. These figures show the positive correlation between fluency and comprehension which has been documented elsewhere: as a student’s reading level (CWPM) increases, so does their ability to comprehend the text they are reading.

In Mtwara, the proportion of emergent and fluent readers who scored least 80% comprehension increased from midline to endline. For emergent readers, the proportion grew from 2% to 4%, but for the fluent readers, it grew from zero at midline to 61% at endline. However, it is important to remember that though these are positive changes, they represent a small proportion of the total number of students.

Table 13: Percent of students by reading performance level and comprehension level - midline, Mtwara

Category (Performance level)	CWPM	# of students	%	% of Students by Comprehension Level						Total
				0%	20%	40%	60%	80%	100%	
Non-Reader	0	266	67%	100%	0%	0%	0%	0%	0%	100%
Early reader	1 to 16	57	14%	37%	60%	4%	0%	0%	0%	100%
Emergent reader	17 to 44	64	16%	9%	42%	30%	17%	2%	0%	100%
Fluent reader	45 and above	13	3%	23%	23%	31%	23%	0%	0%	100%

Table 14: Percent of students by reading performance level and comprehension level - endline, Mtwara

Category (Performance level)	CWPM	# of students	%	% of Students by Comprehension Level						Total
				0%	20%	40%	60%	80%	100%	
Non-Reader	0	222	56%	100%	0%	0%	0%	0%	0%	100%
Early reader	1 to 16	62	16%	26%	73%	2%	0%	0%	0%	100%
Emergent reader	17 to 44	103	26%	10%	23%	37%	26%	3%	1%	100%
Fluent reader	45 and above	13	3%	0%	0%	8%	31%	46%	15%	100%

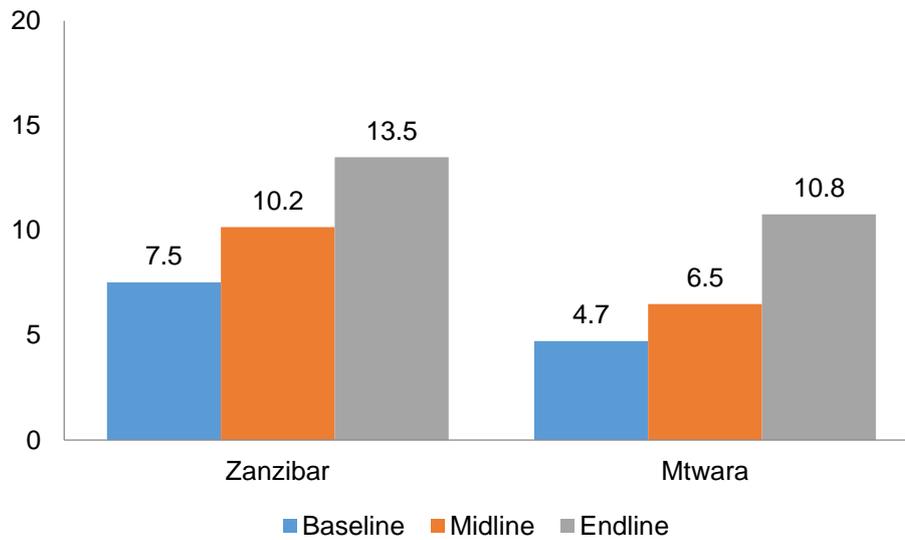
EGRA results, baseline to endline

An analysis of key EGRA variables were undertaken to examine potential changes in reading ability from baseline to endline. Three tasks were common across all three instruments: timed and untimed reading and comprehension, letter sound knowledge, and invented word decoding. The following analyses focus on these tasks.

Averages for fluency scores (CWPM) were calculated for the 60-second timed reading passage. Across time points, means for both Zanzibar and Mtwara showed statistically significant increases from baseline to midline, and again from midline to endline. In Zanzibar, average fluency rates substantially increased from baseline to endline: students read at 7.5 CWPM at baseline, 10.2 CWPM at midline, and 13.5 CWPM at endline, a 6 CWPM gain from baseline ($p < 0.001$).

In Mtwara, average fluency rates rose significantly from 4.7 CWPM at baseline to 6.5 CWPM at midline ($p < 0.001$), and doubled from baseline to endline, where students were reading an average of 10.8 CWPM ($p < 0.001$).

Figure 8: Average correct words read per minute, timed reading task, baseline to endline

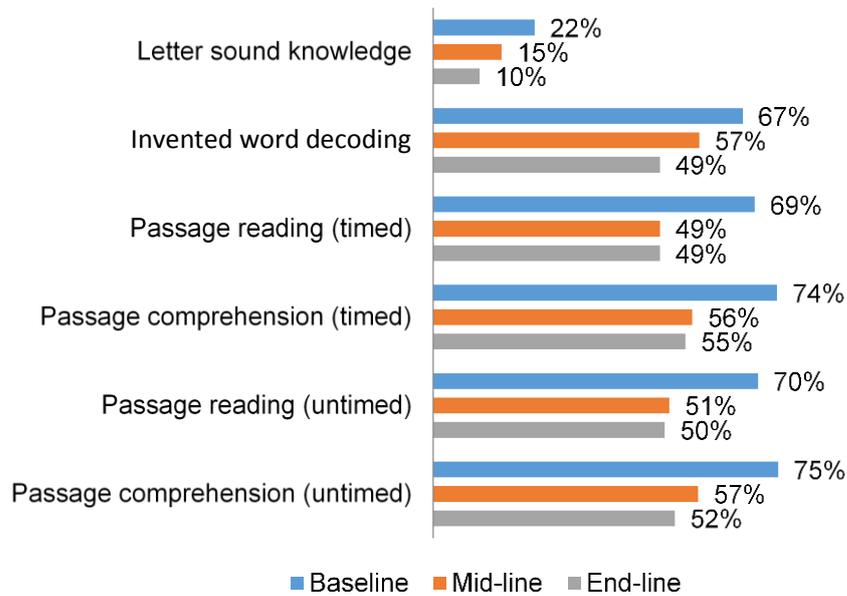


Students receiving “zero scores” on any given task were unable to answer any of the items correctly, or were unable to correctly any of the words in the passages provided. As an indicator of students who are most struggling, Figures 8 and 9 present the proportion of students receiving zero scores from baseline to endline in each Zanzibar and Mtwara.

Substantial decreases in zero scores were observed in Zanzibar from baseline to endline across all tasks analyzed. Decreases in zero scores were greatest with respect to the untimed reading comprehension task, where 75% of students tested received zero scores at baseline and 52% received zero scores at endline--a difference of 23 percentage points. Other notable decreases in zero scores from baseline to endline were observed with respect to the untimed reading passage (20 percentage points), the timed reading passage (20 percentage points), and the timed reading comprehension task (19 percentage points). The lowest zero scores were observed in letter sound knowledge, which fell from about one-quarter of students at baseline to 10% of students at endline. In this case, reductions in zero scores were significant from baseline to midline, as well as from midline to endline.

While these reductions are important, about half of students are still unable to answer a single question or read any words for most of the tasks. This indicates that mastery of the most basic reading skills is still beyond the reach of approximately half the children in Zanzibar.

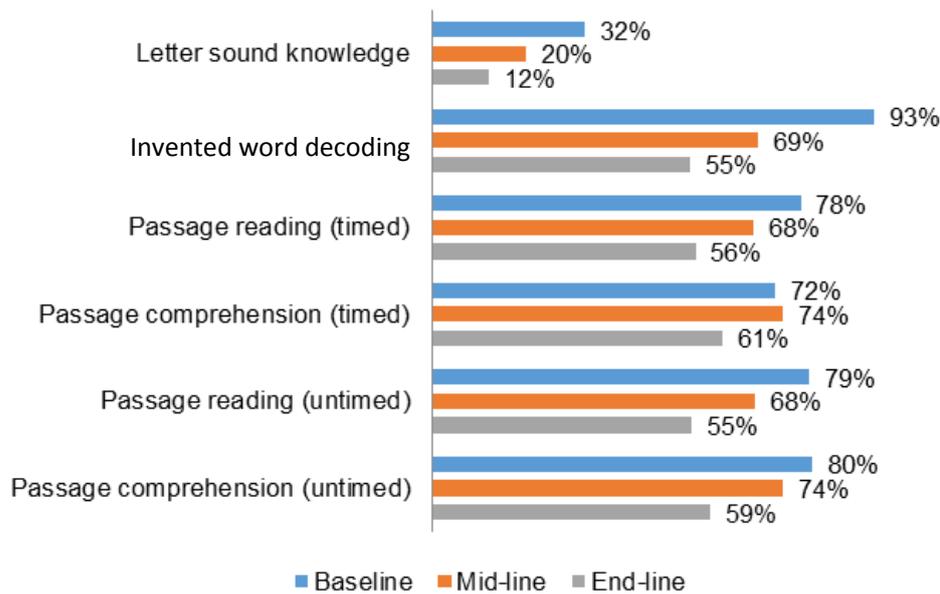
Figure 9: Percentage of zero scores from baseline to endline, Zanzibar



An analysis of zero scores in Mtwara reveal statistically significant reductions in the proportion of struggling students from baseline to endline across all comparable tasks. The biggest reduction of zero scores was on the invented word decoding task, where 93% of students obtained zero scores at baseline and only 55% of students obtained zero scores at endline—a decrease of 38 percentage points. Other big decreases in zero scores were observed in the letter sound knowledge, timed reading passage, untimed reading passage, and untimed reading comprehension tasks.

As with Zanzibar, however, while zero scores decreased from baseline to endline, over half of students still received zero scores on half or more of all tasks assessed at baseline, midline, and endline.

Figure 10: Percentage of zero scores from baseline to endline, Mtwara



Reductions in the proportion of zero scores were statistically significant across all comparable tasks from baseline to midline in both Zanzibar and Mtwara. However, from midline to endline, reductions were only statistically significant in Zanzibar for the letter sound knowledge and invented word decoding passages, and for the letter sound knowledge, invented word decoding and timed reading passage in Mtwara. In both Zanzibar and Mtwara, one of the greatest challenges will be continuing to reduce the proportion of students obtaining zero scores. This becomes particularly difficult as the proportion of zero scores remaining tends to represent students who struggle the most with reading. This is often the case with respect to improvements in learning for students who perform persistently low, as these students will likely need more specific and intensive instruction than their peers more receptive to general improvements in the teaching of reading.

Teaching and learning in TZ21 schools: Zanzibar

This section describes the status of teaching and learning in TZ21 schools in Zanzibar by examining seven aspects of teaching and learning as captured in the observation instrument:

- Classroom environment
- Teaching and learning methods
- Teachers’ use of reading instruction methods
- Teacher assessment of student learning during the lesson
- The use of learning and instructional materials
- Students’ reading behavior, and
- Teacher-student interactions.

In each school Zanzibar, midline enumerators observed a Standard 2 reading class in pairs. Following midline protocol, raters were asked to discuss their findings and reach consensus on each observation item. Inconsistencies between raters found in the data were, therefore, discarded and marked as “missing.”

Classroom environment: As at the midline, conditions observed in classrooms were generally viewed as conducive to instruction, with 95% of observers reporting classrooms were clean and tidy and 98% reporting having sufficient air and light in the classroom. At endline, all students had seats in only 19 of 40 – or 48% – of classes observed, a decrease from 24 classrooms at midline. The observance of girls and boys being equally likely to have (or not have) a seat was similar at the endline as compared to the midline (55% at midline vs. 53% at endline). More classrooms at endline showed boys and girls equally distributed in the front and back of the classroom (30 cases – or 75% – at endline compared to 24 cases – or 60% – during the midline). Similar to the midline, in most cases (n=38, or 95%), students were seated in rows.

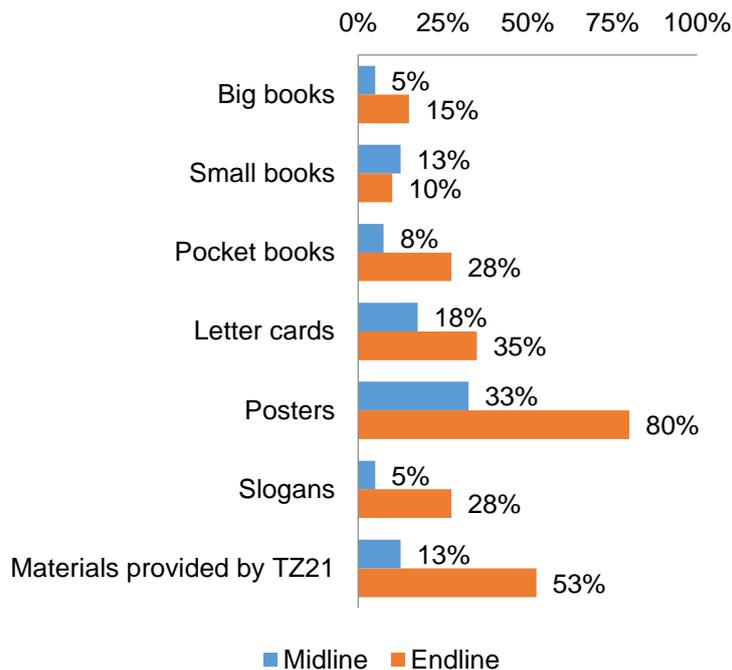
Teacher's general methodological approach: In all classrooms observed, some or all of the reading lesson was conducted in Kiswahili. Teachers were more frequently observed beginning the lesson with prior knowledge or things taught (20% at midline to 58% at endline), and substantially reduced the practice of criticizing, scolding or punishing students (78% at midline to 13% at endline).

Teacher's use of reading instruction methods: Importantly, changes in the use of reading-specific teaching methods were observed from midline to endline, including teachers asking students to create or write their own texts (20% midline, 75% endline) and asking students to answer questions about the meaning of text or to draw pictures to show they understand the text (20% midline, 50 endline). Together, these findings suggest that teachers in Zanzibar are using these techniques more often at midline, with a particular focus on students' understanding of text.

Basic study materials: Substantial improvements in classroom reading materials were observed from midline to endline: while only one-quarter of classrooms observed at midline had students with textbooks or printed materials, at endline, 100% of observed classrooms had students with a textbook or printed materials. There was, however, a reduction in classrooms observed where children had an exercise book, paper or slate (78% at midline to 58% at endline). Similar to midline, almost all classrooms had a blackboard and chalk or marker (93% at midline to 98% at endline).

Learning materials: Observers noted the types of literacy materials available in classrooms in which they were observing. There was a large increase in the display of posters (33% at midline to 80% at endline), materials provided by TZ21 (13% at midline to 53% at endline), and slogans (5% at midline to 28% at endline). Student work was also increasingly displayed in classes observed (15% at midline to 38% at endline). There was an improvement in the availability of materials, with the amount of classrooms in which no children had textbooks or printed materials reduced from 17 – or 43% – at midline to only 4 – or 10% – at endline.

Figure 11: Percentage of classrooms in which reading materials were displayed



Teacher’s use of instructional materials: In general, observers did not find a substantial increase in teachers’ use of instructional materials. The proportion of teachers who had a teacher’s guides applicable to the class was similar at endline as compared to midline (53% at midline to 55% at endline). From midline to endline, there was a reduction in the percentage of teachers who could evidence their non-TZ21 lesson plans (65% at midline to 43% at endline). In addition, less than one-third of teachers were able to show at midline (28%) and endline (30%) that they had a scripted lesson from TZ21.

As with midline observations, the most frequently observed instructional material used by the teacher at endline was the blackboard (98% at midline to 90% at endline). The observed use of textbooks in classrooms climbed from 30% at midline to 58% at endline. The use of student exercise books and/or slates also increased from midline (10%) to endline (20%) while the use of small books, computers from Bridge IT provided by TZ21, and other ICT equipment decreased. No manipulatives were observed at midline or endline. The use of computers from mobile labs provided by TZ21 remained steady from midline to endline at 10% of observations. These findings suggests that TZ21-provided project materials were not increasingly adopted into practice after the midline.

Students’ reading behavior: Observers were asked to note students’ reading behavior. The behavior most frequently observed at both midline and endline was repetition and recitation (90% at midline to 95% at endline). Slight increases in a variety of alternative reading practices were observed from midline to endline including, reading out loud to another student (0% at midline to 3% at endline), reading out loud one student at a time (78% from midline to 83% at endline), and students writing on paper, in an exercise book or on a slate (23% at midline to 25% at endline). Larger increases were seen in the behaviors of students listening to the teacher read out loud (73% at midline to 90% at endline), answering teacher’s questions (78% at midline to 90% at endline), and playing learning games, skits or songs organized by the teacher (33% at midline to 43% at endline). There was a

decrease in observations of students reading independently from 13% at midline to 3% at endline. Together, these findings suggest that teachers could be making use of more interactive learning methods to teach reading.

Teaching and learning in TZ21 schools: Mtwara

As was presented in the preceding section, teaching and learning in TZ21 schools in Mtwara will be examined reviewing six aspects of teaching and learning as captured in the observation instrument:

- Classroom environment
- Teaching and learning methods
- Teachers' use of reading instruction methods
- Teacher assessment of student learning during the lesson
- The use of learning and instructional materials, and
- Students' reading behavior.

In each school in Mtwara, endline enumerators observed a Standard 2 reading class in pairs and were asked to reach consensus on each observation item. Where discrepancies were found between observers in the data, therefore, they were marked as "missing."

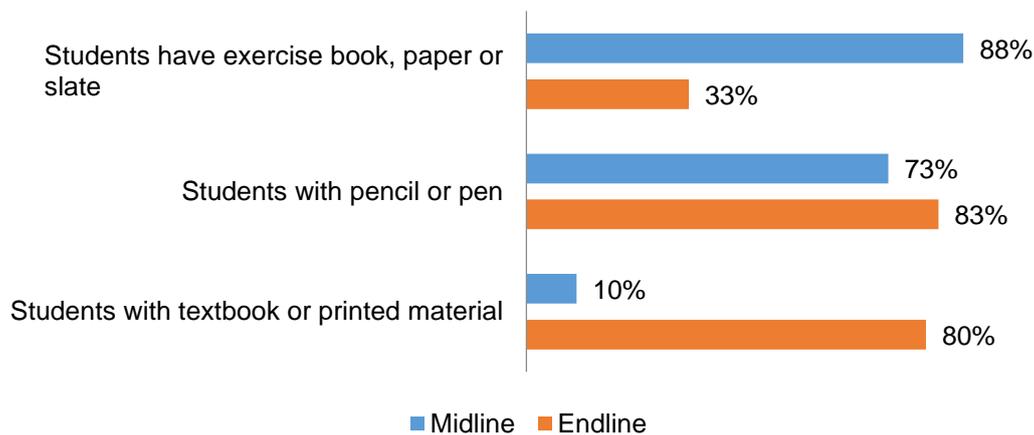
Classroom environment: Conditions observed in classrooms were, in general, improved since the midline, with more observers reporting classrooms were clean and tidy (58% at midline to 93% at endline) and had sufficient air and light in the room (98% at midline to 100% at endline). All classrooms observed at endline had seats for all students, up from 88% at midline. Students were seated in rows.

General teaching methods: The largest increases in observed teaching methods were teachers demonstrating reading or writing skills (85% at midline to 98% at endline), teachers praising or complimenting students (83% at midline to 95% at endline), and asking students questions about the lesson (93% at midline to 100% at endline). For reasons that are unclear, observers noted a decrease in teachers providing an explanation if student don't understand (85% at midline to 43% at endline), though there was a very slight increase in teachers responding to student questions (0% at midline to 3% at endline). There was also a substantial decrease in teachers criticizing, scolding, or punishing students (30% at midline to 8% at endline). These results suggest that teachers began to increasingly implement the positive teaching methods that were already prominent at midline.

Teacher's use of reading instruction methods: In Mtwara, observers noted a decrease in several reading instruction methods including teachers' identification of the differences and similarities of sounds (83% at midline to 60% at endline), blending letter-sounds to form syllables and words (80% at midline to 58% at endline), dictating to students words or sentences (35% at midline to 18% at endline), and discussing the meaning of vocabulary words (40% at midline to 30% at endline). There were slight increases observed in the proportion of teachers asking students to think critically about text however, such as in asking students to create or write their own texts (28% at midline to 35% at endline) and answering questions about the meaning of texts or drawing pictures to show the students understand the text (12% at midline to 15% at endline). A slightly greater proportion of classrooms also featured students reading printed materials or books (10% at midline to 15% at endline).

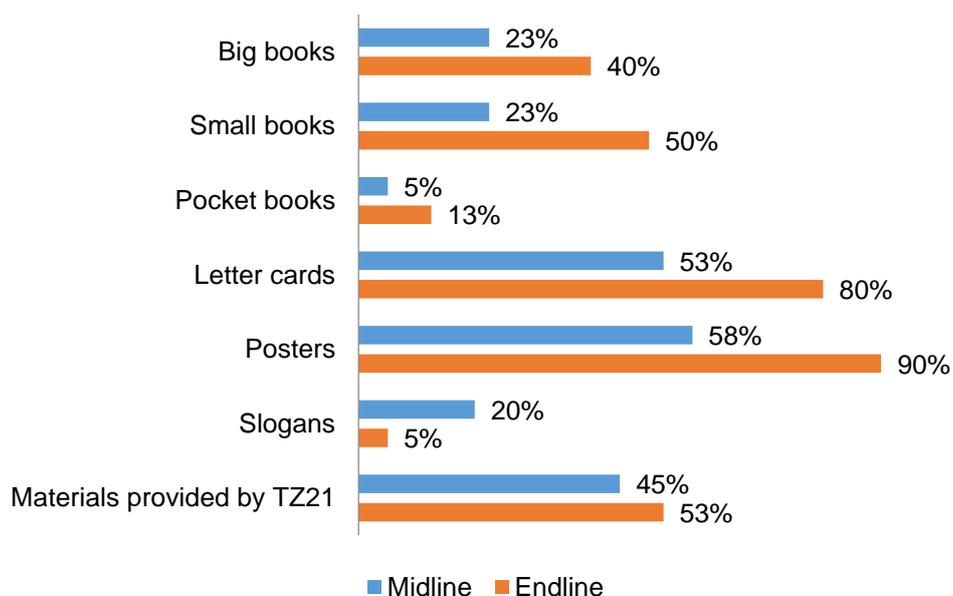
Basic study materials: More students were observed equipped with pencils or pens at endline (83%) than midline (73%). Importantly, while only 10% of classroom featured students with textbooks or printed material at midline, by endline, 80% of classrooms featured these materials. While it is unclear as to whether an increase in textbooks or other printed materials may have affected the use of other materials, a reduced proportion of classrooms were observed where students had an exercise book, paper or slate (88% at midline to 33% at endline).

Figure 12: Percentage of classrooms in which half or more than half of students showed learning materials



Learning materials: Learning materials were increasingly available and on display at endline as compared to midline in all cases excluding slogans (which decreased from 20% at midline to 5% at endline). The largest increases in observed materials were with respect to posters (58% at midline vs. 90% at endline), small books (23% at midline vs. 50% at endline), and big books (23% at midline vs. 40% at endline).

Figure 13: Percentage of classrooms in which reading materials were observed



Teacher’s use of instructional materials: Teachers increasingly utilized a variety of instructional methods at endline including small books (3% at midline to 10% at endline), worksheets (0% at midline to 8% at endline), and manipulatives (5% at

midline to 10% at endline). Smaller increases were seen in the use of blackboards (85% at midline to 90% at endline), big books (5% at midline to 8% at endline), and computers from mobile labs provided by TZ21 (3% at midline to 5% at endline). In general, the most frequently observed instructional materials at midline continued to be the most frequently observed at endline. Most notable was the increased use of computers from Bridge IT provided by TZ21 which rose from 13% of classrooms at midline to 28% at endline.

Question 2: What differences can be observed in the teaching and learning of reading in TZ21 schools in Mtwara as compared to schools in Lindi, a non-TZ21 region, from baseline to endline?

TZ21 continues to implement programming in Mtwara and Zanzibar, and the focus of the endline evaluation remains in these two places. Schools in the Lindi region were again selected to provide a point of comparison to TZ21 interventions in Mtwara. Performance in Lindi schools thus serve as a counterfactual from which to determine whether TZ21 interventions are uniquely influencing teaching and learning in Mtwara schools.

EGRA results

A review of performance by task on the midline and endline EGRA assessments shows that Lindi students score significantly higher than Mtwara students in most tasks at both time points. However, the important difference is in gains made in each region from midline to endline in order to understand the extent to which TZ21 interventions may have had an influence on student performance.

The results in Table 19 show that students in both Mtwara and Lindi made significant gains from midline to endline with most EGRA tasks as well as their overall EGRA scores. Reading gains in both provinces were possibly influenced by the Big Results Now campaign of the Ministry of Education and associated efforts being made by the MOE and numerous partners to boost reading outcomes.

Although gains made in both regions from midline to endline, the gains achieved by Mtwara students were greater than those achieved by Lindi students on one: oral passage comprehension. In this case, student made a gain of almost six percentage points from midline to endline, while the average percent correct for this task dropped by about one percentage point from midline to endline.

Table 15: Percent correct by EGRA task, Mtwara vs. Lindi

Task	Mtwara			Lindi		
	Midline	Endline	Gain (Endline-Midline)	Midline	Endline	Gain (Endline-Midline)
1. Phonemic awareness	26.7%	30.5%	3.7%*	25.7%	28.9%	3.2%
2. Letter name recognition	17.1%	21.6%	4.5%***	19.6%	24.3%	4.7%***
3. Letter sound knowledge	16.2%	17.5%	1.3%	15.1%	17.3%	2.1%*
4. Invented word decoding	10.1%	14.8%	4.6%***	16.3%	22.4%	6.1%***
5a. Passage reading (timed)	11.5%	19.1%	7.6%***	19.2%	26.9%	7.7%***
5b. Passage comprehension (timed)	8.0%	14.8%	6.8%***	12.1%	20.4%	8.3%***
6a. Passage reading (untimed)	22.6%	35.3%	12.7%***	41.0%	46.5%	5.6%
6b. Passage comprehension (untimed)	11.9%	20.9%	9.1%***	21.8%	26.8%	5.0%*
7. Oral passage comprehension	38.8%	44.4%	5.6%** ^a	46.5%	45.51%	-1.03% ^a
Overall EGRA Score	18.1%	25.1%	7.0%***	24.1%	30.0%	5.8%***

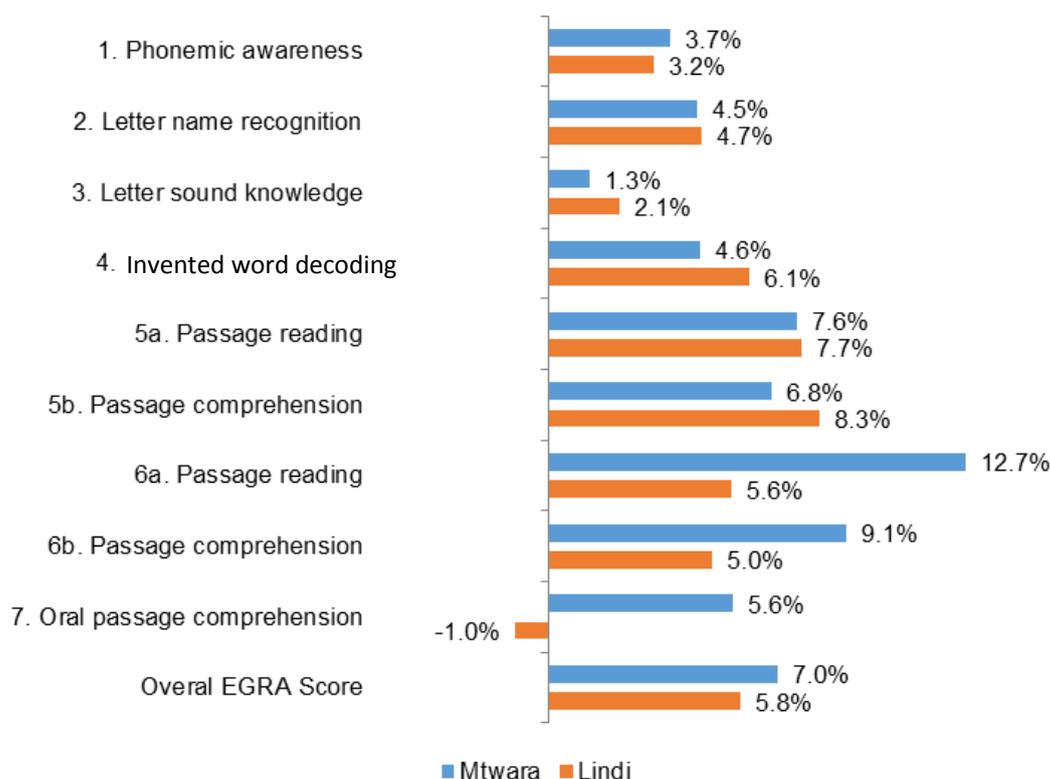
*Significant difference between midline and endline at $p < 0.05$

**Significant between midline and endline at $p < 0.01$

***Significant between midline and endline at $p < 0.001$

^a Significant difference between Mtwara and Lindi at $p < 0.05$

Figure 14: Midline to Endline Gains by Task - Mtwara vs. Lindi



It should be remembered that the schools comprising the Lindi sample were selected purposively, and that student and school performance data were not included in selection criteria. As such, these findings should not be generalized to schools unlike those in the selected sample. Additionally, Lindi schools were introduced as a point of comparison at the time of the midline evaluation. As a result, no baseline data are presented within this section.

An analysis of overall EGRA scores by gender shows that these average scores increased by about 4 percentage points for girls in Mtwara and significantly for boys by about 10 percentage points from midline to endline. Gains for boys in Mtwara were found to be significantly higher than those made by girls. In Lindi, both boys and girls were observed to increase their average overall EGRA scores by almost 7 percentage points, and the difference was significant for girls. No differences between girls' and boys' gain were detected from midline to endline in Lindi.

Table 16: Overall EGRA Scores by Gender, Mtwara vs. Lindi

	Mtwara			Lindi		
	Midline	Endline	Gain	Midline	Endline	Gain
Girls	19.1%	23.2%	4.1%	24.8%	31.3%	6.5%**
Boys	17.2%	27.5%	10.3%***	22.3%	28.8%	6.5%

**Significant difference between midline and endline at $p < 0.01$

***Significant difference between midline and endline at $p < 0.001$

An analysis of gains between midline to endline are presented for both girls and boys by task in Mtwara and Lindi in Table 21. Here, we see that boys in Mtwara made significantly greater gains than girls in Mtwara in the invented word decoding, timed reading and comprehension, and untimed reading and comprehension tasks. In Lindi, while increases were observed across all tasks by boys and girls (excluding the oral passage comprehension task), no significant differences were detected between groups. On average, it appears that boys and girls made equivalent progress across tasks from midline to endline.

Table 17: Midline to Endline EGRA Gains by Task and Gender – Mtwara vs. Lindi

Task	Mtwara Gain (Midline to Endline)		Lindi Gain (Midline to Endline)	
	Girls	Boys	Girls	Boys
1. Phonemic awareness	2.1%	4.9%	5.8%	0.6%
2. Letter name recognition	3.2%	5.9%	5.0%	5.1%
3. Letter sound knowledge	-0.1%	2.6%	2.8%	2.0%
4. Invented word decoding	0.8%	8.8%**	7.4%	5.7%
5a. Passage reading	3.4%	12.4%*	9.0%	7.6%
5b. Passage comprehension	3.7%	10.4%*	9.1%	8.8%
6a. Passage reading	7.2%	19.1%*	5.4%	9.2%
6b. Passage comprehension	7.4%	11.7%	3.8%	9.0%
7. Oral passage comprehension	3.9%	8.2%	-1.7%	0.6%
Overall EGRA Score	4.1%	10.3%*	6.5%	6.5%

*Significant different between boys and girls at $p < 0.05$

**Significant difference between boys and girls at $p < 0.01$

***Significant difference between boys and girls at $p < 0.001$

An analysis of gains from midline to endline in terms of fluency is summarized in Table 22 below. These findings show that, on average, students in Mtwara correctly read about 4 more words per minute on the timed reading passage task at endline than at the midline and that this difference was significant. In Lindi, this gain was also about 4 CWPM and was also significant.

For the untimed reading passage, the average student in Mtwara read almost 9 more words correctly per minute on the endline assessment than the midline assessment. In Lindi, the average student gain was over 6 CWPM from midline to endline. Gains in both regions for fluency in the untimed reading passage were found to be significant.

Despite increases in the average number of words read correctly per minute, no significant differences were detected between the gains made by Mtwara students and the gains made by Lindi. This was true for both the timed and untimed reading passages and suggests that students in both regions made equivalent strides in fluency from midline to endline.

Table 18: Gains in fluency from midline to endline – Mtwara vs. Lindi

	Mtwara			Lindi		
	Midline	Endline	Gain	Midline	Endline	Gain
Task 5a: Reading Passage (Timed) Correct Words per Minute	6.49	10.78	4.29***	10.93	15.15	4.22***
Task 6a: Reading Passage (Untimed) Correct Words per Minute	11.91	20.77	8.86***	21.32	27.99	6.67**

**Significant difference between midline and endline at $p < 0.01$

***Significant difference between midline and endline at $p < 0.001$

Still another way of comparing gains in Mtwara and Lindi from midline to endline is to look at changes in the number of students obtaining zero scores, or in other words, the proportion of students who were unable to answer any item correctly on

a given task. As shown in Table 23, significant reductions in the proportion of students obtaining zero scores were observed in Mtwara across all tasks except for phonemic awareness (where there was a significant increase), and for the oral passage comprehension. In many of these cases, decreases in the proportion of zero score students were quite substantial, and for six of the tasks, exceeded 12 percentage points.

On the contrary, differences in the proportion of students receiving zero scores from midline to endline were not observed to be significant across any of the tasks for students in Lindi. This suggests that, on average, the number of Lindi students obtaining zero scores from midline to endline stayed the same. For both Mtwara and Lindi, no significant changes were observed in the percentage of students obtaining zero scores on the overall EGRA (i.e. students who did not answer a single item correctly across the entire assessment).

In the case of zero scores, significant differences were detected between Mtwara and Lindi by EGRA task. Decreases in the proportion of Mtwara students obtaining a zero score in invented word decoding, the timed reading passage, and the untimed reading passage were greater than the decreases of students receiving zero scores for these tasks in Lindi from baseline to midline ($p < 0.05$). This suggests that participation in TZ21 programming has positively benefitted struggling students in these particular skill areas. Of note, however, is the significant rise in the proportion of Mtwara students obtaining a zero score in phonemic awareness by 14 percentage points from midline to endline. While Lindi students also experienced an increase in zero scores over this same time period, this increase was seen to be significantly larger in Mtwara than in Lindi ($p < 0.05$).

Table 19: Zero Scores from midline to endline – Mtwara vs. Lindi

	Mtwara			Lindi		
	Midline	Endline	Difference (Endline – Midline)	Midline	Endline	Difference (Endline – Midline)
1. Phonemic awareness	20.5%	34.5%	14.0%***,a	28.6%	32.8%	4.2% ^a
2. Letter name recognition	17.5%	10.8%	-6.8%**	12.2%	10.4%	-1.8%
3. Letter sound knowledge	19.8%	12.0%	-7.8%**	14.9%	11.9%	-3.1%
4. Invented word decoding	68.8%	54.5%	-14.3%***,a	47.3%	43.2%	-4.2% ^a
5a. Passage reading (timed)	67.8%	55.5%	-12.3%** ^a	45.6%	45.7%	0.1% ^a
5b. Passage comprehension (timed)	74.0%	61.3%	-12.8%***	58.5%	51.8%	-6.7%
6a. Passage reading (untimed)	68.0%	54.8%	-13.2%***,a	46.6%	46.2%	-0.4% ^a
6b. Passage comprehension (untimed)	74.0%	58.8%	-15.2%***	54.4%	48.7%	-5.7%
7. Oral passage comprehension	13.3%	9.7%	-3.5%	10.9%	9.3%	-1.5%
Overall EGRA Score	2.3%	2.0%	-0.3%	1.0%	1.0%	0.0%

*Significant difference between midline and endline at $p < 0.05$

**Significant difference between midline and endline at $p < 0.01$

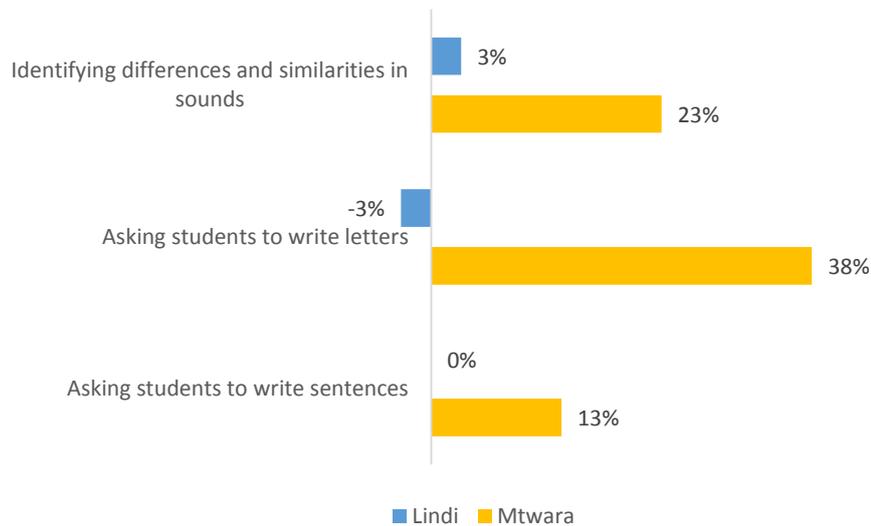
***Significant difference between midline and endline at $p < 0.001$

^aSignificant difference between Mtwara and Lindi at $p < 0.05$

Results from classroom observations

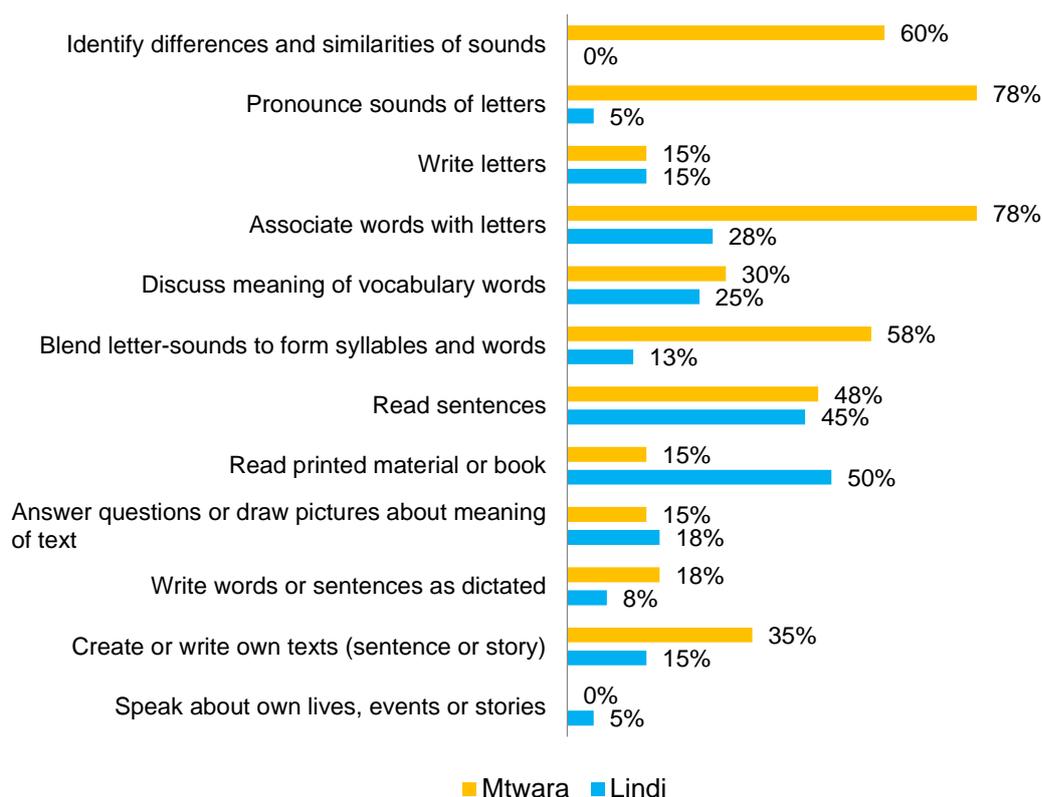
From midline to endline, teachers' use of selected teaching strategies was observed more frequently in Mtwara than in Lindi. These include helping students identify differences and similarities of sounds, asking students to write letters, and asking students to write sentences.

Figure 15: Use of reading strategies in Mtwara and Lindi: Changes in frequency of observations from midline to endline



As shown in Figure xxx, classroom observations at the time of endline exhibited greater proportions of Mtwara teachers than Lindi teachers using phonics-oriented teaching methods, such as identifying differences and similarities of sounds, pronouncing the sounds of letters, blend letter-sounds to form syllables and words, and associate words with letters. These differences are similar to those observed at midline, however, and distinct changes in teacher practice were not observed in either region from midline to endline. As a result, it is difficult to link changes in teachers' practices to changes in student reading performance. However, these persistent differences between Mtwara and Lindi may have contributed to the greater reductions in zero scores observed in Mtwara.

Figure 16: Percentage of teachers using reading techniques and methods – Mtwara vs. Lindi



Question 3: To what extent are students, teachers, head teacher, schools and communities implementing practices or activities to support reading?

In this section, we consider the accounts in Zanzibar and Mtwara of Head Teachers, teachers, and SMC members concerning the conditions of teaching and learning and their roles in improving reading learning outcomes. Specifically, we examine the following:

- Head Teachers: Lesson planning, scheduled time for reading, their role as instructional leaders, and classroom observations and feedback to teachers
- Teachers: How they know when their students are having trouble reading, and conduct participating in teacher meetings
- SMCs: Their main functions and support they provide to improve quality

For more details on the results of these interviews, see Annex D.

Practices & activities supporting reading, Zanzibar

Lesson planning: At the time of the midline, of 40 Head Teachers interviewed, only thirteen (33%) said their teachers are required to prepare daily lesson plans. By the time of the endline evaluation, 32 of 40 Head Teachers (80%) reported that their teachers are required to prepare daily lesson plans. All of these 32 Head Teachers also said that they reviewed those lesson plans.

Time for reading: Similar to the midline, all teachers interviewed reported having a scheduled time for teaching reading in Standard 2 at endline. For this endline report, frequency of reading lessons taught is not being reported in light of the fact

that teachers showed some confusion when asked how many reading lessons they taught each week, some stating that they didn't have a separate class called "reading" and that they taught reading in several classes, including language. Because of this ambiguity, this item was removed from the analysis.

Practices & activities supporting reading, Mtwara: Head Teachers' views

Lesson planning: As with the midline, while a majority of Head Teachers reported at midline that their teachers are required to prepare daily lesson plans (n=26, or 65%), over one-third (n=14, 35%) do not. Almost all Head Teachers who require lesson plans reported to review them (n=24). Half of the Head Teachers interviewed at endline said that their school taught Standard 2 reading every day (n=20, or 50%), a decrease from the 33 Head Teachers reporting the same at midline. Fifteen Head Teachers (38%) said Standard 2 reading was taught two to four times per week at endline. A strong majority reported that reading lessons tended to last 16-30 minutes (n=32, or 80%), suggesting a potential reduction in the amount of class time reportedly reserved for reading from midline, where half of Head Teachers stated that reading instruction lasted 31 to 45 minutes.

Additional findings on factors that supported reading in TZ21 schools can be found in Annex E.

Question 4: Which factors at the student, classroom, and school levels are associated with strong reading performance?

Following the EGRA assessment, each student was asked a number of questions about school attendance, home life, and whether they had attended preschool. Their teachers were also observed and interviewed, and Head Teachers were interviewed. The results of these data sources were correlated with students who scored highest on the EGRA. The results of the student questionnaire show that in both places, students scored higher on EGRA when they went to a nursery or preschool, when they had an English and Kiswahili reading textbook in school, when they got their exercise books checked at home by their parents, when they had books or reading materials at home, and when they had parents read them stories at home.

Interestingly, the results of the teacher and Head Teacher questionnaires and teacher observations reveal different outcomes for each place. Students in Zanzibar, but not in Mtwara, performed better when their teachers were observed using materials like letter or word tiles, program materials, exercise books or slates; when they had letter cards in the classroom (though were not observed using them); and when they were observed leading students in basic phonics activities like sounding out words and associating words with letters – all skills typically taught at the initial stages of literacy acquisition (i.e., Standard 1). In contrast, students performed better in Mtwara, but not in Zanzibar, when their teachers read out loud to their students and guided students to read printed material or books – activities that are designed to increase students' comprehension and become more autonomous readers (i.e., Standard 2). Moreover, EGRA scores were higher in Mtwara, but not in Zanzibar, when teachers identified students who were having trouble reading, and when their teachers had received feedback on the content of their lessons. Finally, students in Mtwara, but not in Zanzibar, performed better when their Head Teachers reported having communicated often with their teachers as a way of helping them teach better and holding reading or book fairs at school to support reading – signs that some schools in Mtwara, efforts by Head Teacher to practice supportive supervision and cultivate

a “culture of reading” were beginning to have an impact. These differences suggest a profile where the teaching of reading at the end of Standard 2 in Zanzibar tends to focus on Standard 1 skills, whereas in Mtwara, it focuses more on higher level skills, and in Mtwara schools, practices like formative assessment, coaching, and developing a culture of reading are showing results in reading ability (discussed more in the next section):

Table 20: EGRA performance and student, teacher, and Head Teacher characteristics, Mtwara and Zanzibar

	Mtwara	Zanzibar
Student characteristics: Children had significantly higher EGRA scores when they:		
• Went to a nursery or pre-school	x	x
• Had an English reading textbook in school	x	x
• Had a Kiswahili reading textbook in school	x	x
• Got their exercise books checked at home by their parents	x	x
• Had books or reading materials at home	x	x
• Had father that can read and write		x
• Said that their parents read them stories at home	x	x
<i>Source: Student interviews correlated with EGRA scores</i>		
Teacher characteristics: Children had significantly higher EGRA scores when their teacher:		
• Arranges letter or word tiles to help children learn to manipulate/use the letters of the alphabet		x
• Had students sound out words		x
• Used specific reading program materials as criteria to assess whether students achieved the learning outcomes of the reading lesson		x
• Used exercise books or slates to teach reading		x
• Had letter cards available in the classroom		x
• Guided students to associate words with letters		x
• Guided students to read printed material or books	x	
• Read out loud to their students	x	
• Used a textbook	x	
• Identified students as having trouble with reading as those who can't answer questions about text s/he just read	x	
• Taught reading lessons 3-4 times/week (as opposed to every day or 1-2 times per week)	x	
• Received observation feedback on the content of lessons (as opposed to the use of teaching and learning materials)	x	
<i>Source: Classroom observations and teacher interviews correlated with EGRA scores</i>		
Head Teacher characteristics: Children had significantly higher EGRA scores when their Head Teacher:		
• Worked as a teacher for 6-10 years before taking a Head Teacher position (as opposed to 0-5 years or more than 10 years)	x	
• Communicated often with teachers as a way of helping them teach better	x	
• Held reading or book fairs at school to support reading	x	
<i>Source: Head Teacher interviews correlated with EGRA scores</i>		

Discussion

Gains in reading skills: The most important finding of this endline evaluation can be found in the substantial progress in reading ability – a pattern continued from the midline. Fluency gains were statistically significant in both Zanzibar and Mtwara. Unlike the midline evaluation, where no significant gains in comprehension were found over baseline, students showed statistically significant improvement on the EGRA comprehension tasks from midline to endline. Moreover, this evaluation found large reductions in zero scores in both Zanzibar, and the proportion of emergent and fluent readers who were able to read with at least 80% comprehension increased from midline to endline in both Mtwara and Zanzibar. This suggests that TZ21 project interventions may have resulted in improved fluency and comprehension – a fact reinforced by the fact that differences in timed passage reading comprehension from midline to endline were significant in Mtwara, but not in Lindi.

Differences from midline to endline: What were some possible reasons for these improvements? At midline, teachers were observed in both Mtwara and Zanzibar using a large percentage of traditional teaching practices in the classroom, such as choral repetition and recitation. While some of these practices still endured at endline, increases in instructional methods that engage students in thinking about their learning were also observed in both Zanzibar and Mtwara, such as beginning the lesson with students' prior knowledge and asking students questions about the lesson. In both Zanzibar and Mtwara, increases in learner-friendly practices such as praising and complimenting students were observed, as well as decreases in criticizing, scolding or punishing students. Improved practices in the teaching of reading were also found. For example, more teachers were observed at endline than at midline in both Zanzibar and Mtwara asking students to think about text, such as creating or writing their own texts, asking students to answer questions about the meaning of text, drawing pictures to show they understand the text, and playing learning games, skits or songs. However, none of these changes in practice was positively correlated with improved reading scores.

Differences between Mtwara and Zanzibar at endline: What does appear to have made a difference in reading scores at endline was the types of teaching practices teachers used, the types of support they received, and the cultivation of a culture of reading. Correlations between teacher characteristics and EGRA scores at endline, for example, showed that students performed better in Zanzibar when teachers used certain basic phonics materials and strategies – practices typically used at the initial stages of literacy acquisition (i.e., Standard 1). In contrast, students performed better in Mtwara when their teachers encouraged them to read printed material or books – practices typically used once students have acquired basic reading skills (e.g., Standard 1). Moreover, students performed better in Mtwara when their teachers identified students who were having trouble reading, and when their teachers had received feedback on the content of their lessons – attributes of schools where formative assessment and coaching practices appear to have been taking root. Finally, students in Mtwara performed better when their Head Teachers reported having communicated often with their teachers as a way of helping them teach better and holding reading or book fairs at school to support reading – signs that some schools in Mtwara, efforts by Head Teacher to practice supportive supervision and cultivate a “culture of reading” were beginning to have an impact.

The effect of TZ21: Differences between Mtwara and Lindi: A comparison of scores in Mtwara and Lindi reveals differences between the two regions and, by extension, possible effects of TZ21 in Mtwara. Several signs point to such effects. First, zero scores declined from midline to endline in Mtwara overall and for each task except phonemic awareness, and differences from midline to endline were statistically significant in all cases. In contrast, though Lindi registered some declines in zero scores, none were statistically significant. Some differences in teaching practices were also observed between the two regions. For example, the practice of identifying differences and similarities in sounds increased from midline to endline in Mtwara by 23 percentage points, but by only 3 percentage points in Lindi. Asking students to write letters increased by 38 percentage points in Mtwara and only 3 percentage points in Lindi.

These developments suggest that TZ21's interventions focusing on the training of teachers in the use of effective literacy instruction skills and materials may have had an impact on student learning.

Differing gender results: As at midline, girls performed better overall than boys in Zanzibar and the difference was statistically significant. The gains realized by both girls and boys from midline to endline were also statistically significant, as were selected tasks within the EGRA for each group. And in Mtwara, boys made significantly greater gains than girls in the invented word decoding, timed reading and comprehension, and untimed reading and comprehension tasks. In Lindi, no significant differences were detected between boys and girls groups. The reasons for these different results by gender are unclear.

Warning signs

The use of new strategies and the availability and use of materials: A mixed record. While some desired changes in the use of new teaching strategies was achieved, others were not. On the positive side, teachers in Zanzibar increasingly used strategies such as beginning the lesson with prior knowledge or things taught increased in Zanzibar, asking students to create or write their own texts, and asking students to answer questions about the meaning of text or to draw pictures to show they understand the text – all key strategies in building students' comprehension skills. The practice of teachers reading out loud remained frequent in Mtwara and increased in Zanzibar, and the frequency of students playing learning games, skits or songs also increased in both Zanzibar and Mtwara. However, some declines in the use of target strategies were also observed, principally in Mtwara, where declines were observed in the identification of the differences and similarities of sounds (83% at midline, 60% at endline), blending letter-sounds to form syllables and words (80% at midline, 58% at endline), and discussing the meaning of vocabulary words (40% at midline, 30% at endline). Similarly, targeted availability and use of materials was achieved in some instances, but not in others. Among the most dramatic positive changes were the availability of student materials, the use of small books and big books, and the evidence of posters in classrooms in both places. However, teacher's guides were rarely found at both midline and endline, and the percentage of teachers in Zanzibar who could show they had a scripted TZ21 lesson plan remained low.

Continued reliance of traditional teaching methods: Though teachers increased their use of a number of effective teaching strategies described above, a number of traditional practices remained. Repetition and recitation remained the most prevalent reading behaviors in both regions. In Zanzibar, incidents where students were observed reading independently declined from 13% at midline to 3%, and in

Mtwara, the use of other approaches such as the identification of the differences and similarities of sounds, the blending letter-sounds to form syllables and words, the teaching of vocabulary also decreased, as well as the use of formative assessment methods such as monitoring student work and observing student activities. In Mtwara, most teachers observed used some of the same assessment strategies at both midline and endline, though others declined, such as the practice of listening to students read aloud and using a reading assessment tool.

The challenge of the resilient half. In both Zanzibar and Mtwara, though zero scores overall went down, half or more of the student population continued to obtain zero scores at endline – a troubling pattern that suggests that while instruction is improving, a large proportion of students is still not benefitting from these improvements. This pattern suggests a critical need for additional teacher support in the use of strategies such as formative assessment, individualized instruction, and the use of remedial strategies so that all children can benefit from support provided by efforts such as TZ21.

Recommendations

1. Continue to provide support to teachers to assist them with the transition from more traditional, whole-class methods to ones that enable learners to work independently in groups, pairs, and individually.
2. Continue to develop students' reading comprehension skills: Assist teachers with the transition from teaching foundational skills to strengthening the reading of connected text and improved comprehension.
3. Continue to provide teachers with regular school-based support focused on early grade reading, including follow-up at school level. Reinforce mechanisms for mentoring & coaching as well as support for literacy acquisition in the home and community. These include WEC observations and support as well as training provided by experts from the MOEVT and education partners.
4. Provide teachers with support in strategies to help struggling students (especially ones with zero scores), including formative and summative assessment strategies, remediation strategies (see next recommendation), individualized instruction, and identification of early warning signs that students are struggling.
5. Provide teachers with explicit support in remediation strategies: Establish a remedial program to inform teachers, Head Teachers, education administrators, and parents of “early warning signs” of children who are struggling to become early grade readers. Help teachers learn to diagnose reading problems, and provide them with tools for addressing these problems, with special adaptations for use in large classes, where a large percentage of children are likely to need support.
6. Ensure the timely distribution and appropriate use of instructional materials. Promote the distribution of textbooks and supplementary materials to Tanzanian primary schools, and systematically monitor their use in the classroom. Establish standards for the use of materials instruction to ensure their regular and effective use.

Annex A: EGRA passages

Task 5a: Timed reading

Kiswahili	English
<p>Bahati anapenda kusoma. Yeye huamka asubuhi na mapema kwenda skuli.</p> <p>Wazazi wake humwambia asome kwa bidii. Wao humnunulia penseli, vitabu na maandazi.</p> <p>Bahati ni mtoto mzuri. Lakini siku moja, rafiki yake alimwambia wakaibe maembe kwa jirani.</p> <p>Walipanda mwembe kwa ngazi. Jirani akaja. Aliwaambia washuke.</p> <p>Bahati na rafiki yake walishuka. Jirani aliwaeleza ubaya wa kuiba. Kisha akawasamehe.</p>	<p>Bahati likes to study. She wakes up early in the morning to go to school.</p> <p>Her parents told her to study hard. They bought her pencils, books and maandazi.</p> <p>Bahati is a nice child. But one day her friend told her to steal a mango from the neighbor.</p> <p>She climbed the mango with a ladder. The neighbor came and told them to get off.</p> <p>Bahati and her friend got off the tree. The neighbor explained them that stealing is bad and in the end he pardoned them.</p>

Task 5b: Comprehension questions

Kiswahili	English
<p>1. Bahati anapenda kufanya nini? [kusoma]</p>	<p>What does Bahati like to do? [To study]</p>
<p>2. Wazazi humwambia Bahati afanye nini? [asome kwa bidiii]</p>	<p>What did the parents tell Bahati to do? [To study hard]</p>
<p>3. Rafiki ya Bahati alimwambia nini? [Wakaibe maembe kwa jirani]</p>	<p>What did her friend tell her? [To steal a mango from the neighbor]</p>
<p>4. Walitumia nini kupanda mwembe? [Ngazi]</p>	<p>What did they use to climb the mango tree? [A ladder]</p>
<p>5. Jirani aliwaeleza nini? [ubaya wa kuiba]</p>	<p>What did the neighbor explain them? [that stealing is bad]</p>

Task 6a: Untimed reading

Kiswahili	English
Amina anapenda kusoma. Yeye anaishi na wazazi wake. Wao wanakula nyama kila siku.	Amina likes to study. She lives with her parents. They eat meat every day.
Amina ni mwanafunzi hodari. Siku moja aliwaeleza wazazi wake faida za vyakula bora.	Amina is a diligent student. One day she explained to her parents about the importance of eating healthy.
Tokea siku hiyo wazazi wake waliamua kubadilisha kula nyama na kununua samaki au mboga za majani ili kujenga mwili.	From that day her parents stopped eating meat and started buying fish and vegetables in order to build their body.
Wazazi walifurahi sana kwa sababu walifahamu umuhimu wa kusoma.	The parents were very happy because they understood the importance of studying.

Task 6b: Comprehension questions

Kiswahili	English
1. Amina anapenda kufanya nini? [Anapenda kusoma]	What does Amina like to do? [Amina likes studying]
2. Amina aliwaeleza nini wazazi wake? [Faida za vyakula bora]	What did Amina explain to her parents? [The importance of eating healthy]
3. Wazazi wa Amina walianza kununua nini? [samaki na mboga za majani]	What did Amina's parents start to buy? [fish and vegetables]
4. Kwa nini wazazi walifurahi? [Kwa sababu walifahamu umuhimu wa kusoma; kwa sababu walifahamu umuhimu wa vyakula bora]	Why were her parents happy? [Because they understood the importance of studying; because they understood the importance of eating healthy]

Task 7a: Oral comprehension

Kiswahili	English
<p>Ali anaishi karibu na msitu. Siku moja alisikia ndege akiimba. Alimfuata ndege huyo hadi kati kati ya msitu. Ghafila, alihisi njaa. Akapanda juu ya mti wa matunda na mara alimuona nyoka. Aliteremka chini haraka. Kabla hajafika chini, alimuona mbwa mwitu. Alivunja tawi la mti na kumrushia. Wakati mbwa mwitu alipokimbia mbali Ali aliteremka chini na kukimbia nyumbani.</p>	<p>Ali lives near a forest. One day he heard a bird singing. He followed the bird up to the heart of the forest. Suddenly, he felt hungry. He climbed a fruit tree and immediately he saw a snake. He climbed down the tree fast. Before reaching the ground, he saw a wild dog. He broke a branch of the tree and threw it at the dog in order to make him leave. When the wild dog ran far, Ali climbed down the tree and ran home.</p>

Task 7b: Comprehension questions

Kiswahili	English
<p>1. Ali anaishi wapi? [karibu na msitu]</p>	<p>Where does Ali live? [near a forest]</p>
<p>2. Ali alisikia sauti ya nani? [Ya ndege]</p>	<p>Which sound did Ali hear? [of a bird]</p>
<p>3. Kwa nini Ali alipanda mti wa matunda? [Kwa sababu alihisi njaa, kwa sababu aliona matunda]</p>	<p>Why did Ali climb a fruit tree? [Because he felt hungry, because he saw the fruits]</p>
<p>4. Kwa nini Ali alishuka haraka? [Alimuona nyoka]</p>	<p>Why did Ali climb down fast? [He saw a snake]</p>
<p>5. Kwa nini Ali ni kijana hodari? [kwa sababu alimrushia tawi mbwa mwitu, kwa sababu aliweza kumfukuza mbwa mwitu]</p>	<p>Why is Ali a brave boy? [Because he threw a branch at the wild dog, because he managed to make the wild dog run away]</p>

Annex B: Test and item statistics, EGRA

Zanzibar: The quality of the EGRA test in Zanzibar was strong, with a high reliability rating (Cronbach's Alpha) of 0.914. The mean percentages correct across tasks were in the acceptable range - for this kind of assessment, above .10 and below .90. Item-total correlations – a measure of discrimination power of each item – were also acceptable at .30 and above. (An item-total correlation closer to 1.0 confirms that stronger students performed better on the more difficult items. Item-total correlations near zero are an indication that the item cannot discriminate between stronger and weaker students and should therefore be excluded from analysis.) The item statistics for the EGRA test in Zanzibar were as follows:

Table 21: Task statistics: Means and item-total correlations

Task (Subtest)	Mean percentage correct	Item-Total Correlation
1. Phonemic awareness	30.43	0.30
2. Letter name recognition	26.17	0.73
3. Letter sound knowledge	21.04	0.77
4. Invented word decoding	19.99	0.91
5a. Passage reading	23.46	0.94
5b. Passage comprehension	19.35	0.87
6a. Passage reading	40.99	0.86
6b. Passage comprehension	26.06	0.84
7. Oral passage comprehension	50.35	0.37
Overall EGRA Score	29.72	

Mtwara: As with Zanzibar, test statistics for the EGRA test in Mtwara were strong, with a high reliability rating (Cronbach's Alpha) of 0.911. The mean percentage correct for each task were in the acceptable range, between 0.10 and 0.90, and item-total correlations were acceptable (and in some cases very strong) at 0.44 and above. Means and item-total correlations are presented in the following table:

Table 22: Task statistics: p-values and item-totals, Mtwara

Task (Subtest)	Mean percentage correct	Item-Total Correlation
1. Phonemic awareness	30.45	0.48
2. Letter name recognition	21.61	0.70
3. Letter sound knowledge	17.49	0.58
4. Invented word decoding	14.76	0.91
5a. Passage reading, timed	19.08	0.90
5b. Passage comprehension	14.80	0.85
6a. Passage reading, untimed	35.30	0.90
6b. Passage comprehension	20.94	0.84
7. Oral passage comprehension	44.40	0.44
Reading Summary Score	25.06	

Annex C: Results of T-tests of differences between time points, EGRA Standard 2, Timed Reading Passage (60 seconds)

Statistical significance (generalization to a population) is based on the computation of an estimate of population parameters (mean) and its standard error. To perform tests of difference between fluency rates on the timed reading passage (60 seconds), standard errors were computed for the mean correct words read per minute (CWPM) for each EGRA administration (baseline, midline and endline). As an example, the following formula shows how the standard error is derived for the difference between fluency rates at endline and the fluency rates at midline.

$$\sigma_{(\text{mean endline} - \text{mean midline})} = \sqrt{\sigma_{(\text{mean endline})}^2 + \sigma_{(\text{mean midline})}^2}$$

To perform a statistical test of difference between fluency rates at endline and at midline, the difference between the mean CWPM is divided by this standard error. This “standardized difference” is used to compute a p-value based on a normal distribution of student fluency rates.

As can be seen in the tables below, in Zanzibar, the mean fluency score for students on the timed reading passage (60 seconds) at baseline was 7.52 CWPM, 10.17 CWPM at midline, and 13.50 CWPM at endline. The T-test of difference between the baseline and endline fluency rates shows the standardized difference between means was nearly 6 CWPM. The p-value, calculated on the basis of a 2-tailed analysis, is .006, indicating that the difference in means is statistically significant and the chances of there being no differences in baseline and endline fluency rates is very small. The standardized difference between midline and endline fluency rates is almost 3 CWPM, and this is also statistically significant at $p < 0.000$.

In Mtwara, the mean fluency score for students on the timed reading passage (60 seconds) at baseline was 4.73 CWPM, 6.49 CWPM at midline, and 10.78 CWPM at endline. The T-test of difference between the baseline and endline fluency rates shows the standardized difference between means was about 7 CWPM. The p-value, calculated on the basis of a 2-tailed analysis, is 0.000, indicating that the difference in means is statistically significant and the chances of there being no differences in baseline and endline fluency rates is very small. The standardized difference between midline and endline fluency rates is about 4.5 CWPM, and this is also statistically significant at $p < 0.000$.

Lindi students did not participate in the baseline EGRA, however the average student read about 11 words correctly per minute on the midline. At endline, the average fluency rate was about 15 CWPM. The standardized difference between midline and endline was almost 4 CWPM, and this difference was found to be statistically significant at $p < 0.000$.

Table 23: Task 3a: CWPM adjusted for time remaining, baseline

TZ21 Region	N	Minimum	Maximum	Mean	Std. Deviation	S.E.
Zanzibar	799	0	57.93	7.52	13.46	0.48
Mtwara	748	0	63.4	4.73	10.78	0.39

Table 24: 5a: CWPM adjusted for time remaining, midline

TZ21 Region	N	Minimum	Maximum	Mean	Std. Deviation	S.E.
Zanzibar	386	0.00	61.09	10.17	13.93	0.71
Mtwara	391	0.00	49.00	6.49	11.39	0.58
Lindi	378	0.00	58.95	10.93	13.12	0.67

Table 25: CWPM adjusted for time remaining, endline

TZ21 Region	N	Minimum	Maximum	Mean	Std. Deviation	S.E.
Zanzibar	399	0.00	105.00	13.50	17.91	0.90
Mtwara	400	0.00	71.74	10.78	15.23	0.76
Lindi	396	0.00	84.00	15.15	16.89	0.85

Table 26: T-test of Difference between baseline and endline fluency rates (Zanzibar and Mtwara Only), Task 3a/5a: CWPM adjusted for time remaining

Region	S.E. Baseline-Endline	Diff. Baseline-Endline	Standardized Diff. Baseline-Endline	p
Zanzibar	1.015	5.978	5.888	0.000
Mtwara	0.857	6.053	7.059	0.000

Table 27: T-test of Difference between midline and endline fluency rates (All Places), Task 5a: CWPM adjusted for time remaining

	S.E. Midline- Endline	Diff. Midline- Endline	Standardized Diff. Midline-Endline	p
Zanzibar	1.143	3.331	2.914	0.006
Mtwara	0.955	4.293	4.496	0.000
Lindi	1.084	4.220	3.892	0.000

Annex D: Minimum Performance Criteria

Minimum Performance Criteria: Earlier in the project, TZ21 had established a set of Minimum Performance Criteria in which selected indicators would be tracked from midline to endline in order to determine the extent to minimum conditions had been met in TZ21 schools. Three categories were established: student behavior (10 items), teacher behavior (5 items), and learning materials (5 items).

In the category of student behavior, Mtwara's average score decreased by 6% from midline to endline, while Lindi's decreased 22%. In the category of teacher behavior, Mtwara's average score increased by 53% while Lindi's decreased by 20%. And in the category of learning materials, Mtwara's average score increased by 321% while Lindi's decreased by 33% - see the following tables:

Table 28: Student behaviors observed in the classroom, midline to endline, Mtwara vs. Lindi

Note: Green=meeting 50% competency threshold, pink=falls below threshold

Student behavior	Mtwara			Lindi		
	Midline	Endline	% change	Midline	Endline	% change
1. Identify differences and similarities of sounds	83%	60%	-27%	3%	0%	-100%
2. Blend letter-sounds to form syllables and words	80%	58%	-28%	35%	13%	-64%
3. Read sentences	48%	48%	0%	58%	45%	-22%
4. Read out loud together (choral reading)	90%	65%	-28%	85%	68%	-21%
5. Read out loud one student at a time	83%	75%	-9%	83%	83%	0%
6. Read out loud to another student (paired reading)	5%	5%	NA	5%	0%	NA
7. Listening to teacher read out loud	95%	95%	0%	88%	75%	-14%
8. Answer teacher's questions	93%	100%	8%	80%	90%	13%
9. Answer questions about meaning of text, or draw picture to show they understand a text	12%	15%	25%	10%	18%	75%
10. Play learning games, skits or songs organized by the teacher	58%	60%	4%	28%	10%	-64%
Average	64%	58%	-6%	47%	40%	-22%
Change in average: Midline to endline	-6%			-7%		

Teacher behavior: Five types of teacher behavior promoted by TZ21 were identified for this analysis, including the use of Kiswahili in instruction (in accordance with the national language policy), reading aloud to students, organizing students in small groups, organizing independent reading, and moving around the classroom.

Of the five targeted behaviors being tracked, teachers were observed using only three at above the 50% threshold, and two of these had already been observed above the 50% threshold at mid-line. In Zanzibar, the percent change from midline to endline was a decrease of 8% while Mtwara saw an increase of 53%. The biggest changes were in observed in Zanzibar where teachers were reading aloud to students (24% increase) and moving throughout the classroom (22% increase). The practices of organizing students in small groups and independent reading was rarely observed:

Table 29: Teacher behaviors observed in the classroom, midline to endline, Mtwara vs. Lindi

Note: Green=meeting 50% competency threshold, pink=falls below threshold

Teacher behavior	Mtwara			Lindi		
	Midline	Endline	% change	Midline	Endline	% change
1. Conducts some or all of the lesson in Kiswahili	93%	98%	5%	80%	100%	25%
2. Reads aloud to students	95%	95%	0%	88%	75%	-14%
3. Organizes students in small groups	10%	8%	NA	0%	5%	NA
4. Organizes independent reading	3%	8%	200%	10%	3%	-75%
5. Moves throughout the classroom/space	60%	65%	8%	73%	60%	-17%
Average	52%	55%	53%	50%	49%	-20%
Change in average: Midline to endline	3%			-2%		

Learning materials and competencies: Of the materials discussed above, TZ21 tracked the use of five kinds of materials to determine the attainment of key competencies in the project – i.e. the extent to which these materials were used in instruction. These materials included big books, letter and flash cards, small books (readers), and posters or wall charts. TZ21 also tracked students’ use of exercise books or slates because of its critical importance in learning to read.

In this area, Zanzibar saw a big increase in use of materials (69% overall) owing to the increased use of student exercise books or slates (100%), and the increased use of small books (186%). Similarly, Mtwara saw a big increase in the use of materials (321% overall), owing to the large increase in the use of big books (1400% increase) and small books (300%):

Table 30: Learning materials used in classroom observation lesson, midline to endline, Mtwara vs. Lindi

Note: Green=meeting 50% competency threshold, pink=falls below threshold

Learning materials	Mtwara			Lindi		
	Midline	Endline	% change	Midline	Endline	% change
1. Big books	5%	75%	1400%	0%	0%	NA
2. Letter cards/flash cards	60%	50%	-17%	20%	10%	-50%
3. Student exercise books and/or slates	23%	8%	-67%	65%	18%	NA
4. Small books	3%	10%	300%	35%	40%	14%
5. Poster/wall charts (with letters, words, pictures)	25%	23%	-10%	20%	8%	-63%
Average	23%	33%	321%	28%	15%	-33%
Change in average: Midline to endline	10%			-13%		

Annex E: Practices supporting reading

Additional findings from teacher, Head Teacher and SMC interviews are summarized here.

Zanzibar

How Head Teachers help their teachers - role as instructional leader: When asked how they help teachers teach better, most Head Teachers (n=34, or 85%) said they communicate often with teachers, followed by encouraging teamwork among their teachers (n=33, or 83%). Head Teachers least often reported that they tried to stay abreast of current innovations in pedagogy (n=13, or 33%). As can be seen in Table 27, patterns of responses were the same as those from the midline.

Table 31: How Head Teachers help teachers teach better - Head Teacher responses from most frequent to least frequent, Zanzibar

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Communicate often with teachers	Communicate often with teachers
2	Encourage teamwork among teachers	Encourage teamwork among teachers
3	Share my knowledge with teachers	Share my knowledge with teachers
4	Communicate often with students and parents	Communicate often with students and parents
5	Stay abreast of current innovations in pedagogy	Stay abreast of current innovations in pedagogy

Frequency of classroom observations & feedback to teachers: Thirty-eight Head Teachers and Deputies (95%) said they had conducted classroom observations this academic year up from 30—or 75%--at midline. Nearly half of the Head Teachers interviewed (48%) said they observe a classroom teacher three times per year and 25% said they conduct observations four times per year (down from 55% at the time of midline). As at midline, most Head Teachers (48%) said they conduct classroom observations for purposes of official performance reviews and 25% said they did so to meet official requirements. Seven Head Teachers suggested that classroom observations were to help the teacher teach better—patterns similar to those seen at midline.

The same number of Head Teachers reported meeting with their teachers individually to discuss their performance once per month or more at endline as at midline—28, or 70% of Head Teachers interviewed.

The frequency with which Head Teacher cited strategies to help their teacher improve reading instruction were very similar to those patterns observed at midline. Most commonly, Head Teachers at endline said they observe their reading lessons (55%) or pair them with a good teacher (53%). Least commonly, they cited discussing reading in Standards 1 and 2 in teacher meetings and asking an inspector for training or assistance.

Table 32: Steps Head Teachers take to help their teachers improve reading instruction - Head Teacher responses from most frequent to least frequent, Zanzibar

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Pair them with a good teacher	Observe their reading lessons
2	Observe their reading lessons	Pair them with a good teacher
3	Provide materials	Coach them
4	Coach them	Provide materials
5	Discuss reading in S1 and S2 in teacher meetings	Discuss reading in S1 and S2 in teacher meetings
6	Ask an inspector for training or assistance	Ask an inspector for training or assistance

Practices & activities supporting reading, Zanzibar: Teachers' views

Enumerators asked teachers if anyone ever observes them or gives them support or feedback. Similar to the midline, just over half (n=21, or 53%) of teachers reported being observed by their Head Teacher, six by a District Education Officer, and seven by other teachers. The greatest proportion of teachers in Zanzibar (n=13, or 33%) reported being observed by school administration personnel with at least once a month some frequency, with eight (20%) reporting at least once a week. Only one teacher reported never having been observed by school administration personnel.

Over half of teachers (n=26, or 65%) reported being mentored by school inspectors or a Ward educational coordinator on reading practices at endline (up from 21 teachers at midline). All but one teacher, or 98%, reported receiving guidance from their Head Teacher at endline (up from 32 at midline). Thirty-eight teachers (95%) said that they received training on how to teach reading using phonics or letter sounds at endline—an increase of seven teachers since midline. Similar to midline, 38 teachers confirmed they had been trained by TZ21 in the teaching of reading, and 23 (58%) had been trained by TZ21 in the use of computers.

How teachers know when their students are having trouble: When asked how they know when a student is having trouble with reading, most teachers said their students can't sound out words (78%) and can't make letter sounds (68%), followed by the student reads very slowly (40%). These patterns at endline mirror those of the most frequently reported midline responses, as can be seen in Table 29, where

responses at both time points are listed in order of most frequently cited to least frequently cited. While teachers at both the midline and endline least often reported that they identify students having trouble with reading as those who do not volunteer for reading activities or who do not participate in class or group reading activities, fewer teachers at endline suggested that struggling readers are those who skip words when reading.

Table 33: How teachers know when students are having trouble with reading, teacher responses from most frequent to least frequent

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Can't sound out words	Can't sound out words
2	Reads very slowly	Can't make letter sounds
3	Can't make letter sounds	Reads very slowly
4	Skips word when reading	Can't complete reading assignments
5	Can't complete reading assignments	Can't answer questions about text s/he just read
6	Can't answer questions about text s/he just read	Can't string words into sentence
7	Does not volunteer for reading activities	Skips word when reading
8	Can't string words into sentence	Does not volunteer for reading activities
9	Does not participate in class or group reading activities	Does not participate in class or group reading activities
10	Other	Other

When asked how they help a student who is having trouble learning to read, an increase from midline to endline were observed in the proportion of teachers saying they work one-on-one with their students from 63% to 75% or who find someone to tutor or work with that student (rising from 18% at midline to 28% at endline). Nearly three-quarters of teachers at endline (73%) said they give extra assignments, and more than half (55%) reported that they pair students who are having trouble learning to read with good students. These patterns were similar to those seen at midline. Interestingly, there was a substantial reduction in the number of teachers talking to parents was observed from midline to endline, such that “talking to parents” fell into the bottom three responses at endline. Fewer teachers also said they reported the student to the head teacher at endline than midline, although “taking no action” remained the least frequent response at both timepoints.

Table 34: How teachers help students who are having trouble learn to read - teacher responses from most frequent to least frequent, Zanzibar

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Give extra assignments	Work one-on-one with student
2	Work one-on-one with student	Give extra assignments
3	Pair with a good students	Pair with a good students
4	Talk to parents	Find someone to work with/tutor the students
5	Recommend outside tutoring	Recommend outside tutoring
6	Report to head teacher	Talk to parents
7	Find someone to work with/tutor the students	Report to head teacher
8	Take no action	Take no action

Teacher meetings: In 38 cases (95%), teachers reported holding teacher group meetings at their school, evidencing an increase from midline (n=33, or 83%). About 35% of teachers reported meeting weekly and 30% reported meeting monthly. In most instances, the School Principal organizes the meetings (80%), although teachers reported organizing meetings in 6 cases (15%).

When asked what topics are typically discussed, most teachers (63%) cited subject content--a topic area that rose from rank order #3 at midline to rank order #1 at endline--and student learning (48%). Only two teachers (5%) mentioned training methods, showing a sharp decline from midline when training methods was most frequently listed as a topic of discussion. As with midline patterns, teachers mentioned administrative issues least often.

Table 35: Topics discussed at teacher meetings – teacher responses from most frequent to least frequent, Zanzibar

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Training methods	Subject content
2	Student learning	Student learning
3	Subject content	Student discipline/behavior
4	Student discipline/behavior	School improvement
5	School improvement	Classroom management
6	Classroom management	Administrative issues
7	Administrative issues	Training methods

Practices & activities supporting reading, Zanzibar: SMCs' views

Main functions: In a departure from midline patterns, when asked what their main functions were, SMC members most frequently said at endline that they support and encourage students (75%), participate in budgeting (70%), help develop school strategy (58%), and improve and monitor school quality (55%). In contrast, SMC

members interviewed at midline most frequently stated that one of their major functions was to help decide expenditures (75%), while supporting and encouraging pupils was rank order #4. While encouraging and supporting the community to enroll school age children was rank order #9 at midline, at endline, this function rose to rank order #7, suggesting an increase in SMC’s observance of this activity as a major function. Raising extra funds, on the other hand, was less frequently mentioned at endline than at midline. Help to set staff incentives and allocating funds for new classrooms remained at the bottom of the list from midline to endline.

Table 36: SMC main functions – Member responses from most frequent to least frequent, Zanzibar

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Help decide expenditures	Support/encourage pupils
2	Improve and monitor school quality	Participate in budgeting
3	Inform/communicate with community	Help develop school strategy
4	Support/encourage pupils	Improve and monitor school quality
5	Participate in budgeting	Monitor and improve school environment
6	Help develop school strategy	Help decide expenditures
7	Monitor and improve school environment	Encourage and support the community to enroll school age children
8	Raise extra funds	Inform/communicate with community
9	Encourage and support the community to enroll school age children	Organize extra educational services
10	Organize extra educational services	Raise extra funds
11	Help set staff incentives	To allocate funds for new classrooms
12	To allocate funds for new classrooms	Help set staff incentives
13	Other	Other

Issues raised at the most recent SMC meetings appear concurrent with SMC members’ views of their major functions. SMC members most frequently reported that they dealt with issues of student attendance (43%), school building and grounds (35%), and students’ contributions (13%). These patterns are very similar to those observed at midline. About half of members at endline (48%) cited “other” issues handled by the SMC, and these tended to relate to concerns about school infrastructure, students’ health, exam results, and financing school projects. SMC members least frequently mentioned discussing pupil/teacher ratios (8%), the distribution of scholarships or grants (3%), and none of the SMCs recently discussed midday meals (0%), or the distribution of books at endline (0%). Again, the rank order patterns of these, “least discussed” topics were similar to those observed at midline.

Table 37: Issues raised at most recent SMC meeting – Member responses from most frequent to least frequent, Zanzibar

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Student attendance	Other
2	Other	Student attendance
3	School building/ground	School building/ground
4	Students' contribution	Students' contribution
5	School security	School security
6	Teacher absenteeism	Teacher absenteeism
7	Other aspects of teacher behavior	Other aspects of teacher behavior
8	School materials	School materials
9	Pupil/teacher ratios	Pupil/teacher ratios
10	Midday meals	Distribution of scholarships/grants
11	Distribution of scholarships/grants	Midday meals
12	Distribution of books	Distribution of books

Support for quality: As at the midline, all 40 SMC's reported being responsible for monitoring teacher behavior and performance. When asked what their SMC had done to improve the quality of teaching and learning at their school last year, members most frequently cited creating professional development opportunities for teachers (48%), providing teaching-learning equipment (43%), providing teaching-learning aids and materials (43%), and supporting in-class teacher assistance by parents (40%). These responses differed from those observed at the midline, when textbook purchases (48%) and improved school infrastructure (43%) were most frequently cited as activities implemented to promote the quality of teaching and learning. Instead, these activities fell to rank order 7 and rank order 5 at endline, respectively.

Cultural clubs and events, the purchase of furniture, and Olympiad coaching and preparation remained the least cited activities of focus for SMCs looking to improve the quality of teaching and learning from midline to endline. Sports clubs and events fell from rank order #8 at midline to rank order #10 at endline. Together, these findings suggest that SMCs consistently seem to focus primarily on the procurement of physical materials and equipment for their school, as well as the support of parents in classrooms.

Table 38: SMC roles in promoting quality in school - Member responses from most frequent to least frequent, Zanzibar

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Professional development opportunities for teacher	Professional development opportunities for teacher
2	Textbook purchases	Teaching-learning aids and materials
3	In-class teacher assistance by parents	Teacher-learning equipment
4	Improved school infrastructure	In-class teacher assistance by parents
5	Teaching-learning aids and materials	Improved school infrastructure
6	Teacher-learning equipment	Extra-curricular educational services and lessons
7	Professional journals or books to teachers	Textbook purchases
8	Sports clubs and events	Other
9	Extra-curricular educational services and lessons	Professional journals or books to teachers
10	Cultural clubs and events	Sports clubs and events
11	Furniture	Cultural clubs and events
12	Other	Furniture
13	Olympiad coaching and participation	Olympiad coaching and participation

Mtwara

How Head Teachers help their teachers - role as instructional leader: When asked how they help teachers teach better (across all subjects), almost all said they communicate often with teachers (80%) and encourage teamwork among teachers (70%). The frequency of these responses mirror patterns observed at midline, which can be seen in Table 35 listing Head Teacher responses in order of most frequently cited to least frequently cited at midline and endline. Interestingly, while communicating often with students and parents was a rank order #3 response at midline, this fell to the least frequently cited response at endline suggesting that Head Teachers now use this strategy least often as a way of supporting teachers' improvement in practice.

Table 39: How Head Teachers help teachers teach better – Head Teacher responses from most frequent to least frequent, Mtwara

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Communicate often with teachers	Communicate often with teachers
2	Encourage teamwork among teachers	Encourage teamwork among teachers
3	Communicate often with students and parents	Share my knowledge with teachers
4	Share my knowledge with teachers	Stay abreast of current innovations in pedagogy
5	Stay abreast of current innovations in pedagogy	Communicate often with students and parents

Frequency of classroom observations & feedback to teachers: Similar to midline results, 37 (93%) Head Teachers and Deputies said they had conducted classroom observations this academic year. Again, over half said they observe more than three times per year (n=25, 64%), the others less frequently. At endline, 15 teachers, or 38%, said that the main purpose of observations was to help teachers improve teaching, although another 17 (43%) said they did so for official performance reviews and to meet officially required number. These patterns are also consistent with those observed at midline.

When asked whether they meet with their teachers individually to discuss their performance, most (n=34, or 85%) Head Teachers said that they did so once per month or more. Again, similar to the midline, most Head Teachers said that they offer feedback to teachers in one-on-one discussions (n=31, 78%), and over one-third did so in group discussions (n=14, or 35%). As with the midline, only some of the Head Teachers interviewed said that they provided written feedback to teachers in the form of a report (n=7, or 18%).

When asked what steps they take to help their teachers improve reading instruction, Head Teachers most frequently said that they discuss reading in their Standard 1 and 2 teacher meetings (45%) and provide their teachers with materials (43%). This shows some difference in comparison with midline patterns, where Head Teachers most frequently stated that they provide materials to support teacher improvement in reading instruction (73%) and then observe reading lessons (65%). Head Teachers more frequently stated that they coach their teachers at endline (rank order #4) than at midline (rank order #5), and at both midline and endline, the least frequently cited response was to ask an inspector for training or assistance.

Table 40: Steps Head Teachers take to help their teachers improve reading instruction - Head Teacher responses from most frequent to least frequent, Mtwara

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Provide materials	Discuss reading in S1 and S2 in teacher meetings
2	Observe their reading lessons	Provide materials
3	Discuss reading in S1 and S2 in teacher meetings	Observe their reading lessons
4	Pair them with a good teacher	Coach them
5	Coach them	Pair them with a good teacher
6	Ask an inspector for training or assistance	Ask an inspector for training or assistance

Practices & activities supporting reading, Mtwara: Teachers’ views

Enumerators asked teachers if any had observed them teach a reading lesson within the school year. A strong majority (n=35, or 88%) said that they had been observed, and of these teachers, 26 (65%) said they had been observed by their Head Teacher. Four teachers mentioned having been observed by a District Education Officer (10%), and three were observed by another teacher (8%)—results similar to those observed at midline. One-third of interviewed teachers (n=13) said that they were observed by school administration personnel at least once per week, while 38% of teachers (n=15) said they were observed at least once a month. These figures are slightly reduced from those observed at midline, where almost half of teachers (n=19) reported being observed by school administration personnel at least once per week. However, while 11 teachers (28%) at midline reported never having been observed, only 3 (8%) reported the same at endline.

When asked who mentors them, 31 teachers (78%) said they had been mentored by school inspectors or Ward educational coordinator on reading practice—an increase from 24 teachers (60%) at midline. Of note, thirty-seven (94%) teachers reported receiving guidance from the Head Teacher to help them teach more effectively, an increase from 28 teachers reporting the same at midline. Nearly all teachers in Mtwara had received training in how to teach reading (n=38, 95%).

When asked how they know when a student is having trouble with reading, most teachers said their students can’t sound out words (65%), can’t make letter sounds (43%)—similar to midline patterns. The third most frequently cited response at endline was “can’t complete reading assignments,” moving up from being the fifth most frequently cited response at midline. While at midline, the proportion of teachers saying that they identify struggling readers as those who do not participate in class or group reading activities was very low relative to other responses (rank order #9), at endline, this response moved up to rank order #6, suggesting that teachers rely on this indicator more often. Although at midline, students who “did not volunteer for reading activities” was a rank order #6 indicator, this was the least frequently cited response at endline (rank order #10).

Table 41: How teachers know students are having trouble learning to read - Teacher responses from most frequent to least frequent, Mtwara

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Can't make letter sounds	Can't sound out words
2	Can't sound out words	Can't make letter sounds
3	Can't string words into sentence	Can't complete reading assignments
4	Skips word when reading	Skips word when reading
5	Can't complete reading assignments	Reads very slowly
6	Does not volunteer for reading activities	Does not participate in class or group reading activities
7	Can't answer questions about text s/he just read	Can't string words into sentence
8	Reads very slowly	Can't answer questions about text s/he just read
9	Does not participate in class or group reading activities	Other
10	Other	Does not volunteer for reading activities

When asked how they help a student who is having trouble learning to read, teachers most frequently reported that they give students extra assignments (63%), pair the student with a good student (53%), and work one-on-one with the student (40%)—patterns similar to midline. Also similar to midline patterns, the teachers least frequently mentioned that they talk to parents, report to the Head Teacher, or take no action at all.

Table 42: How teachers help students who are having trouble learn to read - Teacher responses from most frequent to least frequent, Mtwara

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Pair with a good students	Give extra assignments
2	Give extra assignments	Pair with a good students
3	Work one-on-one with student	Work one-on-one with student
4	Find someone to work with/tutor the students	Recommend outside tutoring
5	Recommend outside tutoring	Find someone to work with/tutor the students
6	Talk to parents	Talk to parents
7	Report to head teacher	Report to head teacher
8	Take no action	Take no action

Teacher meetings: All 40 teachers reported that organized teacher meetings were held at their school, and in 23 of these cases (58%), teachers reported holding teacher group meetings on a monthly basis. Almost one-quarter (n=23, 58%) said teachers met once each term. As at midline, school principals organize the meetings.

At these meetings, teachers most frequently reported discussing training methods (63%), school improvement (55%), and student learning (53%). This is slightly different than midline trends, where the top responses included student discipline and behavior (mentioned 73% of the time). In fact, this issue fell from rank order #3 at midline to rank order #5 at endline. While subject content was mentioned least often at midline, this topic moved to rank order #4 at endline. Administrative issues still remained one of the least frequently cited issues discussed at teacher meetings from midline to endline.

Table 43: Topics discussed at teacher meetings - Teacher responses from most frequent to least frequent, Mtwara

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	School improvement	Training methods
2	Training methods	School improvement
3	Student discipline/behavior	Student learning
4	Student learning	Subject content
5	Classroom management	Student discipline/behavior
6	Administrative issues	Classroom management
7	Subject content	Administrative issues

Practices & activities supporting reading, Mtwara: SMCs’ views

Main functions: When asked what their main functions were, the two most frequent responses cited by SMC members were to improve and monitor school quality (73%), as well as monitor and improve the school environment (68%). While helping to develop school strategy was the most frequently cited response at midline, this dropped to rank order #3 at endline. Also a departure from midline trends, the frequency with which “raising extra funds” was mentioned as a primary function of SMC’s at midline, this was one of the least cited functions at endline. Supporting and encouraging students also moved from rank order #9 at midline to rank order #7 at endline. Allocating funds for new classrooms and helping set staff incentives remained the least cited functions from midline to endline.

Table 44: SMC main functions - Member responses from most frequent to least frequent, Mtwara

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Help develop school strategy	Improve and monitor school quality
2	Improve and monitor school quality	Monitor and improve school environment
3	Monitor and improve school environment	Help develop school strategy
4	Participate in budgeting	Encourage and support the community to enroll school age children
5	Encourage and support the community to enroll school age children	Participate in budgeting
6	Help decide expenditures	Inform/communicate with community
7	Raise extra funds	Support/encourage pupils
8	Inform/communicate with community	Help decide expenditures
9	Support/encourage pupils	Organize extra educational services
10	Organize extra educational services	Raise extra funds
11	To allocate funds for new classrooms	Other
12	Help set staff incentives	To allocate funds for new classrooms
13	Other	Help set staff incentives

When asked what issues had been raised at their most recent SMC meeting, members most frequently cited school building and grounds issues (43%), student attendance (38%), and midday meals (30%)—patterns similar to those observed at midline. Thirty-eight percent of SMC members also mentioned “other” issues, including concerns about agriculture, school infrastructure, exam results, and financing school operations. As with the midline, teacher behavior and absenteeism were least cited by SMC members as issues discussed at their most recent SMC meeting.

Table 45: Issues raised at most recent SMC meeting - Member responses from most frequent to least frequent, Mtwara

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Student attendance	School building/ground
2	School building/ground	Student attendance
3	Midday meals	Other
4	Other	Midday meals
5	School security	School security
6	Students' contribution	Students' contribution
7	Distribution of scholarships/grants	School materials
8	School materials	Distribution of scholarships/grants
9	Pupil/teacher ratios	Distribution of books
10	Other aspects of teacher behavior	Teacher absenteeism
11	Distribution of books	Pupil/teacher ratios
12	Teacher absenteeism	Other aspects of teacher behavior

Support for quality: Finally, when asked what their SMC has done to improve the quality of teaching and learning last year, interviewees most frequently cited at endline two SMC roles: the provision of teaching-learning aids and materials (48%) and the provision of teaching-learning equipment (45%). This was a departure from midline trends, where SMC members most frequently emphasized improving school infrastructure (55%) and textbooks purchases (50%). Indeed, the decrease in focus on textbook purchases (falling from rank order #2 at midline to rank order #5 at endline), in-class teacher assistance by parents (falling from rank order #4 to rank order #7), and professional opportunities for teachers (falling from rank order #6 to rank order #9) as strategies taken by the SMC to improve teaching and learning is evident. Of the 35% of SMC members citing “other” strategies at endline, these primarily included the provision of materials to teachers, but also included the provision of student with materials as well as remedial lessons. Overall, the focus of the SMC this academic year seems to be primarily focused on the improvements in school environment rather than on programs surrounding instruction.

Table 46: SMC roles in promoting quality in school - Member responses from most frequent to least frequent, Mtwara

Rank Order (1 = Most frequently cited)	Response Options	
	Midline	Endline
1	Improved school infrastructure	Teaching-learning aids and materials
2	Textbook purchases	Teacher-learning equipment
3	Teaching-learning aids and materials	Improved school infrastructure
4	In-class teacher assistance by parents	Other
5	Furniture	Textbook purchases
6	Professional development opportunities for teacher	Furniture
7	Extra-curricular educational services and lessons	In-class teacher assistance by parents
8	Sports clubs and events	Sports clubs and events
9	Cultrual clubs and events	Professional development opportunities for teacher
10	Teacher-learning equipment	Extra-curricular educational services and lessons
11	Other	Professional journals or books to teachers
12	Professional journals or books to teachers	Cultrual clubs and events
13	Olympiad coaching and participation	Olympiad coaching and participation

Annex F: Additional discussion of findings: student, teacher, and Head Teacher characteristics

Zanzibar

Table 47: EGRA summary scores correlated with student characteristics, Zanzibar

Variable: EGRA summary score by students who:	Group	N	Mean ‡	SD
Speak the same language at home as at school	Yes	396	29.8	23.8
	No	2	10.8	7.6
Go to a nursery or pre-school	Yes*	213	33.5	24.7
	No	187	25.4	21.8
Were absent from school for more than one week last year	Yes	92	27.7	24.0
	No	305	30.4	23.7
Have an English reading textbook in school	Yes*	73	36.9	27.2
	No	327	28.1	22.6
Have a Kiswahili reading textbook in school	Yes*	164	37.9	25.5
	No	236	24.0	20.7
Get their exercise books check at home by their parents	Yes*	308	31.5	24.3
	No	91	23.8	20.9
Have books or reading materials at home	Yes*	154	38.1	26.0
	No	246	24.5	20.5
Said that his or her mother can read and write	Yes	359	30.3	23.8
	No	38	25.4	23.3
Said that his or her father can read and write	Yes*	361	30.7	23.6
	No	30	17.2	20.3
Said that their parents read them stories at home	Yes*	203	34.3	25.1
	No	196	24.8	21.2

*Difference is statistically significant at $p < .05$

‡ EGRA mean score

Positive correlations were also found with respect to indicators of student wealth (i.e., having a radio, electricity at home, a television, a refrigerator, a toilet inside the house, or a motorcycle at home). Combined, these assets had significant correlation of .219 with overall EGRA scores, suggesting that students with more of these resources at home tended also to have higher overall EGRA scores on average. It should be noted, however, that this correlation is relatively weak.

Teacher characteristics and performance, Zanzibar

Potential correlations of average overall EGRA scores with teacher and classroom characteristics were also analyzed. Table 44 lists those correlations found to be significant and positive.

Table 48: EGRA summary scores correlated with teacher characteristics, Zanzibar

Variable: EGRA summary score by teachers who:	Group	N	Mean ‡	SD
Arrange letter or word tiles to help children learn to manipulate/use the letters of the alphabet	Yes*	16	34.9	14.0
	No	22	24.5	11.6
Have students sound out words	Yes*	13	37.3	15.1
	No	25	24.5	10.5
Use specific reading program materials as criteria to assess whether students achieved the learning outcomes of the reading lesson	Yes*	6	40.3	17.6
	No	32	26.7	11.7
Uses exercise books or slates to teach reading	Yes*	14	35.2	15.1
	No	24	25.2	11.2
Have letter cards available in the classroom	Yes*	16	38.3	14.0
	No	23	23.5	10.1
Guide students to associate words with letters	Yes*	25	33.7	13.4
	No	14	24.2	11.3

*Difference is statistically significant at $p < .05$

‡ EGRA mean score

These findings suggest that on average, students with higher overall EGRA scores also had teachers who had letter cards and exercise books or slates to teach reading. Teachers who were observed to use techniques like arranging letter or word tiles, asking their students to sound out words, and guiding their students to associate words with letters also had students with higher overall EGRA scores, on average. Students tended also to have higher overall EGRA scores if their teacher used specific reading program materials to assess their learning outcomes on a particular reading lesson.

Head Teacher characteristics and performance, Zanzibar

Finally, potential correlations between Head Teacher interview responses and overall EGRA scores were reviewed. Although some positive correlations were found (i.e., Head Teachers who listed management as one of their most important duties, and those who reported their teachers ask students to read to them tended to have students with higher overall EGRA scores), overarching patterns across Head Teacher characteristics and student EGRA scores were not observed.

Mtwara

Table 49: EGRA summary scores correlated with student characteristics, Mtwara

Variable: EGRA summary score by students who:	Group	N	Mean ‡	SD
Speak the same language at home as at school	Yes	292	24.9	21.5
	No	99	27.3	23.6
Go to a nursery or pre-school	Yes*	257	28.1	22.0
	No	143	19.6	20.9
Were absent from school for more than one week last year	Yes	141	21.6	20.7
	No*	246	27.3	22.5
Have an English reading textbook in school	Yes*	34	34.0	24.2
	No	365	24.3	21.6
Have a Kiswahili reading textbook in school	Yes*	101	32.2	23.5
	No	325	28.4	22.7
Get their exercise books check at home by their parents	Yes*	302	26.6	22.0
	No	95	20.5	21.3
Have books or reading materials at home	Yes*	73	31.3	21.0
	No	314	24.0	22.2
Said that his or her mother can read and write	Yes	351	25.6	22.0
	No	42	22.4	20.8
Said that his or her father can read and write	Yes	351	25.4	21.7
	No	35	24.3	24.1
Said that their parents read them stories at home	Yes*	125	29.2	22.9
	No	266	23.4	21.3

*Difference is statistically significant at $p < .05$

‡ EGRA mean score

In Mtwara, students were asked as to whether they had certain assets at home as an indicator of wealth, such as a radio, electricity, a television, etc. No significant correlation was detected between the total number of these assets students reported and their overall EGRA scores.

Teacher characteristics and performance, Mtwara

Potential positive relationships between EGRA scores and teacher characteristics were analyzed in Mtwara, and revealed several significant findings.

Table 50: EGRA summary scores correlated with teacher characteristics, Mtwara

EGRA summary score by teachers who:	Group	N	Mean ‡	SD
Guide students to read printed material or books	Yes*	6	39.9	16.1
	No	34	22.4	10.7
Read out loud to their students	Yes*	38	26.0	12.7
	No	2	7.9	1.9
Use a textbook	Yes*	5	36.4	13.6
	No	34	23.9	12.2
Identify students as having trouble with reading as those who can't answer questions about text they just read	Yes*	4	29.5	20.6
	No	39	24.5	12.4
Teach reading lessons:	Every day	29	23.7	13.1
	3-4 times/week*	7	36.6	11.2
	1-2 times/week	7	18.8	8.1
Receive observation feedback on the content of lessons	Yes*	20	29.4	14.2
	No	23	21.2	11.0
Receive observation feedback on the use of teaching and learning materials	Yes	8	35.4	15.7
	No	35	22.6	11.4

*Difference is statistically significant at $p < .05$

‡ EGRA mean score

These correlations can be interpreted to mean that, on average, students overall EGRA scores tended to be higher if their teacher guided them to read printed materials or books, read out loud, or used a textbook. Students also tended to have higher overall EGRA scores, on average, if their teachers identified struggling readers as those students who could not answer questions about text that he or she just read. Teachers teaching reading lessons 3-4 times per week tended to have students with a higher average EGRA score. In addition, teacher who reported being both observed and receiving feedback on the content of their lessons or on their use of teaching and learning materials also tended to have students with higher overall EGRA scores.

Head Teacher characteristics and performance, Mtwara

The relationships between Head Teacher characteristics and student performance on the EGRA were also explored. Although there were no overarching patterns linking them together, it was found that Head Teachers who worked as a teacher for 6-10 years before taking their current position also tended to have students with higher overall EGRA scores, on average. Head Teachers who communicate often with their teachers as strategy of support, as well as those hosting a reading or book fair at their school also tended to have students with overall EGRA scores. As correlations, these results suggest that while there may be some links between Head Teacher characteristics and EGRA performance, these specific actions of the Head Teacher do not necessarily lead to improvements in reading ability.

Table 51: EGRA summary scores correlated to Head Teacher characteristics, Mtwara

EGRA summary score by Head Teachers who:	Group	N	Mean ‡	SD
Worked as a teacher before taking Head Teacher position for:	0-5 years	10	19.2	5.7
	6-10 years*	15	31.1	11.4
	More than 10 years	15	22.9	15.8
Communicate often with teachers as a way of helping them teach better	Yes*	32	26.5	13.8
	No	8	19.2	7.1
Held reading or book fair at school to support reading	Yes*	8	33.2	12.9
	No	32	23.0	12.4

*Difference is statistically significant at $p < .05$

‡ EGRA mean score