USAID | TAJIKISTAN
FROM THE AMERICAN PEOPLE

USAID Quality Reading Project (QRP): Tajikistan

Early Grade Reading Assessment (EGRA) Baseline Data Analytic Report with Addendum

November 2014

Updated February 2016

A partnership with
American Institutes for Research and Save the Children International

Contract No.: AID-176-C-13-00001-00

USAID Quality Reading Project (QRP): Tajikistan

Early Grade Reading Assessment (EGRA) Baseline Data Analytic Report with Addendum

Submitted by:
American Institutes for Research

15 November 2014

Updated February 2016

This baseline study of early grade reading assessment is made possible by the support of the American people through the United States Agency for International Development (USAID). The contents are the sole responsibility of the American Institutes for Research and Save the Children International and do not necessarily reflect the views of USAID or the United States Government.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

AIR	American Institutes for Research
AOE	Academy of Education
CPD	Continuous Professional Development
DIBELS	Dynamic Indicators for Basic Early Literacy Skills
DRS	District of Republican Subordination
EDI-AOE	Education Development Institute under the Academy of Education
EGRA	Early Grade Reading Assessment
IST	In-service Training
MOES	Ministry of Education and Science
NTC	National Testing Center
QRP	Quality Reading Project
RTMC	Republican Teaching and Methodological Center
RTTI	Republican Teacher Training Institutes
TTI	Teacher Training Institutes
UNICEF	United Nations Children‘s Fund
USAID	United States Agency for International Development

DEFINITIONS

Alphabetic knowledge	Familiarity with the alphabet and with the principle that written spellings systematically represent sounds that can be blended into meaningful words.
Cross-sectional design	A research design that utilizes different groups of people who differ in measure of interest but share other characteristics (e.g., socioeconomic status, educational background, and ethnicity).
Fluency	The bridge between decoding and comprehension. Fluency in word recognition so that the reader is no longer aware of or needs to concentrate on the mental effort of translating letters to sounds and forming sounds into words. At that point, the reader is decoding quickly enough to be able to focus on comprehension.
Fluency analysis	A measure of overall reading competence reflecting the ability to read accurately and quickly (see Fluency).
Grapheme	The most basic unit in an alphabetic written system. Graphemes combine to create phonemes (see Phoneme). A grapheme might be composed of one or more letters or of a letter with a diacritic mark (such as *é* vs. *e* in French).
Longitudinal design	A longitudinal design is a design in which a researcher measures a particular group of people for a long period of time. It can also be defined as a correlation research study in which there are many observations (e.g., about reading ability) being conducted over long periods of time on the same group, or number, of people.
Morpheme	Smallest linguistic unit with meaning. Different from a word, as words can be made up of several morphemes (unbreakable can be divided into un-, break, and -able). There are bound and unbound morphemes. A word is an unbound morpheme, meaning that it can stand-alone. A bound morpheme cannot stand-alone (e.g., prefixes, such as un-).
Metaphonology	See Phonological awareness.
Orthographic	The art of writing words with the proper letters according to usage; spelling.
Performance standards	Knowledge, skills, and abilities that students are expected to demonstrate at their grade level in reading.
Phoneme	The smallest linguistically distinctive unit of sound allowing for differentiation of two words within a specific language (e.g., *top* and *mop* differ by only one phoneme, but the meaning changes).
Phonological awareness	A general appreciation of the sound structure of language, as demonstrated by the awareness of sounds at three levels of structure: syllables, onsets and rhymes, and phonemes.
Phonics	Instructional practices that emphasize how spellings are related to speech sounds in systematic ways.
Reading standards	Knowledge, skills, and abilities that students are expected to learn at their grade level in reading.

FOREWORD

The United States Agency for International Development (USAID) is strategically focused on improving early grade reading outcomes, especially in Central Asia Region. Early grade reading is essential to educational success. Studies like this one are providing strong technical data to Ministries of Education and USAID around the world so that education development partners can work towards policy reforms, making informed decisions and programmatic choices to improve the reading skills lacking in millions of students.

This study is the result of the strong partnership between USAID Central Asia/Tajikistan, the Ministry of Education and Science (MOES), Republic of Tajikistan, and USAID Quality Reading Project. The joint efforts in development of the tools with gender equity made the study a comprehensive success. The study offers great insight into what is working in Tajik and Russian language teaching and learning in Tajikistan, and what is not working and needs to be strengthened. This student-based assessment is an important initiative that will serve as a practical tool for policy makers, school leaders, professors, parents and teachers. All educators and partners working for improving the quality of education in Tajikistan will benefit from this important data and strong analyses. USAID is pleased to support and endorse this Early Grade Reading Assessment (EGRA) in Tajikistan.

Mrs. Mavjuda Nabieva
Education Management Specialist, USAID Central Asia/Tajikistan

---

In recent years, there has been a lot of research and analyses related to the quality and assessment of the education sector conducted by the government, non-governmental organizations, and international agencies that express diverse opinions. The Early Grade Reading Assessment has attracted a lot of attention in the last few years, because the quality of reading in the early grades lays the groundwork for studying all other subjects and learning in general. Identification of the major components of quality of education in early grades is made possible by utilizing a common evaluation tool grounded in competency to form a complete and perfect system of evaluation. Systematic utilization of reading assessment in early grades allows transparent reporting of reading assessment indicators at class, school, district, and republic levels.

This assessment carried out by USAID Quality Reading Project was guided and monitored by the Ministry of Education and Science (MOES), National Testing Center (NTC), Education Academy of Tajikistan, Institute of Education Development of Academy of Education, Teaching and Methodological Center under the MOES, Republican Teacher Training Institute (RTTI), different teacher training institutes, and regional and district education departments. The purpose of the assessment was to gather baseline data on the level of reading skills, reading fluency, comprehension, as well as the identification of the relationships between those skills and different factors in school and at home.

There is no doubt that the assessment carried out for the project purposes using random sampling, as well as other assessments in the future will serve not only for project purposes but also will serve for identifying the level of reading skills on the national level by covering all school and will provide transparent and timely data.

Mrs. Tojiniso Mahmadova,
Deputy Minister, Ministry of Education and Science

ACKNOWLEDGMENTS

The early grade reading assessment (EGRA) has been successfully completed as a result of contributions made by many organizations and individuals. The Ministry of Education and Science (MOES), National Testing Center (NTC), Academy of Education (AOE), Education Development Institute under the Academy of Education (EDI-AOE), Republican Teaching and Methodological Center (RTMC), Republican Teacher Training Institute (RTTI), and various regional Teacher Training Institutes and Education Departments provided vital guidance and oversight during the whole assessment process. In this regard, special gratitude goes to the Minister of Education and Science Mr. Nuriddin Said, and Deputy Minister of Education and Science Ms. Tojinisso Mahmadova for their continuous guidance and support. Also highly acknowledged are Mr. Khurshed Teshaev, Director, and Mr. Jafarov Sabzali, Deputy Director of the National Testing Center for their assistance at the beginning of the project.

This EGRA would have been impossible without the generosity of the American people through the United States Agency for International Development Quality Reading Project. USAID not only provided the funding required but also gave critical input for the implementation process through Mrs. Mavjuda Nabieva, Education Management Specialist at USAID Central Asia/Tajikistan.

The EGRA was implemented with the technical and logistical assistance of American Institutes for Research (AIR) and Save the Children International through the USAID Quality Reading Project. AIR planned and executed the process with the support of staff members in both the Washington, DC-based headquarters and the USAID Quality Reading Project office in Tajikistan. Many thanks goes to Mr. Jerome Mindes, Project Manager based in Washington, DC, for his overall role in supporting the activities, Dr. Abdullah Ferdous for his technical lead and expertise in designing, conducting, and analyzing the assessment, Ms. Barbara Greenwood, Chief of Party, and Ms. Kathryn Fleming, Deputy Chief of Party for technical and administrative leadership, and all the USAID Quality Reading Project staff for their support.

I. EXECUTIVE SUMMARY

The objectives of early grade reading assessment (EGRA) in the Tajik and Russian languages are to set baselines for the United States Agency for International Development (USAID) Quality Reading Project on student reading fluency and comprehension skills, and to determine the relationship of these skills to selected factors in their school and home environment. We utilized an assessment approach that is based on widely-accepted research and best-known practices for competency-based education. This approach supports developing, implementing, and sustaining a system that can be used to (a) determine what students at Grades 1 and 2 know and are able to do with key competencies of the Tajik and Russian reading standards, and (b) inform educational policy, program planning, and decision making. This approach has three core elements that are necessary for improving student reading performance, as follows:

1. Countries must have alignment between academic content standards (i.e., what students are expected to learn at their grade level), performance standards (i.e., how students are expected to perform on the content standards), classroom instruction, and student assessments;
2. The pedagogical factors must be accompanied by support structures, such as political leadership, management systems, and professional training programs; and
3. Schools, districts, and regions must be held accountable through the tracking of student achievement over time.

INSTRUMENT

To achieve more accurate measures of student reading outcomes, the USAID Quality Reading Project utilizes vertically equated common-matrix sampled design for Grades 1 and 2. Each grade and language has a single form at the baseline containing a set of core matrix items unique to grade level and a set of items common to all three grades' instruments, which appear at exactly the same location in each grade's instrument. The common set of items brings Grades 1 and 2's reading outcome measures onto the same reporting scale and also allows for the tracking of students' reading progress from grade to grade. The instrument has nine sections. Four of the nine sections are timed; students are given a maximum of 2 minutes to finish each timed section; however, their ***reading fluency is recorded both at the end of 1 and 2 minutes***. Within 2 minutes, it is expected that students with both lower and higher ability will have had enough time to adequately demonstrate what they know and are able to do. The student instrument, administered orally by a trained administrator in one-on-one sessions with individual students, requires about 25 to 30 minutes per student. The tenth section has been added to include the demographic information of the students. Each section is outlined below.

SECTION 1: LETTER NAME KNOWLEDGE (TIMED)

The purpose of this section is to assess whether students in Grades 1 and 2 know and are able to read aloud both capital and small letters in the Tajik and Russian languages and to determine how fast they can read. A full set of letters is listed in random order. Randomization is used to prevent students from reciting a memorized alphabet.

SECTION 2: LETTER SOUND KNOWLEDGE (NOT TIMED)

The purpose of this section is to examine whether students in Grades 1 and 2 know and are able to demonstrate the sounds of the letters. A list of 10 most frequently used letters identified in primary grade textbooks (Grades 1 to 4) are listed in a row in a clear, large and familiar font. This is not a timed section. Every student is asked to make the sound of each of the letters that is typically taught through a phonic-based approach.

SECTION 3: INITIAL SOUND IDENTIFICATION (NOT TIMED)
The purpose of this section is to examine whether students in Grades 1 and 2 can identify the initial sound of common words used at their grade level. This is a listening exercise. The administrator reads aloud 10 simple words that are grade level appropriate, one word at a time. The student is asked to make the initial sound of each of the words.

SECTION 4: FAMILIAR WORD IDENTIFICATION (TIMED)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to read aloud words that should be familiar at their grade levels. A list of 25 familiar words for Grade 1 and 40 words for Grade 2are selected from primary grade textbooks.

SECTION 5: UNFAMILIAR WORD IDENTIFICATION (TIMED)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to decode unfamiliar words appropriate at their grade levels. A list of 25 unfamiliar words for Grade 1 and 40 words for Grade 2 are selected.

SECTION 6: ORAL VOCABULARY (NOT TIMED)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to understand the meaning of words that should be familiar at their grade levels. This is a listening exercise. The administrator reads 10 words aloud, one word at a time. Students are presented with a set of four pictures for each word and are asked to identify the picture that best matches the word.

SECTION 7A: PASSAGE READING (TIMED)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to read a passage aloud and comprehend it. This section includes one short paragraph, which is around 25 words for Grade 1 and around 40 words for Grade 2.

SECTION 7B: PASSAGE READING COMPREHENSION (NOT TIMED)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to comprehend the passage they have just read. After the student reads the passage aloud, the administrator asks the student four to five simple questions about the passage.

SECTION 8: LISTENING COMPREHENSION (NOT TIMED)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to comprehend the passage they have just heard. This section includes one short paragraph, which is around 25 words for Grade 1 and around 40 words for Grade 2. This is a listening exercise. The test administrator reads a passage aloud to the student only once, slowly (about one word per second), after which he or she asks three to five oral comprehension questions about the passage.

SECTION 9: DICTATION (NOT TIMED)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to write a complete sentence correctly using appropriate formation, size, signs, symbols, and spacing. Students write a dictated sentence on a lined page. The dictation sentence for Grade 1consists of four to five words, and the sentence for grade 2 consists of five to six words.

SAMPLE

The baseline administration of the EGRA was set to be administered in 130 schools randomly selected from the six regions (i.e., Dushanbe, Kurghonteppa, Kulob, Soghd, Zarafhsan, and the District of Republican Subordination [DRS]), with 65 treatment schools receiving the USAID Quality Reading Project interventions and 65 control schools not receiving any USAID Quality Reading Project interventions. The 65 treatment schools (33 large, 29 medium, and 3 small; 47 rural and 18 urban) were randomly selected from 1,721 USAID Quality Reading Project

program schools, and the 65 control schools (33 large, 29 medium, and 3 small; 42 rural and 23 urban) were randomly selected from the remaining non-USAID Quality Reading Project program schools. A systematic sampling procedure was used to select 20 students to be tested from Grades 1 and 2 in each school, for a total of **3,626** students. Out of these, **596** were tested in Russian and **3,030** were tested in Tajik.

DATA ANALYSIS

The EGRA results are primarily reported through students' performance in reading fluency, phonological awareness, comprehension, and dictation. For *reading fluency*, we calculated students' reading fluency at the end of 1 and 2 minutes. If a student attempted to read all words within a given reading section in less than 1 minute or less than 2 minutes, we estimated the corrected reading fluency at the end of minute 1 and 2, respectively, as if the student had used the entire 1 minute or 2 minutes. We also calculated the reading fluency rate for slow readers, i.e., students who took more than 1 minute but less than or equal to 2 minutes to finish each timed section. For *reading comprehension*, we calculated a composite raw score for each student, which included their scores in oral vocabulary knowledge, reading comprehension, and listening comprehension sections. The reliability of internal consistency (Cronbach's alpha) for *reading comprehension* in Tajik and Russian were 0.72–0.75 and 0.77–0.79, respectively. For *phonological awareness and writing,* we calculated a composite score of the letter sound, initial letter sound, and dictation sections. Again, the rates of reliability of internal consistency for that measure in Tajik and Russian were 0.82–0.88 and 0.73–0.85, respectively.

SETTING NATIONAL READING BENCHMARKS

The USAID Quality Reading Project in collaboration with the Ministry of Education and Science (MOES) and other stakeholders set national benchmarks for reading fluency, comprehension, phonological awareness, and dictation. The following four-step, standard-setting procedure was implemented in order to define and establish the reading performance benchmarks:

(1) Developing general descriptions of performance-level categories (e.g., non-satisfactory, satisfactory, good, and excellent);
(2) Developing detailed definitions of performance-level categories to describe the expectations of student performance in each performance-level category, taking into account competency level and grade-level standards;
(3) Establishing cut scores for performance-level categories; and
(4) Recommending national performance benchmarks.

According to MOES's approved benchmarks, students classified as satisfactory and/or above are considered to be meeting the national benchmark. The following table presents approved benchmarks for each grade and skill.

TABLE 1: PERFORMANCE-LEVEL BENCHMARKS BY GRADE AND LANGUAGE SKILL

Language Skill	Grade	Non satisfactory	Satisfactory	Good	Excellent
Reading Fluency: Tajik and Russian (words per minute)	1	1–24 words	25–29 words	30–34 words	35+ words
	2	1–39 words	40–44 words	45–49 words	50+ words
Reading Comprehension: Tajik (marks)	1	0–49%	50–69%	70–84%	85–100%
	2	0–49%	50–59%	60–79%	80–100%
Reading Comprehension: Russian (marks)	1	0–54%	55–77%	78–88%	89–100%
	2	0–52%	53–78%	79–88%	89–100%

Language Skill	Grade	Non satisfactory	Satisfactory	Good	Excellent
Phonological Awareness and Dictation: Tajik (marks)	1	0–62%	63–77%	78–87%	88–100%
	2	0–63%	64–80%	81–88%	89–100%
Phonological Awareness and Dictation: Russian (marks)	1	0–63%	64–71%	72–77%	78–100%
	2	0–65%	66–73%	74–78%	79–100%

For the national *reading fluency* benchmark in Tajik and Russian, students at Grades 1 and 2, must be able to read at least 25 and 40 grade-appropriate words per minute, respectively. For national *reading comprehension*, the benchmarks are defined with respect to the percent of comprehension questions (i.e., comprising reading comprehension, listening comprehension and oral vocabulary knowledge) that students at varying grade levels must answer correctly. To meet national reading comprehension benchmarks in Tajik, students must obtain at least 50 percent marks in Grades 1 and 2. To meet national reading comprehension benchmark in Russian, students must get at least 55 percent marks in Grade 1 and 53 percent in Grade 2. For *phonological awareness and dictation*, the benchmarks are defined by the percentage that students at varying grade levels must answer correctly in the categories of letter sound, initial letter sound, and dictation questions. In Tajik, students must obtain at least 63 percent in Grade 1 and 64 percent in Grade 2. In Russian, students must get 64 percent in Grade 1 and 66 percent in Grade 2. A conjunctive model, in which all students were measured against all national performance benchmarks, was used to find out what percentage of students were meeting benchmarks for certain combinations of skills, such as the combination of reading fluency and comprehension.

In addition to setting national benchmarks, the MOES also set multiple benchmarks relative to the national standards on a four-point performance level categorical scale. The table above presents the benchmarks. For example, if students in Grade 1 obtain 0 percent to 49 percent marks in Tajik reading comprehension, then they would be classified as non-satisfactory; 50 percent to 69 percent marks classify them as satisfactory; 70 percent to 84 percent marks classify them as good; and 85 percent to 100 percent marks classify them as excellent.

RESULTS

1. **NATIONAL READING FLUENCY BENCHMARKS:** Based on the EGRA, more than one-fourth of students in Grade 1 and one-seventh in Grade 2 met national reading fluency benchmarks in Tajik and Russian. The percentage was low because students did not meet the benchmark for unfamiliar words. Students were able to read only familiar words and reading passages containing familiar words faster, which could be due to rote memorization. They struggled noticeably in reading unfamiliar words, suggesting that students have difficulty with decoding. However, it was quite distinct that female students consistently performed better than their male counterparts did in both Tajik and Russian, except for Grade 1 Russian. When it was compared by school type (rural and urban), more students in urban schools met both Tajik and Russian national reading fluency benchmarks than students in rural schools.

2. **NATIONAL READING COMPREHENSION BENCHMARKS:** Although a higher percentage of students met the national reading comprehension benchmark, strong performance in oral vocabulary knowledge compensated for relatively weak performance in reading and listening comprehension. In Tajik, 93 percent in Grade 1 and 78 percent in Grade 2 met national reading comprehension benchmarks; whereas in Russian, 90 percent in Grade 1 and 84 percent in Grade 2 met the benchmarks. Female students performed

better than male students did across all grades and languages, except for Grade 2 Tajik. When compared by school type, students in urban schools performed relatively higher than their counterparts in rural schools.

3. **NATIONAL PHONOLOGICAL AWARENESS AND DICTATION BENCHMARKS:** In general, students performed well in phonological awareness and dictation. More than three-fourths of Tajik language students and four-fifths of Russian language students met national phonological and dictation benchmarks in both grade levels. However, strong performance in letter sound and initial letter sound compensated highly for relatively weak performance in dictation. In Tajik, more than three-fourths of Grade 1 (75 percent male and 83 percent female) and four-fifths of Grade 2 (82 percent male and 84 percent female) met national phonological and dictation benchmarks. In Russian, about four-fifths of Grade 1 (79 percent male and 83 percent female) and eight-ninths of Grade 2 (79 percent male and 87 percent female) met the benchmarks. A higher percentage of urban students met the national benchmarks than their peers in rural schools across all grades and languages.

4. **NATIONAL READING FLUENCY AND COMPREHENSION BENCHMARKS:** A student is said to be meeting the national fluency and comprehension benchmark if he or she meets both fluency and comprehension benchmarks. In Tajik, only 26 percent of students in Grade 1 and 14 percent in Grade 2 met both national benchmarks. In Russian, 36 percent in Grade 1 and 12 percent in Grade 2 met the benchmarks. Again, female students outperformed male students consistently across grade levels, except for Grade 1 Russian. There were no meaningful differences between rural and urban schools except for Grades 1 and 2 Tajik and Grade 2 Russian, for which urban schools performed higher than their counterparts in rural schools did.

5. **NATIONAL LITERACY BENCHMARKS:** A student is said to be meeting the national literacy benchmark if he or she meets the fluency, comprehension, phonological awareness, and dictation benchmarks. More than one-fourth (25 percent in Tajik and 32 percent in Russian) of Grade 1 and one-tenth (13 percent in Tajik and 12 percent in Russian) of Grade 2 met national literacy benchmarks, though a large number of students met comprehension, phonological, and dictation benchmarks. This is due to their poor performance in reading fluency, particularly in reading unfamiliar words. In Tajik, about 28 percent of Grade 1 female students (compared to 22 percent of male students) and 15 percent of Grade 2 female students (compared to 11 percent of male students) met the benchmarks. In Russian, about 33 percent of Grade 1 female students (compared to 32 percent of male students) and 16 percent of Grade 2 female students (compared to 9 percent of male students) met the benchmarks. As shown, the rates of meeting the literacy benchmark for males and females were significantly different in Grades 1 and 2 Tajik and Grade 2 Russian. A higher percentage of urban students met the national benchmarks than their peers in rural schools across all grades and languages, except for Grade 1 Russian.

6. **READING COMPREHENSION PERFORMANCE LEVEL BENCHMARKS:** When student performance in reading comprehension was separated into performance level categories, it was revealed that in Tajik, 7 percent of students in Grade 1 were classified as non-satisfactory, 28 percent as satisfactory, 40 percent as good, and 25 percent as excellent. In Grade 2, 22 percent of students were classified as non-satisfactory, 19 percent as satisfactory, 39 percent as good, and 20 percent as excellent. On the other hand, in Russian, 10 percent of the students in Grade 1 were classified as non-satisfactory, 40 percent as satisfactory, 28 percent as good, and 22 percent as excellent. In Grade 2, 16 percent of

students were classified as non-satisfactory, 46 percent as satisfactory, 18 percent as good, and 20 percent as excellent.

7. **PHONOLOGICAL AWARENESS AND DICTATION PERFORMANCE LEVEL BENCHMARKS:** In Tajik, the majority of the students were classified as excellent in their respective grade levels, across grades 1 and 2. In Grade 1, 21 percent of the students were classified as non-satisfactory, 19 percent as satisfactory, 20 percent as good, and 40 percent as excellent. In Grade 2, 17 percent of the students were classified as non-satisfactory, 19 percent as satisfactory, 18 percent as good, and 46 percent as excellent. On the other hand, in Grade 1 Russian, about 62 percent in Grade 1 and 64 percent in Grade 2 were classified as excellent. The remaining students in Grade 1 were classified as 19 percent non-satisfactory, 9 percent satisfactory, and 10 percent good. In Grade 2, 12 percent of the students were classified as non-satisfactory, the same percentages were classified as satisfactory and good.

8. **PERFORMANCE IN SUBTASKS:** Students' performances in various subtasks of EGRA are presented in the following sections.
   - *Letter naming knowledge:* The majority of the students were able to read Tajik and Russian letter names. In Tajik, students in Grades 1 and 2 read 58 and 72 letters per minute, respectively. About 39 percent of Grade 1 and 62 percent of Grade 2 students attempted to finish the entire section in less than 1 minute, and only 10 students (eight in Grade 1 and two in Grade 2) got a score of zero. In Russian, students in Grades 1 and 2 read 65 and 72 letters per minute, respectively. About 41 percent of Grade 1 and 74 percent of Grade 2 students attempted the entire section in less than 1 minute, and no students got a score of zero. The students in Grade 2 Tajik and Grades 1 and 2 Russian who took more than 1 minute to finish the entire section read much faster in the second minute than in the first minute. Overall, female students performed better than male students did in both Grades 1 and 2 Tajik, and Grade 2 Russian; the differences were only statistically significant for Grades 1 and 2 Tajik.
   - *Letter sound knowledge:* Students in Grades 1 and 2 obtained moderately high scores, irrespective of the languages. About one-fourth to one-third of the students in Tajik and more than two-fifths of the students in Russian obtained perfect scores in letter sound knowledge. In Tajik, 5 in Grade 1 and 16 in Grade 2 obtained zero score. In Russian, only one student in Grade 1 obtained zero score. No meaningful difference in scores was observed between male and female students.
   - *Initial letter sound knowledge:* A similar pattern was detected for initial letter sound knowledge as for letter sound knowledge. In Tajik, students in Grades 1 and 2 obtained average scores of 7.5 (7.2 for males, 7.7 for females) and 7.6 (7.4 for males, 7.7 for females) out of possible scores of 10, respectively. However, a large number of students both in Grades 1 (52 students) and 2 (81 students) obtained a score of zero. In Russian, students in Grades 1 and 2 obtained average scores of 8.1 (7.8 for males, 8.4 for females) and 8.5 (8.5 for males, 8.4 for females) out of possible scores of 10, respectively. There were only two students in Grade 1 and one student in Grade 2 who obtained a score of zero. When compared by gender, female students in Grades 1 and 2 Tajik performed statistically significantly higher than their peers did.
   - *Familiar word identification:* In Tajik, students in Grades 1 and 2 read 27 and 37 familiar words correctly in 1 minute, respectively. About 55 percent of Grade 1 and 48 percent of Grade 2 students attempted the entire section in less than 1 minute; 13 students in Grade 1 and 11 students in Grade 2 got a score of zero. In Russian, students in Grades 1 and 2 read 35 and 54 familiar words correctly in 1 minute, respectively.

About 75 percent of Grade 1 and 76 percent of Grade 2 students attempted the entire section in less than 1 minute; three students in Grade 1 and two students in Grade 2 got a score of zero. Like in other subtasks, female students outperformed male students in both grades in Tajik and Grade 2 Russian, but the differences were only statistically significant for Grade 2 Tajik.

- *Unfamiliar word identification:* Students' overall performance in unfamiliar word identification was very poor as compared to their performance in familiar word identification. In Tajik, students at Grades 1 and 2 read about 25 and 27 unfamiliar words per minute, respectively. About 54 percent of Grade 1 and 24 percent of Grade 2 students attempted the entire unfamiliar word section in less than 1 minute; 13 students in Grade 1 and 16 students in Grade 2 got a score of zero. In Russian, students in Grade 1 and 2 read 21 and 29 unfamiliar words in one minute, respectively. About one-half of Grade 1 and one-fourth of Grade 2 students attempted the entire unfamiliar word section in less than 1 minute; six students in Grade 1 and four students in Grade 2 got a score of zero. In both Tajik and Russian, students who took more than 1 minute but less than or equal to 2 minutes read at a faster pace in the second minute than in the first minute. When compared by gender, female students performed statistically significantly better than male students did for all grades in Tajik and for Grade 2 in Russian.
- *Oral vocabulary knowledge:* Overall, students in Grades 1 and 2 performed very similarly in both Tajik and Russian. In Tajik, students in Grades 1 and 2 obtained average scores of 8.9 and 8.0, respectively, out of a possible score of 10. Only two students in Grade 1 and three students in Grade 2 obtained zero score. About 45 percent of students in Grade 1 and 19 percent in Grade 2 obtained perfect scores (i.e., 10 out of 10) in the oral vocabulary knowledge section. In Russian, students in Grades 1 and 2 received average scores of 9.5, 8.8, and 8.8 out of a possible score of 10, respectively. No students obtained a score of zero, while about 77 percent of the students in Grade 1 and 43 percent in Grade 2 obtained perfect scores (i.e., 10 out of 10). When compared by gender, no meaningful differences were observed between the scores of male and female students for both Tajik and Russian.
- *Reading passage:* In Tajik, students in Grades 1 and 2 read at a rate of 21 and 36 words per minute, respectively. About 42 percent of Grade 1 and 33 percent of Grade 2 students attempted to finish the entire reading passage section in less than 1 minute; a total of 19 Grade 1 and 15 Grade 2 students obtained scores of zero. In Russian, students in Grades 1 and 2 read 31 and 42 words in 1 minute, respectively. A total of 72 percent in Grade 1 and 38 percent in Grade 2 attempted to finish the entire reading passage section in less than 1 minute, and only two students in Grade 1 and one in Grade 2 obtained a score of zero. When compared by gender, female students performed better than male students did in Tajik in all grades and in Russian in Grade 2. However, the differences between male and female students were only significant for Grade 2 Tajik.
- *Reading comprehension:* Students' overall performance in reading comprehension was moderate. In Tajik, students in Grades 1 and 2 obtained average scores of 2.1 out of a possible score of 5 and 2.0 out of 5, respectively. About 19 percent of Grade 1 and 28 percent of Grade 2 students obtained a score of zero; whereas 6 percent of Grade 1 and 9 percent Grade 2 students secured perfect scores (i.e., 5 out of 5). In Russian, students in Grades 1 and 2 obtained average scores of 1.3 out of a possible score of 4 and 2.3 out of 5, respectively. A total of 36 percent of Grade 1 and 21 percent of Grade 2 students obtained scores of zero in Russian reading comprehension; whereas 7 percent of Grade 1 and 12 percent of Grade 2 students received perfect scores.

- *Listening comprehension:* Students performed relatively better in listening comprehension questions than in reading comprehension questions. In Tajik, students in Grades 1 and 2 obtained an average score of 3.2 (3.2 for males, 3.3 for females) out of possible score of 5 and 2.5 (2.5 for males, 2.4 for females) out of 5, respectively. About 5 percent of Grade 1 and 9 percent of Grade 2 students obtained scores of zero, while 27 percent of Grade 1 and 11 percent of Grade 2 students secured perfect scores in their respective listening comprehension sections. In Russian, students in Grades 1 and 2 obtained average scores of 2.3 (2.2 for males, 2.3 for females) out of a possible score of 4 and 2.1 (2.0 for males, 2.1 for females) out of 4, respectively. A total of 19 percent of Grade 1 and 15 percent of Grade 2 students obtained a score of zero, while 29 percent of Grade 1 and 15 percent of Grade 2 students received perfect scores in their respective listening comprehension sections. Female students for Grade 2 Tajik tested statistically significantly better than male students did.
- *Dictation:* Students' overall performance in dictation was good. In Tajik, students at Grades 1 and 2 obtained average scores of 8.7 (73 percent) out of a possible score of 12 and 13.3 (83 percent) out of 16, respectively. About 9 percent of Grade and 2 percent of Grade 2 students obtained a score of zero in dictation. More than one-third of Grade 1 (31 percent) and Grade 2 (34 percent) students secured perfect scores. In Russian, students in Grades 1 and 2 obtained average scores of 11.4 (71 percent) out of a possible score of 16 and 13.4 (74 percent) out of 18, respectively. Only 1 percent of each Grade 1 and 2 students obtained a score of zero in writing. Nine percent of Grade 1 and 15 percent of Grade 2 students received perfect scores in their respective writing sections. For both Grades in Tajik and Grade 2 in Russian, the female students tested statistically significantly better than male students did.

RECOMMENDATIONS

1. Ample research demonstrates that teacher subject knowledge and participation in in-service training (IST) have a positive impact on student performance. Therefore, teachers at both pre-service and in-service levels need to be introduced to proven research-based methods and strategies for teaching languages to students.
2. Research also shows that teachers need long-term guidance and support in order to sufficiently understand and utilize information received through IST programs. Additional in-service and continuous professional development (CPD) activities should take place to ensure teachers continue to develop their skills in utilizing these methodologies in the classroom on a daily basis. Furthermore, these strategies and methods should be incorporated into the pre-service curriculum at every teacher training institute (TTI) so that all graduating teachers are equipped with the methodologies they will need to be successful in teaching literacy skills to their students.
3. Students have difficulty with reading fluency and comprehension. That suggests students do not have adequate opportunity to practice reading, due to a lack of learning materials. Significant efforts need to be made to procure, develop, and distribute quality reading materials and teaching aids so that both students and teachers have easy access to the materials to increase teaching and learning literacy outcomes.
4. This baseline EGRA was aligned with reading standards and performance benchmarks for measuring student progress. Therefore, it is expected that future EGRAs must employ a standards-based approach and must make a strong connection with the baseline EGRA when constructing future EGRA tests, analyzing the data, and reporting the progress.
5. Simple formative assessment tools for literacy learning should be developed and incorporated into pre-service, in-service, and CPD training programs so teachers can better understand student learning in the classroom and adjust their lesson planning accordingly.

II. BACKGROUND AND CONTEXT

INTRODUCTION

The EGRA report has been divided into five sections. The first section provides background information about the USAID Quality Reading Project, including the status of Tajik and Russian languages in Tajikistan and the objectives of EGRA in Tajik and Russian. The second section explains the standards-based reading approach and standards. The third section describes the methodology of the EGRA, including an overview of the EGRA instruments and design; the process used to develop and pilot the instruments; sampling procedures; the procedures of test administration; the process of setting performance benchmark procedures for reading fluency, comprehension, phonological awareness, and dictation; and the process of data analysis. The fourth section provides the findings of the baseline EGRA. The fifth section provides recommendations based on the findings.

USAID QUALITY READING PROJECT

The USAID Quality Reading Project works to improve reading skills among primary grade students in Kyrgyzstan and Tajikistan within four years (June 2013 to September 2017). By drawing on existing structures in both countries, the USAID Quality Reading Project is building capacity from the national level down to the classroom level and supporting the common goal of improving student reading skills. The USAID Quality Reading Project is working with the MOES of both countries to create a set of measurable, uniform standards for teachers, students, and other education officials. Based on these standards, the USAID Quality Reading Project's major activities include teacher training based on reading skills, reading material dissemination, community activities, and increasing government capacity around primary grade reading education. In Tajikistan, the USAID Quality Reading Project is rolling out the activities in three phases (also called training cohorts): phase 1 includes Dushanbe, Kulob, Kurgonteppa, and some part of the Sughd region, phase 2 includes the remaining part of Sughd and Zarafshon, and phase 3 includes the DRS region.

Some of the key goals of the USAID Quality Reading Project in Tajikistan include the following:

- Reinforce the importance of teaching and learning reading in 1,721 Tajik schools, and reach more than 10,000 Tajik teachers;
- Establish grade-level minimum standards for key early literacy skills, such as for phonics and reading comprehension;
- Strengthen national systems to administer standardized and classroom-based assessments in order to track student learning and use data for decision making;
- Assist local institutions and communities to develop and utilize grade-level reading materials; and
- Deliver services to more than 400,000 students in Tajikistan (60 percent of the Tajik and Russian primary school population).

TAJIKISTAN CONTEXT

A number of government assessments and donor-supported interventions each separately revealed that reading levels in Tajikistan are low. In 2008, for example, the government conducted a World Bank–sponsored National Assessment that showed that literacy and numeracy skills of Grade 4 students were below acceptable levels. More recently in 2010, the USAID Quality Learning Project (QLP) conducted a baseline study for Grades 4 and 7 students in Tajikistan and Kyrgyzstan that further validated the reading trends outlined above.

In partnership with Tajik governments, USAID supported an EGRA in 2012 for students in Grades 2, 3, and 4. A sample of more than 4,000 students were tested and a complementary qualitative review of current teaching practices intended to "scratch the surface" of reading pedagogy was undertaken. The tests were administered in the Tajik and Russian languages. Although students who took the EGRA test in Tajikistan have some of the early skills necessary for basic literacy, including letter recognition, the students performed below national and international benchmarks in recognizing phonemes, word decoding, and reading fluency. Students were not reaching sufficient fluency levels to transition to reading comprehension, particularly in Grades 3 and 4. When compared to Dynamic Indicators for Basic Early Literacy Skills (DIBELS), a majority of the students in Grade 4 could not read at their grade level. Students also struggled with inferential questions, indicating low levels of critical thinking and reading comprehension. These difficulties in comprehension also reflected poor fluency, as students were more focused on reading the words and were less able to comprehend what they were reading. It was also evident that almost 41 percent of the students across all grades were not able to read at national standards for reading fluency, the only available benchmark of reading skills. The outcomes in reading comprehension indicated that students performed better on literal questions than inferential questions, indicating difficulties with reading comprehension and critical understanding of text, an indicator of functional literacy.

OBJECTIVES OF EGRA

The objectives of the EGRA in the Tajik and Russian languages are to set baselines for the USAID Quality Reading Project on student reading fluency and comprehension skills and to determine the relationship of these skills to selected factors in their school and home environment. The assessment also provides valid and reliable baseline data on student reading and learning outcomes in the Tajik and Russian languages for Grades 1 and 2,[1] disaggregated at the national level.

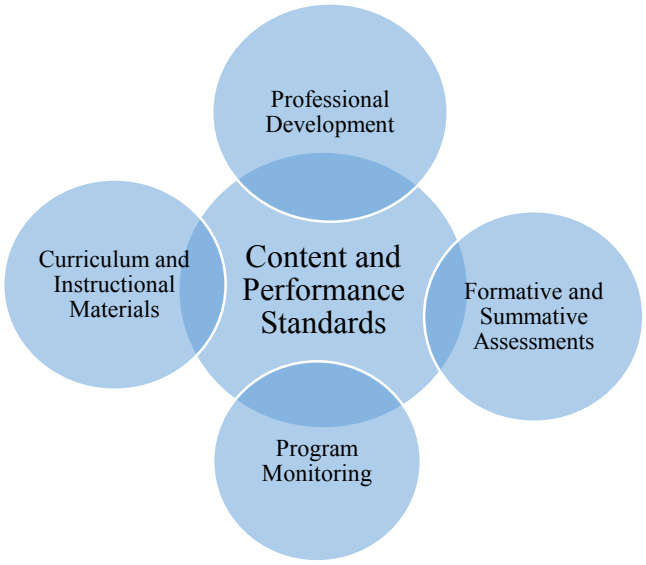
[1] It would have been most feasible and optimized for results if EGRA was administered only to students in Grades 2 and 4 to examine what students were able to do in reading after 2 years of schooling and at the end of the project cycle. The the USAID Quality Reading Project included Grade 1 in the baseline to track the same students over the life of the project and to study their reading learning trajectories. Collecting reading performance data from Grade 3 students would not provide any additional information needed to make reading policy intervention decisions.

III. DESCRIPTION OF APPROACH

STANDARDS-BASED RESEARCH

The USAID Quality Reading Project utilized an approach for the Tajik and Russian EGRA that was based on widely accepted research and best-known practices for standards-based (also called competency-based) education. This approach supports the development, implementation, and sustainability of a system that can be used to (a) determine what students in Grades 1 and 2 know and are able to do with key competencies of the Tajik and Russian reading standards, and (b) inform educational policy, program planning, and decision-making. This approach has been successfully applied in many developed countries (e.g., Canada, China, Finland, Netherlands, and the United States) and developing countries (e.g., Egypt, Ethiopia, Honduras, Namibia, and Pakistan).

FIGURE 1: ELEMENTS OF AN ALIGNED STANDARDS-BASED EDUCATION SYSTEM



The figure above shows the core elements that are necessary for improved student reading performance (Briars & Resnick, 2000; Linn, 2001; McKinsey & Company, 2007), which is defined as follows:

1. Countries must have alignment between academic content standards (i.e., what students are expected to learn at their grade level) and performance standards (i.e., how well students are expected to perform on the content standards), classroom instruction, and student assessments.
2. The pedagogical factors must be accompanied by support structures, such as political leadership, management systems, and professional training programs.
3. Schools, districts, and regions must be held accountable through tracking student achievement over time. Tajikistan has already made significant progress toward the establishment of a standards-based education system. With assistance from the USAID Quality Reading Project, it has recently developed national reading content standards and reading performance benchmarks, aligned in-service teacher training reading materials (also called IST package) and EGRA with these reading standards, and provided training to teachers on teaching reading and comprehension.

NATIONAL READING STANDARDS

The first draft of "Mother Tongue" Subject Standards, which included all language skills (i.e., reading, writing, speaking, and listening), was developed by the MOES under the Fast Track Initiative (FTI) -3. The USAID Quality Reading Project initiated the revision of these standards and involved local and international experts to revise them to meet competency-based learning approaches. As a result, four standards documents—Standards and Syllabus for the Mother Tongue Subject for Grades 1, 2, 3, and 4—were revised and presented to the MOES for curriculum board review. The curriculum board feedback and comments were used accordingly to adjust the documents. The development and revision stages were as follows:

1) Primary education experts, both local and international, helped to revise draft standards. This initiative was focused on the revision of each standard document by grade and subject learning strands (i.e., reading, writing, speaking, listening, grammar, and life and learning skills) to ensure the format and quality of competencies described in these documents.
2) Standards were submitted to the Academy of Education, Education Development Institute, Republican Teacher Training Institute, and the Republican Teaching and Methodological Center for review and comments, which were incorporated where possible.
3) Revised standards were submitted to the Education Development Institute for final review and approval.

Tajikistan Primary Standards were developed based on the approved Conceptual Framework, and each document (Grades 1, 2, 3, and 4) consists of the following sections:

1) Introduction
2) General conditions
3) Application
4) Legislative and regulatory documents
5) General description of teaching the Mother Tongue in primary grades
6) Status of the Mother Tongue subject in primary education
7) Aims and objectives of the Mother Tongue subject in primary grades
8) Standards and syllabi of the Mother Tongue subject in primary grades
9) Requirements for knowledge, skills, and competencies of primary grade students in the Mother Tongue subject
10) Requirements for the compulsory contents of the Mother Tongue subject
11) Requirements for the volume of reading texts
12) Syllabus of the Mother Tongue subject
13) Technology of teaching the Mother Tongue subject
14) Calendar plan for the Mother Tongue subject
15) Requirements for the Mother Tongue subject lesson plans
16) Assessment criteria for students' learning achievements in the Mother Tongue subject
17) Assessment criteria for Grade 4 students' knowledge, skills and abilities in the Mother Tongue subject
18) Requirements to the organizing and implementation of the learning process for the Mother Tongue subject

Language competencies are presented in the section, Requirements for knowledge, skills and competencies of primary grade students in the Mother Tongue, which defines the following skills and sub-skills:

1) Pre-reading and writing (only for Grade 1 students)
2) Phonemic awareness
3) Letter knowledge and decoding
4) Vocabulary
5) Reading fluency
6) Reading comprehension
7) Writing
8) Oral speech development—communication (i.e., speaking and listening)
9) Life and generic learning skills

Due to the delays in making a policy decision regarding splitting the Mother Tongue subject into two separate subjects (Language and Reading), the approval and introduction of the Primary Education Standards package (including other subjects) at the school level has been postponed. However, there is a strong will to make this decision soon. Once an official policy decision has been made, the USAID Quality Reading Project will adjust the materials accordingly and work with the MOES to support their rollout.

IV. PROCEDURE

RESEARCH DESIGN

For the USAID Quality Reading Project, we collected relevant data corresponding to student reading learning outcomes at the beginning (before any significant implementation occurs) to establish a baseline. We will collect student reading progress in the middle of the program to monitor interim and at the end to evaluate final changes in reading outcomes. We will use reliable, valid, and fair tools appropriately aligned with reading learning outcomes and project interventions.

TABLE 2: CROSS-SECTIONAL AND LONGITUDINAL DESIGN

Cohort	2014	2015	2016	2017
Cross Sectional Design				
1	G2		G3	G2
				G4
2 & 3	G2		G2	G2
			G4	G4
Longitudinal Design				
1	G1			
			G3	
				G4
2 & 3	G2			
			G4	

We have utilized cross-sectional and longitudinal research designs for the EGRA study. For the cross-sectional design (training cohorts 1, 2, and 3) covering Grade 2, the baseline group will be compared to different groups of students at the same schools and grade levels in subsequent years. The hypothesis is that the scores will increase from the baseline to the post-tests due to the positive effects of the project interventions on literacy.

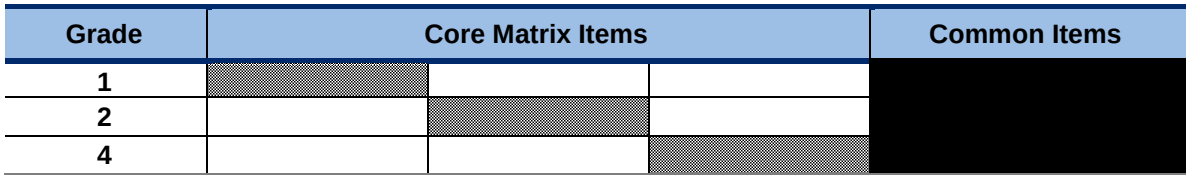
For longitudinal design, the same students' reading performance (in training cohort 1) in Grade 1 in 2014 will be compared with their performances in Grade 2 in 2015 and in Grade 4 in 2017. A key feature of this design is that student reading performances in Grades 1, 2, and 4 are tracked and reported on the same measurement scale. The same student's reading performance (in training cohort 2 and 3) in Grade 2 in 2014 will be compared with their performance in Grade 4 in 2016. The process of bringing Grades 1, 2, and 4's reading performances onto the same scale is called vertical scaling.

INSTRUMENT DEVELOPMENT

ASSESSMENT DESIGN

To achieve a more accurate measure of student reading outcomes, the USAID Quality Reading Project utilizes a vertically equated common-matrix sampled design for Grades 1, 2, and 4 (Figure 2). This means that there is a single form for each grade and language in the baseline, containing a set of core matrix items unique to each grade level, and a set of common items that appear at the exact same locations in all grades' instruments. The common set of items brings Grades 1 and 2 reading outcome measures in the baseline onto the same reporting scale and allows tracking students' reading progress from grade to grade.

FIGURE 2: THE VISUAL REPRESENTATION OF THE VERTICALLY EQUATED ASSESSMENT DESIGN



To measure student progress accurately, cross-sectionally (i.e., different cohort of students at the same schools in different years), and longitudinally (i.e., same students in different years) without the tests being exposed, familiarized, or memorized, the USAID Quality Reading Project will use different sets of instruments in the baseline, at the mid-term, and at the end of the project. However, the instruments across different years will be linked through a set of common items as well. Therefore, a total of two three-set instruments will be developed for EGRA; one set for Tajik and one set for Russian, with each set consisting of three instruments, one each for Grades 1 2, and 4.

The EGRA instrument has nine sections and a background information section as described below. Four of the nine sections are timed; students are given a maximum of 2 minutes to finish each timed section. However, their ***reading fluency is recorded at the end of both 1 and 2 minutes***. Within 2 minutes, it is expected that both students with lower and higher ability would be able to demonstrate adequately what they know and are able to do. The student instrument, administered orally by a trained administrator in one-on-one sessions with the individual student, requires about 25 minutes for each student. The tenth section is added to include background information on the students. Each section is outlined below.

Section 1: Letter Name Knowledge (Timed)
The purpose of this section is to assess whether students in Grades 1 and 2 know and are able to read aloud both capital and small letters in Tajik and Russian languages and to determine how fast they can read. A full set of letters is listed in random order. Randomization is used to prevent students from reciting a memorized alphabet.

Section 2: Letter Sound Knowledge (Not Timed)
The purpose of this section is to examine whether students in Grades 1 and 2 know and are able to sound out the letters. A list of 10 most frequently used letters identified in primary grade textbooks (Grades 1 to 4) are listed in a row in a clear, large, and familiar font. This is not a timed section. Every student is asked to make the sound of the letters that are typically taught in phonic-based approaches.

Section 3: Initial Sound Identification (Not Timed)
The purpose of this section is to examine whether students in Grades 1 and 2 can identify the initial sound of common words used at their grade level. This is a listening exercise. The administrator reads aloud 10 simple words appropriate at grade level, one word at a time. The student is asked to make the initial sound of each of the words.

Section 4: Familiar Word Identification (Timed)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to read aloud familiar words at their grade levels. A list of 25 familiar words for Grade 1 and 40 familiar words for Grade 2 are selected from primary grade textbooks.

Section 5: Unfamiliar Word Identification (Timed)
The purpose of this section is to examine whether students in Grades 1 and 2are able to decode unfamiliar words appropriate to their grade levels. A list of 25 unfamiliar words for Grade 1 and 40 unfamiliar words for Grade 2 are selected.

Section 6: Oral Vocabulary (Not Timed)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to understand the meanings of familiar words at their grade levels. This is a listening exercise. The administrator reads aloud 10 words, one word at a time. Students are presented with a set of four pictures for each word read and asked to identify the picture that best matches the word.

Section 7a: Passage Reading (Timed)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to read aloud a passage with comprehension. This section includes one short paragraph, which is around 25 words for Grade 1 and around 40 words for Grade 2.

Section 7b: Passage Reading Comprehension (Not Timed)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to comprehend the passage they just read. After the student read the passage aloud, the administrator asks the student four to five simple questions about the passage.

Section 8: Listening comprehension (Not Timed)
The purpose of this section is also to examine whether students in Grades 1 and 2 are able to comprehend the passage they just heard. This section includes one short paragraph, which is around 25 words for Grade 1 and around 40 words for Grade 2. This is a listening exercise. The test administrator reads aloud a passage to the student only once, slowly, about one word per second. After the administrator reads the passage, he or she asks three to five oral comprehension questions about the passage.

Section 9: Dictation (Not Timed)
The purpose of this section is to examine whether students in Grades 1 and 2 are able to write a complete sentence correctly, using appropriate formation, size, signs, symbols, and spacing. Student will write the dictation sentence on a lined page. The dictation sentence for Grades 1 and 2 will consists of four to five words and five to six words, respectively.

Section 10: Student Background Questions (Not Timed)
The purpose of this section is to collect more information about student background (e.g., home language, reading culture at home, reading materials and resources available at home) so that the relationship between student performance in reading and factors influencing reading outcomes can be explained.

EGRA ITEM DEVELOPMENT

The USAID Quality Reading Project employed the EGRA instruments in Tajik and Russian, adapted for the Tajikistan context, after reviewing Tajik and Russian primary grade reading standards. The reading standards were the basis for the development of test items. The USAID Quality Reading Project conducted a four-day item development workshop for Tajik and Russian languages concurrently from January 28 to 31, 2014. A total of 37 participants including teachers, language and reading experts, psychologist, and standard developers attended the workshop. Participants were provided a thorough training on item development principles and procedure before they were involved in item writing activity. Items were strongly aligned with national reading standards and had varied cognitive complexity (i.e., knowledge, comprehension, and application) and difficulty levels (i.e., easy, moderate, and hard). Because it was evident from previous USAID (2012) studies that students had difficulty with reading comprehension and critical understanding of text (particularly with inferential questions), participants were given special instruction on writing inferential questions related to reading and listening passages.

Participants were provided with lists of most frequently used words for Grades 1 and 2, constructed using respective grade level textbooks; they used those lists for choosing letters and words for letter naming, letter sound, familiar words, and reading passage sections of the EGRA instruments. Moreover, they received more than 300 pictures (developed under USAID's Facilitating Reading Acquisition in Multilingual Environments (FRAME)/India Project) to develop oral vocabulary questions for all three grades. At the end of the workshop, the USAID Quality Reading Project had an adequate number of items necessary for assembling at least four pilot EGRA instruments in Tajik and Russian.

PILOT TESTING AND ASSEMBLING OF BASELINE INSTRUMENTS

Following the item development workshop, the USAID Quality Reading Project administered the EGRA pilot test from March 31 to April 3, 2014, to a sample of 800 students at 20 purposively selected schools (10 Tajik schools and 10 Russian schools) located in the four regions. During the data collection, 20 teams of three people each were deployed to pilot test the 12 instruments (i.e., three instruments x two grades x two languages). Each team randomly selected 40 students (20 Grade 1 and 20 Grade 2) from each school and tested them over three days. After we administered the pilot tests, we used an image scanning technology, a cost-effective and sustainable system for speedy, reliable, and accurate data capturing. We then analyzed the pilot test data to examine psychometric properties of the items (e.g., item difficulty[2] and discrimination[3] based on classical and item response theory). In addition, each item was reviewed and analyzed to ensure fairness and balance based on gender, ethnicity, religion, and other factors.

The results of the pilot data analysis were the basis for assembling the baseline instruments to be included in the tests, which were items with acceptable psychometric properties and with varied cognitive complexity and difficulty levels. Before finalizing the instruments, we also looked at how the sequencing of various sections were set to be appropriate and logical, and the common items in three instruments (Grades 1 and 2) were placed identically so that common items would not perform differently in different instruments.

SAMPLING

The baseline administration of the EGRA was set to be administered in 130 randomly selected schools drawn from the six regions (65 pilot schools receiving the USAID Quality Reading Project interventions, and 65 control schools not receiving any USAID Quality Reading Project interventions). The 65 pilot schools (33 large, 29 medium, and 3 small; 47 rural and 18 urban) were selected randomly from the 1,721 USAID Quality Reading Project program schools, and the 65 control schools (33 large, 29 medium, and 3 small; 42 rural and 23 urban) were selected randomly from the remaining non-USAID Quality Reading Project program schools. The distribution of sample of schools is presented in Table 3.

[2] The item difficulty (also called *p*-value) is defined as the percent of students tested who answered the question correctly and is therefore interpreted on the 0–100 scale. If the *p*-value of an item is .60, then it indicates that 60 percent of the students who participated in the test have answered the item correctly.

[3] The item discrimination, defined as how well the item distinguishes between the more knowledgeable and the less knowledgeable students, is also described on the 0–100 scale. If an item has a discrimination value of .35, then it indicates that the top one-third higher performing students have 35 percent higher chances of getting the item right than the bottom one-third students.

TABLE 3: DISTRIBUTION OF SCHOOLS BY REGION

Region	Treatment School	Control School	Total
DRS	20	22	42
Dushanbe	4	4	8
Kulob	8	8	16
Kurgonteppa	13	13	26
Sugd	15	14	29
Zarafshon	5	4	9
Total	**65**	**65**	**130**

A systematic sampling procedure was utilized to select 20 students to be tested from each of the Grades 1 and 2 from every school, for a total of 3,626 students. Out of which, we tested 596 students in Russian and 3,030 in Tajik.

EGRA ADMINISTRATION

TRAINING OF EGRA ADMINISTRATORS

The USAID Quality Reading Project trained supervisors and test administrators on one-to-one EGRA administration procedures and how to record students' oral responses into scannable forms by shading in bubbles for students' correct responses, number of letters and words reached within the first and second minutes, and other demographic information. We conducted a training of test administrators through a two-step cascading process: international consultants conducted a four-day training for supervisors and project staff, and then supervisors conducted a five-day training in their respective regions for test administrators. These training workshops trained a total of 12 supervisors and 157 test administrators.

During the training, the test administrators practiced school-level sampling and test administration procedures. In preparation for the various possible scenarios for school environments in Tajikistan, test administrators practiced drawing the student sample by completing the sample selection forms and calculating the sample intervals to select the necessary 20 students for each grade. The final part of the test administrators' training audited the roles and responsibilities of the test administrators, team supervisors, and the USAID Quality Reading Project office, as explicitly described in the administration manual.

EGRA ADMINISTRATION AND MONITORING

The 157 EGRA test administrators were deployed in 35 teams to collect data in the 130 schools. Each team of four of five administrators was responsible for administering the assessment in five schools. Data collection commenced from May 15 to 31, 2014. During the data collection, the test administrators were instructed to thoroughly check the instruments each school completed before returning them to their regional supervisor each evening; and regional supervisors were instructed to review the instruments thoroughly before signing off.

Each cohort of EGRA administrators was deployed following their respective round of training, resulting in a staged rollout of the EGRA administration to ensure timely completion. Academy of Education (AOE) coordinators, MOES representatives, and the USAID Quality Reading Project staff were mobilized to conduct monitoring visits of EGRA administration to ensure proper administration of the assessment and to support troubleshooting as necessary. Although it was planned to administer EGRA to a sample of 6,182 students, we managed to administer to 6,050 students. In some of the sample rural schools, there were as few as 11 students per grade level and therefore did not meet the requirement of a minimum of 20 students, hence the

shortfall in the actual students tested. The USAID Quality Reading Project administered the EGRA to students in Grades 1, 2, and 4, as part of the overall project design, but Grade 4 data are not presented in this report.

DATA CLEANING AND SCANNING

Once regions had completed their data collection and reviewed the instruments, the regional supervisors were called to submit the instruments to the USAID Quality Reading Project central office in Dushanbe. Upon collection of the data, the USAID Quality Reading Project team completed a thorough review of the data to ensure neatness and completeness. After ensuring the quality of the collected data, the instruments were handed out to data scanning personnel in the USAID Quality Reading Project office for scanning. It took about two weeks to complete the scanning of the data of 130 schools. The scanned files were then sent to AIR's psychometrician in Washington, DC, for analysis.

RELIABILITY MEASURES OF INTERNAL CONSISTENCY

Cronbach's alpha is a measure of internal consistency, that is, how closely related a set of questions are as a group. A "high" value of alpha is often used (along with substantive arguments and possibly other statistical measures) as evidence that the questions measure the same underlying (or latent) construct (e.g., comprehension skill). Reliability coefficients of 0.70 and above are considered to be adequate levels for educational testing (George & Mallery, 2003). For EGRA comprehension (comprising oral vocabulary, reading comprehension, and listening comprehension) and phonological awareness (comprising letter sound and initial letter sound) and dictation, the reliability coefficients for all grade and language assessment instruments were estimated between 0.72 and 0.88 (Table 4).

TABLE 4: RELIABILITY OF INTERNAL CONSISTENCY FOR COMPREHENSION AND PHONOLOGICAL AWARENESS AND DICTATION

Grade		Tajik	Russian
1	Comprehension	0.72	0.77
	Phonological Awareness and Dictation	0.87	0.85
2	Comprehension	0.75	0.79
	Phonological Awareness and Dictation	0.88	0.78

The reliability coefficient of 0.87 for phonological awareness and dictation in Tajik means that if a student takes a test that has a reliability coefficient of 0.87, he or she will receive a similar score on a test of equal difficulty 87 out of 100 times. For example, given that a student scores 15 out of 32 on a phonological awareness and dictation test in Tajik, if the student takes 100 similar but different tests (with equivalent difficulty), then the student will get about 15 out of 32 in 87 of the 100 tests. Therefore, we were able to estimate students' true ability in comprehension, phonological awareness, and dictation through the 2014 EGRA baseline, as though we had collected 100 similar test data points.

ANALYTIC STRATEGY

The EGRA data was analyzed in three steps. In Step 1, two types of statistics were produced: (1) item-level statistics to examine the psychometric properties of the items (e.g., item difficulty and item discrimination in classical theory, and *b*-value in item response theory), and (2) student statistics to report students' performance (i.e., raw scores) in each separate EGRA section and in combined sections (e.g., reading comprehension, listening comprehension, and oral vocabulary together). The Rasch model (1980) was used for item-response theory based

on concurrent item calibration. In Step 2, we calculated raw scores for each student in the following combination of EGRA sections to address specific research questions.

- Timed Sections
  - Calculated student reading fluency at the end of 1 and 2 minutes. If a student attempted all items before 1 or 2 minutes, we estimated corrected reading fluency as if the student had used the full 1 or 2 minutes. We also calculated the student reading fluency rate between minute 1 and minute 2 for students who took more than 1 minute but less than or equal to 2 minutes to finish each timed section (i.e., slow readers). It was hypothesized that slow readers read at a relatively higher pace in the second minute than in the first minute.
    - Reading fluency rate at the end of minute 1
    - Reading fluency rate at the end of minute 2
    - Reading fluency rate between minute 1 and 2

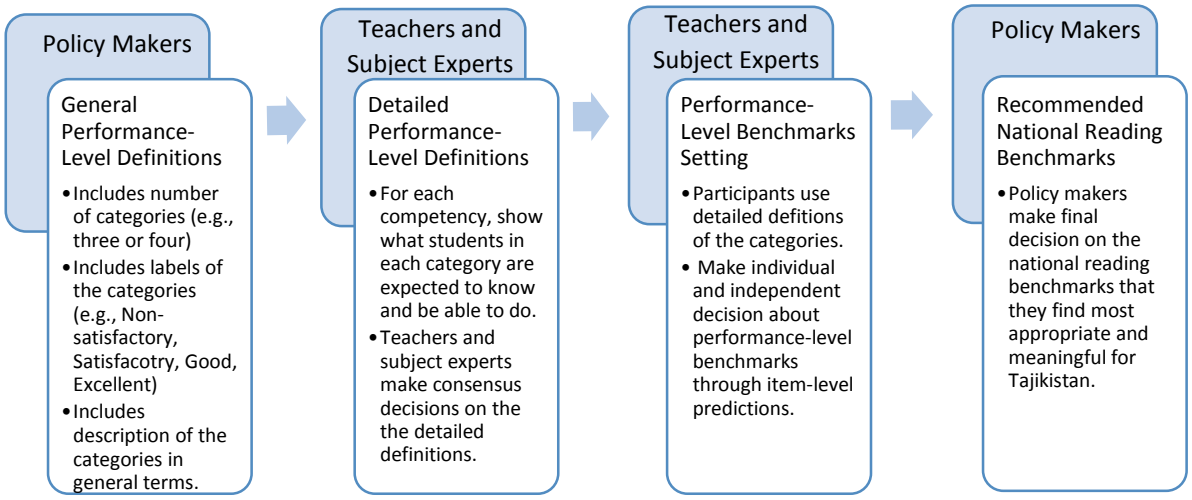
- Untimed Sections
  - Calculated separately the raw scores for each section (e.g., letter sound, initial letter sound, oral vocabulary, reading comprehension, listening comprehension, and dictation).
  - Calculated composite raw scores for comprehension that included oral vocabulary, reading comprehension, and listening comprehension.
  - Calculated composite raw scores for phonological consciousness (letter sound and initial letter sound) and dictation.

In Step 3, we performed a statistical significance test (e.g., *t*-test) to make comparisons between rural and urban students and between female and male students. This is important because simple comparisons are often made between groups without employing tests to ensure that any differences identified are statistically significant. Please note that no regional level analysis was attempted due to a lack of an adequate sample size within each region.

SETTING READING PERFORMANCE BENCHMARKS

Under the USAID Quality Reading Project, the MOES has recently approved primary grade national reading standards, which describe what students in primary grades are **expected to learn** in reading and comprehension. However, it did not define reading **performance benchmarks**, which would describe how students in primary grades are expected to perform in the reading standards, except for reading fluency measures (e.g., 25 words per minute for Grade 1, 40 words for Grade 2, 60 words for Grade 3, and 80 words for Grade 4). In order to set performance benchmarks for reading comprehension, phonological awareness, and dictation for primary grades, the USAID Quality Reading Project implemented a four-step (Figure 3) standard setting procedure (Loomis & Bourque, 2001; Beck, 2003; Cizek & Bunch, 2007; Perie, 2008).

FIGURE 3: THE VISUAL REPRESENTATION OF STANDARD-SETTING PROCESS



1. **DEVELOPING GENERAL DESCRIPTIONS OF PERFORMANCE-LEVEL CATEGORIES:** The USAID Quality Reading Project collaborated with the MOES and other stakeholders on June 17 to 19, 2014, for deciding the number of performance-level categories (that appropriately and meaningfully categorize students based on their performance in the tests) and their category names, and then defined those categories in general terms without necessarily specifying any national reading standards. The MOES decided to classify students into four performance-level categories, labeled them as Non-satisfactory, Satisfactory, Good, and Excellent, and then developed the general descriptions, as presented in Table 5.

TABLE 5: GENERAL DEFINITIONS OF PERFORMANCE-LEVEL CATEGORIES

Categories (Levels)	Definition
Excellent	The knowledge of the student **meets** the standard requirements and has **advance** skills of independent reading and comprehension. (Fluency: Grade 1—35 or more words per minute, Grade 2—50 or more words per minute, Grade 3—370 or more words per minute, Grade 4—90 or more words per minute).
Good	The knowledge of the student **meets** the standard requirements and has **good** skills of independent reading and comprehension. (Fluency: Grade 1—30 to 34 words per minute, Grade 2—45 to 49 words per minute, Grade 3—65 to 69 words per minute, Grade 4—85 to 89 words per minute).
Satisfactory	The knowledge of the student **meets** the standard requirements and has **enough** skills of independent reading and comprehension. (Fluency: Grade 1—25 to 29 words per minute, Grade 2—40 to 44 words per minute, Grade 3—60 to 64 words per minute, Grade 4—80 to 84 words per minute).
Non-satisfactory	The knowledge of the student **does not meet** the standard requirements and has **inefficient** ability of reading and comprehension. (Fluency: Grade 1—up to 25 words per minute, Grade 2—up to 40 words per minute, Grade 3—up to 60 words per minute, Grade 4—up to 80 words per minute).

2. **DEVELOPING DETAILED DEFINITIONS OF PERFORMANCE-LEVEL CATEGORIES:** The USAID Quality Reading Project conducted a four-day workshop from August 17 to 21, 2014, for developing detailed definitions of performance-level categories for each of the Grade 1 through 4 and in both Tajik and Russian. A total of 27 (14 in Tajik and 13 in Russian) teachers, language experts, curriculum experts, and standard developers attended the workshop. The participants developed detailed

consensus definitions for each competency within each component (e.g., phonemic awareness) of each of Grades 1 through 4's reading standards. Table 6 provides an example for Grade 1 in the Tajik language.

TABLE 6: AN EXAMPLE OF A DETAILED DEFINITION OF A COMPETENCY WITHIN A COMPONENT

	Unsatisfactory	Satisfactory	Good	Excellent
Standard 4.2. A student can distinguish and put in conformity printed, handwritten, capital and lowercase letters.	A student cannot distinguish and put in conformity the majority of printed, handwritten, capital and lowercase letters.	A student can distinguish and put in conformity some printed, handwritten, capital and lowercase letters.	A student can distinguish and put in conformity most printed, handwritten, capital and lowercase letters.	A student can distinguish and put in conformity all printed, handwritten, capital and lowercase letters.

3. **ESTABLISHING INTERIM CUT SCORES FOR PERFORMANCE-LEVEL CATEGORIES:** Following developing detailed definitions of performance-level categories, the USAID Quality Reading Project conducted a one-day workshop on August 22, 2014, with the same group of participants who developed detailed performance-level descriptions to establish an interim set of cut scores for each of the Grades 1 and 2 and for both Tajik and Russian. We used a yes–no variation of the Angoff method (Plake & Ferdous, 2005) for establishing the cut scores. Participants provided two rounds of individual and independent ratings of each item of EGRA in both Tajik and Russian, and feedback data was provided to the participants between the rounds. The round 2 ratings were used for estimating the interim cut scores, as it was assumed to be more reliable, robust, and informed than the ratings in round 1. The USAID Quality Reading Project's international consultant analyzed the round 2 rating data, calculated multiple sets of cut scores for each grade and language (i.e., Tajik and Russian) by adjusting varied level of judgmental errors (i.e., standard error of mean) and their corresponding impact data (i.e., percentage of students in performance-level categories). The standard error of mean was calculated using the following formula:

$$Standard\ Error\ of\ Mean\ (SE) = \frac{SD}{\sqrt{(N-1)}}$$

In which,

SD = Standard deviation of participants' item performance ratings within each component of EGRA; therefore, SE was calculated for each section of EGRA separately.

N = Number of participants who attended the standard-setting workshop.

4. **RECOMMENDING CUT SCORES AND DECISION MODEL: T**he USAID Quality Reading Project collaborated with the MOES in deciding the recommended cut scores for classifying student reading ability based on compensatory and conjunctive models. For example, in a compensatory model, weak performance in reading comprehension can be traded off against strong performance on reading fluency when calculating a score for a classification decision. The conjunctive model requires that the individual attain a minimum level of national performance standards (e.g., at cutoff or passing score) in both reading fluency and comprehension sections. A number of meetings, followed by a half-day workshop on August 23, 2014, were conducted with Deputy Minister of MOES, Ms. Tojinisso Mahmadova, and other higher-level officials in finalizing the cut scores. The MOES chose a set of cut scores that they thought meaningful and appropriate for the Tajikistan context in Tajik and Russian.

V. FINDINGS

This section presents the findings of the EGRA in Grade 1 and 2 students' performance in the Tajik and Russian languages. The sections in the EGRA test (i.e., letter name, letter sound, initial letter sound, familiar words, unfamiliar words, reading passage, reading comprehension, listening comprehension, and dictation) are presented throughout, along with findings in school location (rural and urban) and gender comparisons.

TABLE 7: DISTRIBUTION OF EGRA SAMPLE OF STUDENTS

Grade	Tajik			Russian		
	Male	Female	Total	Male	Female	Total
1	455	445	900	124	89	213
2	1062	1068	2130	220	163	383
	Rural	Urban	Total	Rural	Urban	Total
1	642	258	900	0	213	213
2	1714	416	2130	20	363	383

Table 7 represents the distribution sample of students who participated in Tajik and Russian EGRAs. There were 900 students in Grade 1 and 2,130 students in Grade 2, who tested in the Tajik language; and 213 in Grade 1 and 383 in Grade 2 who tested for Russian. When the students were classified into school types, there were 642 rural and 258 urban Grade 1 students and 1,714 rural and 416 urban Grade 2 students who tested in Tajik; and 213 urban and none rural Grade 1 students and 20 rural and 363 urban Grade 2 students who tested in Russian.

NATIONAL READING FLUENCY BENCHMARKS

According to national reading fluency benchmarks, students at Grades 1 and 2 must read at least 25 and 40 grade-appropriate words per minute, respectively. There were three sections in the EGRA that assessed student reading fluency: familiar words, unfamiliar words, and reading passage. The following table presents the percentage of students in Grades 1 and 2 who met the national reading fluency benchmarks in each of the sections, separately and collectively. The last column, Benchmark Met, refers to a conjunctive decision about the students' overall reading fluency on the EGRA test; i.e., the percentage of students in each of these grade levels met national fluency benchmark in all three sections separately. A Grade 1 a student is said to have met the Grade 1 national reading fluency benchmark if the student read at least 25 familiar words, 25 unfamiliar words, and 25 words in a reading passage per minute.

More than one-fourth of Grade 1 and one-seventh of Grade 2 students met the national reading fluency benchmarks in both Tajik and Russian. This was due to not meeting the benchmark for unfamiliar words. Students can read faster only familiar words and reading passages that contain familiar words, which could be caused by rote memorization of the words. They struggled with reading unfamiliar words, suggesting that students have difficulty with decoding.

GENDER COMPARISON

Students' performance in reading fluency was also compared by gender; it was quite distinct that female students consistently performed significantly higher than their male counterparts in all three sections did across both grade levels.

TABLE 8: PERCENTAGE OF STUDENTS MEETING TAJIK NATIONAL READING FLUENCY BENCHMARKS, BY GENDER

Grade	Gender	Tajik			
		Familiar Words	Unfamiliar Words	Reading Passage	Benchmark Met
1	Male	45.3	41.8	29.2	24.2
	Female	51.2	47.4	33.9	30.3
	Total	48.2	44.6	31.6	27.2
2	Male	36.8	15.5	28.7	12.7
	Female	44.4	20.8	39.5	18.2
	Total	40.6	18.2	34.1	15.4

Tajik

Table 8 shows that about 48 percent, 45 percent and 32 percent of Grade 1 students met the national benchmark (25 words per minute) in familiar words, unfamiliar words, and reading passage, respectively. In Grade 2, 41 percent, 18 percent, and 34 percent of the students met the national fluency benchmark (40 words per minute) in familiar words, unfamiliar words, and reading passage sections, respectively. When we calculated what percentage of students in each of the Grades 1 and 2 met national reading fluency benchmarks in all three sections separately and conjunctively, only 27 percent in Grade 1 and 15 percent in Grade 2 met the benchmarks conjunctively. As with Tajik language tests, female students outperformed male students in most sections across both grade levels.

TABLE 9: PERCENTAGE OF STUDENTS MEETING RUSSIAN NATIONAL READING FLUENCY BENCHMARKS, BY GENDER

Grade	Gender	Russian			
		Familiar Words	Unfamiliar Words	Reading Passage	Benchmark Met
1	Male	71.8	40.3	71.0	40.3
	Female	68.5	34.8	65.2	34.8
	Total	70.4	38.0	68.5	38.0
2	Male	68.2	10.0	47.7	9.5
	Female	73.0	17.2	50.9	16.6
	Total	70.2	13.1	49.1	12.5

Russian

A similar pattern was also observed for the Russian language (Table 9); students tend to do relatively well in familiar word and reading passage sections, but most struggle with unfamiliar words. About 70 percent, 38 percent, and 69 percent of Grade 1 and 70 percent, 13 percent, and 49 percent of Grade 2 students met national reading fluency benchmark in familiar words, unfamiliar words, and reading passage sections, respectively (25 words per minute for Grade 1 and 40 words per minute in Grade 2). When students' performance was collectively examined in all three sections together (conjunctively), it was revealed that more than one-third (38 percent) of Grade 1 and one-eighth (13 percent) of Grade 2 students met national reading fluency benchmarks at their respective grade levels. Similar to student performance for the Tajik language, female students in Grade 2 outperformed male students consistently in all three sections, whereas male students in Grade 1 performed higher than female students in familiar and unfamiliar words and reading passage sections.

TABLE 10: PERCENTAGE OF STUDENTS MEETING TAJIK NATIONAL READING FLUENCY BENCHMARKS, BY SCHOOL TYPE

Grade	Gender	Tajik			
		Familiar Words	Unfamiliar Words	Reading Passage	Benchmark Met
1	Rural	43.5	39.6	28.3	23.8
	Urban	60.1	57.0	39.5	35.7
	Total	48.2	44.6	31.5	27.2
2	Rural	37.7	16.3	31.9	13.4
	Urban	52.4	26.0	43.3	23.8
	Total	40.6	18.2	34.1	15.4

SCHOOL TYPE COMPARISON

Tajik

When student reading fluency measure was compared by school type (i.e., rural and urban), it was quite evident that students in urban schools tended to meet national reading fluency benchmarks more consistently, relative to their counterparts in rural schools across both grade levels. About 60 percent of Grade 1 students at urban schools met the national benchmark for familiar words (as compared to 44 percent in rural), 57 percent for unfamiliar words (as compared to 40 percent in rural), and 40 percent for reading passage section (as compared to 28 percent in rural). On the other hand, students in Grade 2 performed very poorly in reading unfamiliar words irrespective of their school types; only one-fourth (26 percent) of Grade 2 (as compared to 16 percent in rural) met the benchmark. When students were evaluated based on the three sections conjunctively, it was found that only 24 percent of rural and 36 percent of urban students for Grade 1 and 13 percent of rural and 24 percent of urban students for Grade 2 met the national reading fluency benchmarks.

Russian

The same pattern obtained for Tajik was also observed for Russian language results. In Grade 2, respectively, 72 percent of the students in urban schools outperformed students in rural schools (45 percent in Grade 2) in familiar words, whereas no comparison between rural and urban (i.e., 70 percent of students in familiar words, 38 percent in unfamiliar words, and 69 percent in reading passage met the national reading fluency benchmark) could be made for Grade 1 because no schools were selected from rural areas[4] (Table 11). In Grade 2, 14 percent and 51 percent of Grade 2 urban students (compared to 5 percent and 10 percent of Grade 2 rural students) met the benchmark in unfamiliar and reading passage, respectively.

TABLE 11: PERCENTAGE OF STUDENTS MEETING RUSSIAN NATIONAL READING FLUENCY BENCHMARKS, BY SCHOOL TYPE

Grade	Gender	Russian			
		Familiar Words	Unfamiliar Words	Reading Passage	Benchmark Met
1	Rural				
	Urban	70.4	38.0	68.5	38.0
	Total	70.4	38.0	68.5	38.0

[4] No schools in rural areas were selelcted through the random selection; this was pure chance.

Grade	Gender	Russian			
		Familiar Words	Unfamiliar Words	Reading Passage	Benchmark Met
2	Rural	45.0	5.0	10.0	5.0
	Urban	71.6	13.5	51.2	12.9
	Total	70.2	13.1	49.1	12.5

When students were evaluated based on three sections conjunctively, it was revealed that 38 percent of urban students for Grade 1 and 5 percent of rural and 13 percent of urban students for Grade 2 met the national reading fluency benchmarks (Table 11).

NATIONAL COMPREHENSION BENCHMARKS

The national reading comprehension benchmarks were set at the benchmark-setting workshop and were later approved by the MOES. The benchmarks were defined with respect to the percentage of comprehension questions (containing reading and listening comprehension and oral vocabulary sections of EGRA) that students at different grade levels are expected to answer correctly. Table 12 shows minimum percentage marks required for students in Grades 1 and 2 to be classified as meeting national comprehension benchmarks in Tajik and Russian.

Although, a higher percentage of students met the national reading comprehension benchmark, strong performance in oral vocabulary knowledge compensated highly for relatively weak performance in reading and listening comprehension.

TABLE 12: PERCENTAGE OF SCORES—NATIONAL COMPREHENSION BENCHMARKS

Grade	Tajik	Russian
1	50%	55%
2	50%	53%

For example, a student in Grade 1 is said to be meeting national comprehension benchmarks in Tajik if he or she receives a 50 percent score on a comprehension test consisting of oral vocabulary knowledge, and reading and listening comprehension questions. On the other hand, a student in the same grade level needs to get a 55 percent score for meeting national comprehension benchmarks in Russian (Table 12).

TABLE 13: PERCENTAGE OF STUDENTS MEETING NATIONAL COMPREHENSION BENCHMARKS, BY GENDER

Grade	Gender	Tajik	Russian
1	Male	91.9	88.7
	Female	94.2	92.1
	Total	93.0	90.1
2	Male	78.8	81.8
	Female	76.7	87.7
	Total	77.7	84.3

GENDER COMPARISON

Tajik

It was revealed that more than nine-tenths (93 percent) of Grade 1 and about four-fifths (78 percent) of Grade 2 students met national reading comprehension benchmarks (Table 13). When compared by gender, male students performed statistically significantly higher than their counterparts in Grades 2. A total of 92 percent of males vs. 94 percent of females in Grade 1

and 79 percent of males vs. 77 percent of females in Grade 2 met the comprehension benchmarks.

Russian

In Russian, 90 percent of students in Grade 1 and 84 percent in Grade 2 met national comprehension benchmarks (Table 13). When compared by gender, the opposite pattern of results was obtained in Russian than in Tajik. Female students performed statistically significantly better than their male peers at both grade levels; 92 percent of Grade 1 and 88 percent of Grade 2 females met the comprehension benchmark, as compared to 89 percent of Grade 1 and 82 percent of Grade 2 males.

SCHOOL TYPE COMPARISON

Tajik

When we compared the percentage of students meeting national comprehension benchmarks by school type, we observed that students in urban schools outperformed students in rural schools in both grade levels, and the differences were quite significant. More than 91 percent of Grade 1 students in rural schools and 97 percent in urban schools met the national benchmark (Table 14). In Grade 2, about 76 percent and 86 percent of students in rural and urban schools, respectively, met the benchmark.

TABLE 14: PERCENTAGE OF STUDENTS MEETING NATIONAL COMPREHENSION BENCHMARKS, BY SCHOOL TYPE

Grade	Type	Tajik	Russian
1	Rural	91.4	
	Urban	97.0	90.1
	Total	93.0	90.1
2	Rural	75.7	55.0
	Urban	86.1	86.0
	Total	77.7	84.3

Russian

In contrast, a substantially higher percentage of students in urban schools met the national comprehension benchmarks than their peers in rural schools (Table 14). In Grade 1, 90 percent of urban students (and no sample from rural schools) met the national benchmark. On the other hand, 86 percent of Grade 2 (as compared to 55 percent of rural) students in urban schools met the benchmark.

NATIONAL PHONOLOGICAL AND DICTATION BENCHMARKS

The national phonological and dictation benchmarks were recommended and approved by the MOES. The benchmarks were defined with respect to the percentage of phonological and dictation scores (containing letter sound, initial letter sound, and dictation sections of EGRA) that students at different grade levels are expected to obtain on the test. Table 15 shows the minimum score required for students in Grades 1 and 2 to be classified as meeting national phonological and dictation benchmarks in Tajik and Russian.

Students performed well in phonological consciousness and dictation. Over three-fourths of students in Tajik and four-fifths in Russian (both grades) languages met national phonological and dictation benchmarks. Strong performance in letter sound and initial letter sound compensated highly for relatively weak performance in dictation.

TABLE 15: PERCENTAGE OF SCORES REQUIRED FOR MEETING NATIONAL PHONOLOGICAL AND DICTATION BENCHMARKS

Grade	Tajik	Russian
1	63%	64%
2	64%	66%

For example, a student in Grade 1 is said to be meeting the national phonological and dictation benchmark in Tajik if he or she receives a 63 percent score on a test that consists of phonological awareness and dictation questions (Table 15). On the other hand, a student in the same grade level needs to get a 64 percent score to be classified as meeting the national benchmark in Russian.

TABLE 16: PERCENTAGE OF STUDENTS MEETING NATIONAL PHONOLOGICAL AND DICTATION BENCHMARKS, BY GENDER

Grade	Gender	Tajik	Russian
1	Male	74.7	79.0
	Female	82.9	83.1
	Total	78.8	80.8
2	Male	81.7	78.5
	Female	83.7	87.1
	Total	82.7	88.0

GENDER COMPARISON

Tajik

Students performed equally well in both the phonological awareness and dictation section and the comprehension sections. More than three-fourths of Grade 1 (75 percent of male and 83 percent of female) and four-fifths of Grades 2 (82 percent of male and 84 percent of female) students met national phonological and dictation benchmarks (Table 16). The female students in each grade level performed statistically significantly higher than the male students did.

Russian

In general, students in both grade levels irrespective of their gender performed higher in Russian phonological awareness and dictation (Table 16). In Grade 1, slightly more than four-fifths of Grade 1 (79 percent of male and 83 percent of female) and Grade 2 (79 percent of male and 87 percent of female) students met national phonological and dictation benchmarks. The percentage of female students who met the benchmark across the grade levels was statistically significantly higher than that of male students.

TABLE 17: PERCENTAGE OF STUDENTS MET NATIONAL PHONOLOGICAL AND DICTATION BENCHMARK, BY SCHOOL TYPE

Grade	Type	Tajik	Russian
1	Rural	76.3	
	Urban	84.9	80.8
	Total	78.8	80.8
2	Rural	81.9	65.0
	Urban	86.1	89.3
	Total	82.7	88.0

SCHOOL TYPE COMPARISON

Tajik

When we compared the percentage of students meeting national phonological and dictation benchmarks by school type, it was revealed that students in urban schools performed noticeably higher than students in rural schools (Table 17). In urban schools, about 85 percent of Grade 1 (as compared to 76 percent in rural schools) and 86 percent of Grade 2 (as compared to 82 percent in rural schools) met the national benchmark. The percentage of urban school students who met the benchmark across both grades was statistically significantly higher than the percentage of rural school students who met the benchmark.

Russian

The same pattern was observed for Russian language as was observed for Tajik. Overall, students in urban schools outperformed students in rural schools in both grade levels, and the differences were found to be statistically significant. About 80 percent of Grade 1 urban (no sample collected from rural schools) and 89 percent of Grade 2 urban (compared to 65 percent in rural) students met the national benchmark (Table 17).

NATIONAL READING FLUENCY AND COMPREHENSION BENCHMARKS

A conjunctive decision model was utilized to find out what percentage of students is meeting both reading fluency and comprehension benchmarks separately. These students read at a pace required at the grade level and understand what they read or listened to. The following table shows the national reading fluency and comprehension benchmarks for the Tajik and Russian languages.

Although three-fourths to nine-tenths of the students in Tajik and Russian languages met national reading comprehension benchmarks due to their strong performance in oral vocabulary, most of them did not meet the reading fluency benchmark due to their weak performance in reading unfamiliar words.

TABLE 18: NATIONAL FLUENCY AND COMPREHENSION BENCHMARKS

Grade	Tajik	Russian
1	*Fluency:* 25 words per minute & *Comprehension:* 50% score	*Fluency:* 25 words per minute & *Comprehension:* 55% scores
2	*Fluency:* 40 words per minute & *Comprehension:* 50% score	*Fluency:* 40 words per minute & *Comprehension:* 53% score

A student is said to have met Grade 1 national reading fluency and comprehension benchmarks in Tajik if he or she read at least 25 words per minute (as the measure of fluency) and obtained at least 50 percent score in comprehension sections (comprising oral vocabulary, reading comprehension, and listening comprehension) on the EGRA test (Table 18).

TABLE 19: PERCENTAGE OF STUDENTS MEETING NATIONAL READING FLUENCY AND COMPREHENSION BENCHMARKS, BY GENDER

Grade	Gender	Tajik	Russian
1	Male	23.7	37.9
	Female	29.0	33.7
	Total	26.3	36.2
2	Male	11.4	9.1
	Female	16.4	16.0
	Total	13.9	12.0

GENDER COMPARISON

Tajik

As expected, a lower percentage of students met both national reading fluency and comprehension benchmarks; only one-fourth (26 percent) of Grade 1 and one-seventh (14 percent) of Grade 2 students met the national benchmarks (Table 19). However, a higher percentage of female students than male students met the benchmarks. About 29 percent of Grade 1 female students (compared to 24 percent of males) and 16 percent of Grade 2 females (compared to 11 percent of males) met the benchmarks.

Russian

Although a higher percentage of students met national reading fluency and comprehension benchmarks in the Russian language than in the Tajik language, the percentages for Russian were very low, too. Only 36 percent of Grade 1 (38 percent males and 34 percent females) and 12 percent of Grade 2 (9 percent males and 16 percent females) students met the national benchmarks (Table 19). However, female students did significant better than male students for both grade levels.

SCHOOL TYPE COMPARISON

Tajik

When the percentage of students meeting national fluency and comprehension benchmarks was compared by school type, a mixed result was observed. Students in Grades 1 and 2 at urban schools performed statistically significantly better than their counterparts in rural schools (Table 20); about 34 percent of Grade 1 urban students (compared to 23 percent in rural students) and 22 percent of Grade 2 urban students (compared to 12 percent rural) met the benchmarks.

TABLE 20: PERCENTAGE OF STUDENTS MEETING NATIONAL READING FLUENCY AND COMPREHENSION BENCHMARKS, BY SCHOOL TYPE

Grade	Type	Tajik	Russian
1	Rural	23.2	
	Urban	34.1	36.2
	Total	26.3	36.2
2	Rural	11.8	5.0
	Urban	22.4	12.4
	Total	13.9	12.0

Russian

Although the pattern for Russian language favored urban schools over rural schools, the difference was only significant for Grade 2. More than one-third (36 percent) of Grade 1 urban students (no sample of students selected from rural schools) and one-eighth (12 percent) of Grade 2 urban students (as compared to 5 percent of students in rural schools) met the national benchmarks.

NATIONAL LITERACY BENCHMARKS

A conjunctive decision model was utilized to define the national literacy benchmark; students' performance in each reading fluency, comprehension, phonological awareness, and dictation sections would be counted separately to find out whether they had met or not met the national

literacy benchmark. The following table shows the national reading fluency, comprehension, phonological awareness, and dictation benchmarks for Tajik and Russian languages.

TABLE 21: NATIONAL LITERACY BENCHMARKS

Grade	Tajik	Russian
1	*Fluency:* 25 words per minute; *Comprehension:* 50% score; & *Phonological and Dictation:* 63% score	*Fluency:* 25 words per minute; *Comprehension:* 55% scores; & *Phonological and Dictation:* 64% score
2	*Fluency:* 40 words per minute; *Comprehension:* 50% score; & *Phonological and Dictation:* 64% score	*Fluency:* 40 words per minute; *Comprehension:* 53% score; & *Phonological and Dictation:* 66% score

A student is said to have met Grade 1 national literacy benchmarks in Tajik if he or she read at least 25 words per minute (as the measure of fluency) and obtained at least a 50 percent score in comprehension sections (comprises oral vocabulary, reading comprehension, and listening comprehension) and a 63 percent score in the phonological awareness and dictation sections of the EGRA test (Table 21).

TABLE 22: PERCENTAGE OF STUDENTS MEETING NATIONAL LITERACY BENCHMARKS, BY GENDER

Grade	Gender	Tajik	Russian
1	Male	22.4	32.3
	Female	27.6	32.6
	Total	25.0	32.4
2	Male	10.8	8.6
	Female	15.4	16.0
	Total	13.1	11.7

GENDER COMPARISON

Tajik

As expected, the percentage of students who met national literacy benchmarks (i.e., fluency, comprehension, phonological awareness, and dictation) was low; only 25 percent of students in Grade 1 and 13 percent in Grade 2 met the national literacy benchmarks (Table 22). The difference between the percentage of male and female students who met the benchmark was significant for Grades 1 and 2. About 28 percent of Grade 1 female students (compared to 22 percent males) and 15 percent of Grade 2 females (compared to 11 percent males) met the benchmarks.

More than one-fourth (25 percent in Tajik and 32 percent in Russian) of Grade 1 and one-tenth (13 percent in Tajik and 12 percent in Russian) of Grade 2 students met national literacy benchmarks, although a large number of students met comprehension, phonological, and dictation benchmarks. This is due to their poor performance in reading fluency, particularly in unfamiliar words.

Russian

Although the percentage of students who met national literacy benchmarks in the Russian language was very similar to that of the Tajik language, the gap between male and female student performance was significant for Grade 2. Only 32 percent of Grade 1 (32 percent of males and 33 percent of females) and 12 percent of Grade 2 (9 percent males and 16 percent females) students met the national benchmarks (Table 22).

SCHOOL TYPE COMPARISON

Tajik

When the percentage of students meeting national literacy benchmarks was compared by school type, a mixed result was observed. Students in Grades 1 and 2 at urban schools performed statistically significantly better than their counterparts in rural schools (Table 23); about 32 percent of Grade 1 urban students (compared to 22 percent of rural students) and 22 percent of Grade 2 urban students (compared to 11 percent of rural students) met the benchmarks.

TABLE 23: PERCENTAGE OF STUDENTS MEETING NATIONAL LITERACY BENCHMARKS, BY SCHOOL TYPE

Grade	Type	Tajik	Russian
1	Rural	22.1	
	Urban	32.2	32.4
	Total	25.0	32.4
2	Rural	11.1	5.0
	Urban	21.6	12.1
	Total	13.1	11.7

Russian

Overall, students in Grade 2 urban schools performed statistically significantly better than students in Grade 2 rural schools did (Table 23). About one-third (32 percent) of Grade 1 and more than one-eighth (12 percent) of Grade 2 students at urban schools met the national literacy benchmark. By contrast, only 5 percent of Grade 2 rural students met the same benchmark.

COMPREHENSION PERFORMANCE-LEVEL CATEGORIES

In addition to setting benchmarks for meeting or not meeting the national standards, the MOES also set multiple benchmarks relative to the national standards on a four-point performance-level categorical scale. The four points are called performance-level categories (e.g., Non-satisfactory, Satisfactory, Good, and Excellent), and students are classified into these categories based on their performance in reading comprehension (comprising oral vocabulary knowledge, reading, and listening comprehension). The following table provides benchmarks for the categories in the Tajik and Russian languages.

TABLE 24: BENCHMARKS FOR COMPREHENSION PERFORMANCE-LEVEL CATEGORIES

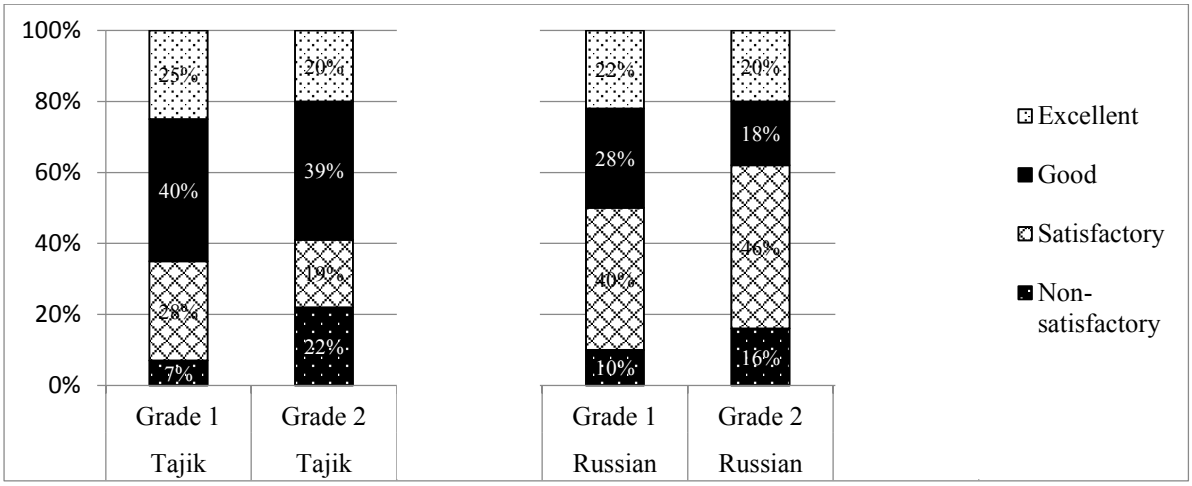
Language	Grade	% Score in Comprehension		
		Non-satisfactory/ Satisfactory	Satisfactory/ Good	Good/Excellent
Tajik	1	50%	70%	85%
	2	50%	60%	80%
Russian	1	55%	78%	89%
	2	53%	79%	89%

For example, a student in Grade 1 is going to be classified as Satisfactory, Good, or Excellent in Tajik reading comprehension if he or she receives, respectively, a 50 percent or 70 percent score on a comprehension test (Table 24). Ranges of scores represent the performance-level categories. For example, if a student receives a score between 0 and 49 percent, 50 and 69 percent, 70 and 84 percent, or 85 and 100 percent in reading comprehension, then he or she will be classified as Non-satisfactory, Satisfactory, Good, or Excellent, respectively.

Tajik

A similar pattern was observed across both grade levels; the majority of the students in their respective grade levels fell into the Good category. In Grade 1, 7 percent of the students were classified into Non-satisfactory, 28 percent into Satisfactory, 40 percent into Good, and 25 percent into the Excellent categories. In Grade 2, 22 percent of students were classified into Non-satisfactory, 19 percent into Satisfactory, 39 percent into Good, and 20 percent into the Excellent categories. In other words, more than three-fourths of Grades 1 and 2 students were in the top two categories (i.e., Good and Excellent) in comprehension (Figure 4).

FIGURE 4: PERCENTAGE DISTRIBUTION OF STUDENTS, BY PERFORMANCE-LEVEL COMPREHENSION CATEGORIES



Russian

In contrast, a higher percentage of students tested in the bottom two categories (i.e., Non-satisfactory and Satisfactory) in the Russian than the Tajik language; more than one-half of Grade 1 and three-fifths of Grade 2 students were either in the Non-satisfactory or Satisfactory categories. In Grade 1, 10 percent of students were classified into Non-satisfactory, 40 percent into Satisfactory, 28 percent into Good, and 22 percent into the Excellent categories. In Grade 2, 16 percent of students were classified into Non-satisfactory, 46 percent into Satisfactory, 18 percent into Good, and 20 percent into the Excellent categories (Figure 4).

PHONOLOGICAL AND DICTATION PERFORMANCE-LEVEL CATEGORIES

Students were also classified into the same four performance-level categories based on their performance in phonological awareness (comprising letter sound knowledge and initial letter sound) and dictation. The following table provides benchmarks for these categories.

For example, a student in Grade 1 is classified as Satisfactory, Good, or Excellent in Tajik phonological awareness and dictation if he or she receives a 63 percent, 78 percent, or 88 percent score on a phonological and dictation test that consists of letter sound, initial letter sound, and dictation sections, respectively (Table 25). As stated earlier, performance-level categories are represented by ranges of scores. For example, if a student receives a score between 63 percent and 77 percent in phonological awareness and dictation, then he or she will be classified as a Satisfactory student.

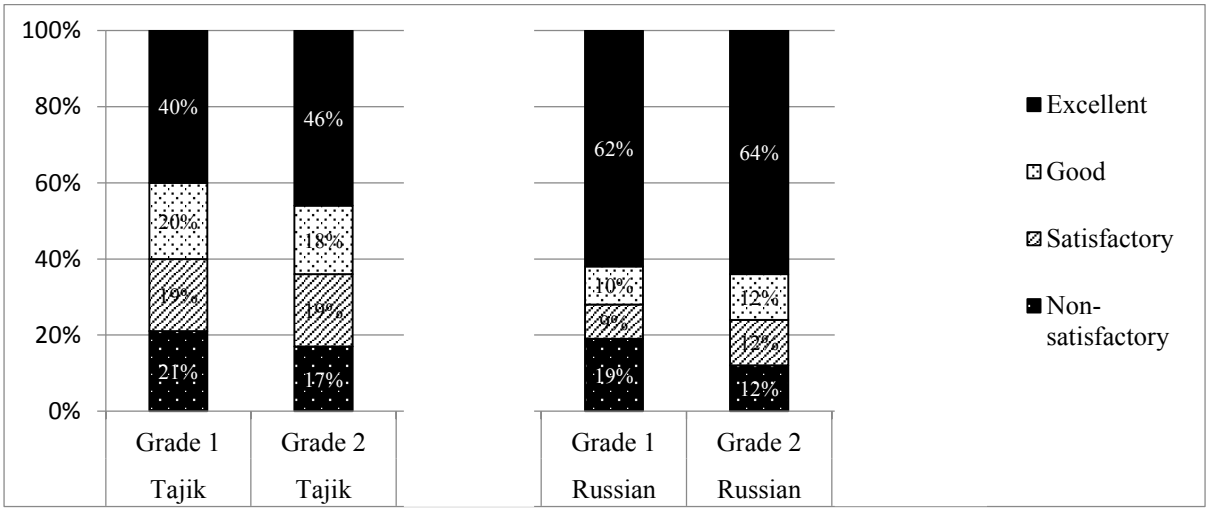
TABLE 25: CUT SCORES FOR PHONOLOGICAL AND DICTATION PERFORMANCE-LEVEL CATEGORIES

Language	Grade	% Score in Phonological and Dictation		
		Non-satisfactory/ Satisfactory	Satisfactory/Good	Good/Excellent
Tajik	1	63%	78%	88%
	2	64%	81%	89%
Russian	1	64%	72%	78%
	2	66%	74%	79%

Tajik

The majority of the students were in the Excellent category in both grade levels. In Grade 1, 21 percent of students were classified into Non-satisfactory, 19 percent into Satisfactory, 20 percent into Good, and 40 percent into the Excellent the categories. In Grade 2, 17 percent of students were classified into Non-satisfactory, 19 percent into Satisfactory, 18 percent into Good, and 46 percent into the Excellent categories (Figure 5).

FIGURE 5: PERCENTAGE DISTRIBUTION OF STUDENTS, BY PHONOLOGICAL AND DICTATION PERFORMANCE-LEVEL CATEGORIES



Russian

In Russian, the majority of the students were classified in the Excellent category. About 62 percent in Grade 1 and 64 percent in Grade 2 were classified as Excellent. The remaining students in Grade 1 were classified as 19 percent Non-satisfactory, 9 percent Satisfactory, and 10 percent Good. In Grade 2, 12 percent of students were classified as Non-satisfactory, Satisfactory, and Good (Figure 5).

SUBTASKS COMPARISON

This section presents the findings of the EGRA on Grades 1 and 2 student performance in the Tajik and Russian languages. The sections in the EGRA test (i.e., letter name, letter sound, initial letter sound, familiar words, unfamiliar words, reading passage, reading comprehension, listening comprehension, and dictation) are presented throughout, along with findings by school location (i.e., rural, semiurban, and urban) and gender comparisons.

LETTER NAME KNOWLEDGE

The first component of the EGRA assessed letter recognition skills. Each student received an EGRA instrument booklet that included all capital and small letters of the Tajik or Russian alphabets. These letters were randomly arranged in the booklets.

The majority of students could read Tajik and Russian letter names. Only 10 students (eight in Grade 1 and two in Grade 2) in Tajik and none in Russian samples obtained a score of zero. The students in Grade 2 who took more than 1 minute to finish the entire section read much faster in the second minute than in the first minute.

TABLE 26: LETTER NAME KNOWLEDGE FLUENCY, BY GENDER

Grade	Gender	Tajik			Russian		
		End of Min. 1	End of Min. 2	Between Min. 1 and 2	End of Min. 1	End of Min. 2	Between Min. 1 and 2
1	Male	55.5	86.5	43.6	67.0	87.0	40.8
	Female	60.0	86.6	42.7	62.5	91.5	54.0
	Total	57.7	86.6	43.2	65.1	89.3	47.6
2	Male	69.0	91.1	50.9	71.1	89.1	51.0
	Female	74.0	97.1	58.4	74.1	95.1	65.6
	Total	71.5	93.8	54.2	72.4	91.5	56.8

Note: End of Min. 1 refers to number of letters students read correctly in 1 minute; if the student took less than 1 minute, then the score was estimated for the entire 1 minute, and if the students took more than 1 minute, then the score at minute 1 was reported. End of Min. 2 refers to number of letters students (those who finished the entire letter naming section in more than 1 minute but less than or equal to 2 minutes) read correctly in 2 minutes; if the student took less than less than 2 minutes but more than 1 minute, and then the score was estimated for entire 2 minutes. Between Min. 1 and 2 refers to reading pace (of students who took more than 1 minute to finish the entire section) during the 2nd minute (i.e., number of letters read correctly during the 2nd minute).

GENDER COMPARISON

Tajik

It was revealed from the baseline data that student performance in letter name knowledge varied significantly across grade levels and gender. The students at Grades 1 and 2 read 58 and 72 Tajik letters correctly in 1 minute (Table 26). About 39 percent of Grade 1 and 62 percent of Grade 2 students attempted to finish the entire section in less than 1 minute, but only 10 students (eight in Grade 1 and two in Grade 2) got a score of zero. On the other hand, students in Grades 1 and 2 those who took more than 1 minute but less than or equal to 2 minutes to finish the entire letter naming section read 87 and 94 letters correctly in 2 minutes, respectively. It was evident from minutes 1 and 2 data that students (those who took more than 1 minute to finish the entire section) do not always read at a linear pace in minute 1 and 2; students at Grade 1 read at a faster (87 - 43 = 44 vs. 43) pace in the first minute than in the second minute. Students in Grade 2 read at a faster pace (94 - 54 = 40 vs. 47) in the second minute than in the first minute. When students' performance in letter name knowledge was compared by gender, it was quite distinct that female students performed significantly higher than their counterparts did consistently at the end of minute 1 (56 vs. 60 for Grade 1 and 69 vs. 74 for Grade 2) and minute 2 (91 vs. 97 for Grade 2 only).

Russian

The students in Grades 1 and 2 read 65 and 89, and 72 and 92 Russian letters correctly at the end of minutes 1 and 2, respectively (Table 26). However, about 41 percent of Grade 1 and 74 percent of Grade 2 students attempted the entire section in less than 1 minute and no students got a score of zero. An opposite pattern was also observed when comparing the reading pace of those took more than 1 minute and less than or equal to 2 minutes to finish the entire section at the end of minutes 1 and 2. Students in both Grades 1 and 2 read at a faster pace (89 – 48 =

41 vs. 48 in Grade 1 and 92 – 57 = 35 vs. 57 in Grade 2) in the second minute than in the first minute. A mixed pattern was observed when letter naming was compared by gender across grades. The male students in Grade 1 performed noticeably better (67 vs. 63 at the end of minute 1) than their female counterparts. The pattern was reversed for Grade 2 students, where the female students obtained higher scores than male students at the end of minute 1 (71 vs. 74). None of the differences between male and females in Grades 1 and 2 was found to be statistically significant.

SCHOOL TYPE COMPARISON

Tajik

When students' letter naming knowledge in Tajik was compared by school type, it was revealed that students in Grades 1 and 2 in urban schools performed better than their peers in the rural schools did, respectively. Moreover, students in Grade 1 rural and urban schools read 56 and 62 letters correctly in 1 minute and 85 and 93 letters in 2 minutes, respectively (Table 27). Whereas students in Grade 2 rural and urban schools read at a rate of 71 and 75 letters in 1 minute and 94 and 95 letters, respectively, in 2 minutes. The differences between rural and urban schools were statistically significant for Grade 1 students both at the end of minutes 1 and 2, and only for Grade 2 students at the end of minute 1. Students (those who took more than 1 minute to finish the entire section) in Grade 1 rural and urban schools read at a uniform pace (85 - 42 = 43 vs. 42 for rural and 93 - 46 = 47 vs. 46 for urban) in both minutes 1 and 2. In contrast, students in Grade 2 rural schools read at a faster pace (94 - 55 = 39 vs. 55 for rural as compared to 95 - 48 = 47 vs. 48 for urban) in the second minute than in the first.

TABLE 27: LETTER NAME KNOWLEDGE FLUENCY, BY SCHOOL TYPE

Grade	Type	Tajik			Russian		
		End of Min. 1	End of Min. 2	Between Min. 1 and 2	End of Min. 1	End of Min. 2	Between Min. 1 and 2
1	Rural	55.9	84.5	42.3			
	Urban	62.3	92.9	46.1	65.1	89.3	47.6
	Total	57.7	86.6	43.2	65.1	89.3	47.6
2	Rural	70.8	93.5	55.3	69.0	65.9	37.0
	Urban	74.8	95.1	48.4	72.6	93.3	58.2
	Total	71.5	93.8	54.2	72.4	91.5	56.8

Note: End of Min. 1 refers to number of letters students read correctly in 1 minute; if the student took less than 1 minute, then the score was estimated for the entire 1 minute, and if the students took more than 1 minute, then the score at minute 1 was reported. End of Min. 2 refers to number of letters students (those who finished the entire letter naming section in more than 1 minute but less than or equal to 2 minutes) read correctly in 2 minutes; if the student took less than less than 2 minutes but more than 1 minute, then the score was estimated for entire 2 minutes. Between Min. 1 and 2 refers to reading pace (of students who took more than 1 minute to finish the entire section) during the 2nd minute (i.e., number of letters read correctly during the 2nd minute).

Russian

In contrast, students in Grade 2 in urban schools performed statistically significantly higher than their counterparts in rural schools did. In Grade 2, students in urban schools read 73 letters (as compared to 69 in rural) at the end of minute 1 and 93 letters (as compared to 66 in rural) at the end of minute 2, respectively (Table 26). No comparison between rural and urban schools could be made for Grade 1, as the sample comprised of only urban schools. On the other hand, students (those who took more than 1 minute to finish the entire section) in Grade 2 rural and urban schools read at a faster pace (66 - 37 = 29 vs. 37 for rural and 93 - 58 = 35 vs. 58 for urban) in the second minute than in the first.

TABLE 28: LETTER SOUND KNOWLEDGE, BY GENDER

Grade	Gender	Tajik	Russian
1	Male	8.1	8.7
	Female	8.2	8.3
	Total	8.2	8.5
2	Male	7.9	8.9
	Female	8.0	8.8
	Total	8.0	8.8

LETTER SOUND

The second component of the EGRA assessed phonological awareness. Each student received an EGRA instrument booklet that included a list of the 10 most frequently used letters in Tajik or Russian alphabets (Table 28). These letters were chosen based on the letter-frequency lists generated by the USAID Quality Reading Project team from Grades 1 and 2 reading textbooks, and letters were randomly arranged in the booklets. The maximum score designated for this section was 10.

In Tajik, 52 (5.8 percent) students in Grade 1 and 81 (3.8 percent) in Grade 2 obtained a score of zero. In Russian, only two students in Grade 1 and one student in Grade 2 obtained a zero score in the initial letter sound section.

GENDER COMPARISON

Tajik

Students in Grades 1 and 2 obtained relatively higher scores, irrespective of the languages. In Tajik, students in Grades 1 and 2 obtained average scores of 8.2 and 8.0 out of possible score of 10, respectively (Table 28). A total of five students in Grade 1 and 16 students in Grade 2 obtained scores of zero in Tajik letter sound. When compared by gender, female students performed better than male students in Grades 1 and 2, but the differences were not statistically significant.

Russian

Students in Grades 1 and 2 received average scores of 8.5 and 8.8 in Russian, respectively (Table 28). Only one student in Grade 1 and no students in Grade 2 received a score of zero in Russian letter sound. When compared by gender, although male students performed relatively better than female students in Grade 1, the differences were not statistically significant.

SCHOOL TYPE COMPARISON

Tajik

When students' performance in letter sound was compared by school type, no meaningful differences between rural and urban schools were observed in Grade 1 (Table 29), but differences were observed in Grade 2. In general, students in both grades irrespective of school type obtained on an average of 7.9 marks or more out of 10; however, students in Grade 2 rural and urban schools secured average scores of 7.9 and 8.2, respectively, and the difference was found statistically significant.

TABLE 29: LETTER SOUND KNOWLEDGE, BY SCHOOL TYPE

Grade	Type	Tajik	Russian
1	Rural	8.1	
	Urban	8.2	8.5
	Total	8.2	8.5

Grade	Type	Tajik	Russian
2	Rural	7.9	8.8
	Urban	8.2	8.8
	Total	8.0	8.8

Russian

By contrast, students in Russian letter sound knowledge obtained average scores 7.8 to 8.9 out of a possible score of 10 (Table 29). Students in Grade 2 earned an average 8.8, irrespective of school type.

INITIAL LETTER SOUND

The third section of EGRA assessed phonological awareness. The purpose of this section was to examine whether students in Grades 1 and 2 could identify the initial sound of common words used at their grade level. This was a listening exercise. The administrator reads aloud 10 one-syllable, simple words (one word at a time) and students are asked to make the initial sound of each of the words. The maximum score allocated for this section was 10.

About one-fourth to one-third students in Tajik and over two-fifth students in Russian obtained the perfect scores in letter sound knowledge. In Tajik, five in Grade 1 and 16 in Grade 2 obtained a score of zero. In Russian, only one student in Grade 1 obtained a score of zero.

TABLE 30: INITIAL LETTER SOUND, BY GENDER

Grade	Gender	Tajik	Russian
1	Male	7.2	7.8
	Female	7.7	8.4
	Total	7.5	8.1
2	Male	7.4	8.5
	Female	7.7	8.4
	Total	7.6	8.5

GENDER COMPARISON

Tajik

The pattern for initial letter sound knowledge was similar to that for letter sound knowledge. Both male and females students performed very similarly in both Tajik and Russian. In Tajik, students in Grades 1 and 2 obtained average scores of 7.5 and 7.6 out of a possible 10. However, a large number of students in both Grades 1 (52 students) and 2 (81 students) obtained a score of zero. When compared by gender, it was revealed that female students in both Grades 1 and 2 performed statistically significantly better than their male counterparts. Male and female students at Grades 1 and 2 obtained average scores of 7.2 and 7.7, and 7.4 and 7.7, respectively (Table 30).

Russian

Female students in Grade 1 and male students in Grade 2 performed higher than their respective counterparts did, but the differences between them were not statistically significant. There were only two students in Grade 1 and one student in Grade 2 who obtained a score of zero.

TABLE 31: INITIAL LETTER SOUND, BY SCHOOL TYPE

Grade	Type	Tajik	Russian
1	Rural	7.3	
	Urban	7.8	8.1
	Total	7.5	8.1
2	Rural	7.6	7.8
	Urban	7.5	8.5
	Total	7.6	8.5

SCHOOL TYPE COMPARISON

Tajik

A similar pattern was also observed when compared by school type. For Tajik, students at Grade 1 in rural and urban schools obtained average scores of 7.3 and 7.8 out of a possible score of 10, respectively (Table 31). On the other hand, students in Grade 2 rural and urban schools received average scores of 7.6 and 7.5, respectively. The difference between rural and urban schools was only significant for Grade 1 students.

Russian

In contrast, for Russian, although there were noticeable differences in students' performance in initial letter sound knowledge between rural and urban schools, no differences were found to be statistically significant. Students' performances across grade level and school type ranged between 7.8 for Grade 2 rural school and 8.5 for Grade 2 urban schools, whereas students in Grade 1 urban schools obtained average scores of 8.1 (Table 31).

FAMILIAR WORD IDENTIFICATION

The fourth section of EGRA assessed familiar word identification. The purpose of this section was to examine whether students in Grades 1 and 2 were able to read familiar words aloud at their grade levels. A list of 25 familiar words at Grade 1 and 40 words at Grade 2 were selected from primary grade textbooks and were randomly arranged in the student booklets.

More than half of the students in Tajik and three-fourths of students in Russian finished the entire familiar word section in less than a 1 minute. In Tajik, 13 students in Grade 1 and 11 in Grade 2 obtained a score of zero. In Russian, three students in Grade 1 and two in Grade 2 scored a zero.

TABLE 32: FAMILIAR WORD IDENTIFICATION, BY GENDER

Grade	Gender	Tajik			Russian		
		End of Min. 1	End of Min. 2	Between Min. 1 and 2	End of Min. 1	End of Min. 2	Between Min. 1 and 2
1	Male	25.9	23.9	14.6	35.5	23.1	15.9
	Female	27.4	25.1	14.3	34.3	24.3	12.9
	Total	26.6	24.5	14.5	35.0	23.6	14.5
2	Male	35.2	42.8	22.9	54.1	51.6	29.2
	Female	39.4	45.4	23.9	54.5	45.9	22.6
	Total	37.3	44.1	23.4	54.3	49.4	26.7

Note: End of Min. 1 refers to number of words students read correctly in 1 minute; if the student took less than 1 minute, then the score was estimated for the entire 1 minute, and if the students took more than 1 minute, then the score at minute 1 was reported. End of Min. 2 refers to number of words students (those who finished the entire familiar word section in more than 1 minute but less than or equal to 2 minutes) read correctly in 2 minutes; if the student took less than less than 2 minutes but more than 1 minute, then the score was estimated for entire 2 minutes. Between Min. 1 and 2 refers to reading pace (of students who took more than 1 minute to finish the entire section) during the 2nd minute (i.e., number of familiar words read correctly during the 2nd minute).

GENDER COMPARISON

Tajik

The students in Grades 1 and 2 read 27 and 37 words in 1 minute, respectively (Table 32). About 55 percent of Grade 1 and 48 percent of Grade 2 students attempted the entire section in less than 1 minute; 13 students at Grade 1 and 11 students at Grade 2 received a score of zero. Of the students who took more than 1 minute but less than or equal to 2 minutes to finish the entire familiar word identification section, Grade 1 students read 25 words and Grade 2 students read 44 words correctly in 2 minutes. It was evident from the data from minutes 1 and 2 that students do not always read at a linear pace; students (those who took more than 1 minute but less than or equal to 2 minutes) read at a faster (25 - 15 = 10 vs. 15 for Grade 1 and 44 - 23 = 21 vs. 23 for Grade 2) pace in the second minute than in the first minute. When students' performance in familiar word identification was compared by gender, it was quite distinct that females performed better than their counterparts did consistently at the end of minutes 1 (26 vs. 27 for Grade 1 and 35 vs. 39 for Grade 2) and 2 (24 vs. 25 for Grade 1 and 43 vs. 45 for Grade 2). However, the differences between males and females were statistically significant for Grade 2 students.

Russian

By contrast, the students in Grades 1 and 2 read 35 and 54 Tajik familiar words and 24 and 49 Russian familiar words correctly at the end of minutes 1 and 2, respectively (Table 32). However, about 75 percent of Grade 1 and 76 percent of Grade 2 students attempted the entire section in less than 1 minute; three students in Grade 1 and two students in Grade 2 received a score of zero. A similar pattern was also observed when comparing the reading pace of those who took more than 1 minute and less than or equal to 2 minutes to finish the entire section at the end of minutes 1 and 2. Students at Grades 1 and 2 read at a faster pace (25 - 15 = 10 vs. 15 in Grade 1, 49 - 27 = 22 vs. 27 in Grade 2) in the second minute than in the first minute. A mixed pattern was observed when it was compared by gender across grades. The male students in Grade 1 performed better than females at the end of minute 1 (36 for male vs. 34 for female), but females performed better than males at the end of minute 2 (23 for male vs. 24 for female). In Grade 2, female students at the end of minute 1 (54 vs. 55) and male students at the end of minute 2 (52 vs. 46) performed better than their respective counterparts. No differences between male and female students in Grades 1 and 2 were statistically significant.

TABLE 33: FAMILIAR WORD IDENTIFICATION, BY SCHOOL TYPE

Grade	Type	Tajik			Russian		
		End of Min. 1	End of Min. 2	Between Min. 1 and 2	End of Min. 1	End of Min. 2	Between Min. 1 and 2
1	Rural	24.9	24.2	14.1			
	Urban	31.0	25.5	15.8	35.0	23.6	14.5
	Total	26.6	24.5	14.5	35.0	23.6	14.5
2	Rural	36.2	43.9	23.8	34.1	38.4	20.7
	Urban	41.9	44.7	21.3	55.4	51.1	27.6
	Total	37.3	44.1	23.4	54.3	49.4	26.7

Note: End of Min. 1 refers to number of words students read correctly in 1 minute; if the student took less than 1 minute, then the score was estimated for the entire 1 minute, and if the students took more than 1 minute, then the score at minute 1 was reported. End of Min. 2 refers to number of words students (those who finished the entire familiar word section in more than 1 minute but less than or equal to 2 minutes) read correctly in 2 minutes; if the student took less than less than 2 minutes but more than 1 minute, then the score was estimated for entire 2 minutes. Between Min. 1 and 2 refers to reading pace (of students who took more than 1 minute to finish the entire section) during the 2nd minute (i.e., number of familiar words read correctly during the 2nd minute).

SCHOOL TYPE COMPARISON

Tajik

When students' familiar word identification in Tajik was compared by school type (i.e., rural and urban), it was revealed that students in Grades 1 and 2 at urban schools (as compared to rural schools) performed relatively higher. Students in Grade 1 of rural and urban schools read 25 and 31 words correctly in minute 1 and 24 and 26 words in minute 2, respectively. Whereas, students in Grade 2 rural and urban schools read at a rate of 36 and 42 words in minute 1 and 44 and 45 words in minute 2, respectively (Table 33). The differences between rural and urban students' scores at the end of minute 1 were statistically significant across both grade levels.

Russian

A similar pattern was also observed for the Russian language as it was obtained for the Tajik language. Students in urban schools outperformed students in rural school both in Grades 2; however, the difference was only statistically significant for Grade 2 students at the end of both minutes 1 and 2. Students at Grade 2 in rural and urban schools read at 34 and 55 words in minute 1, and 38 and 51 words in minute 2, respectively. The comparison for Grade 1 was not possible, as students from rural Russian schools were not selected in the sample.

UNFAMILIAR WORD IDENTIFICATION

The fifth section of EGRA assessed unfamiliar word identification. The purpose of this section was to examine whether students in Grades 1 and 2 were able to decode successfully unfamiliar words appropriate to their grade levels. A list of 25 unfamiliar words for Grade 1 and 40 words for each of Grade 2 were randomly arranged in the student booklets.

Only one-fourth to one-half of the sample students finished the entire unfamiliar word section in less than 1 minute. On average, students read correctly about 21 to 25 words in Grade 1 and 27 to 29 words in Grade 2. In Tajik, 13 students in Grade 1 and 16 in Grade 2 obtained a score of zero. In Russian, six students in Grade 1 and four in Grade 2 scored a zero.

TABLE 34: UNFAMILIAR WORD IDENTIFICATION, BY GENDER

Grade	Gender	Tajik			Russian		
		End of Min. 1	End of Min. 2	Between Min. 1 and 2	End of Min. 1	End of Min. 2	Between Min. 1 and 2
1	Male	23.8	24.5	16.4	21.5	25.3	18.6
	Female	25.9	25.6	13.8	20.3	23.7	15.4
	Total	24.9	25.0	15.2	21.0	24.6	17.2
2	Male	25.5	39.6	24.1	28.4	46.4	28.7
	Female	28.7	42.0	23.8	29.4	45.9	24.7
	Total	27.1	40.7	24.0	28.8	46.2	27.1

Note: End of Min. 1 refers to number of words students read correctly in 1 minute; if the student took less than 1 minute, then the score was estimated for the entire 1 minute, and if the students took more than 1 minute, then the score at minute 1 was reported. End of Min. 2 refers to number of words students (those who finished the entire unfamiliar word section in more than 1 minute but less than or equal to 2 minutes) read correctly in 2 minutes; if the student took less than less than 2 minutes but more than 1 minute, then the score was estimated for an entire 2 minutes. Between Min. 1 and 2 refers to reading pace (of students who took more than 1 minute to finish the entire section) during the 2nd minute (i.e., number of unfamiliar words read correctly during the 2nd minute).

GENDER COMPARISON

Tajik

Overall, student performance in unfamiliar word identification was very poor relative to their performance in familiar word identification. Students in Grades 1 and 2 read about 25 and 27

unfamiliar words per minute (Table 34), respectively. About 54 percent of Grade 1 and 24 percent of Grade 2 students attempted the entire unfamiliar word section in less than 1 minute; 13 students in Grade 1 and 16 students in Grade 2 got a score of zero. Of the students who took more than 1 minute but less than or equal to 2 minutes to read the entire section, Grade 1 and 2 students read 25 and 41 words in 2 minutes, respectively. When their reading pace was compared at the end of the first and second minutes, it was revealed that students in Grades 1 and 2 read faster (25 - 15 = 10 vs. 15; 41-24=17 vs. 24) in the second minute than in the first minute. Female students performed statistically significantly higher than their counterparts in both grades did consistently at the end of minute 1 (24 vs. 26 for Grade 1, 26 vs. 29 for Grade 2) and minute 2 (40 vs. 42 for Grade 2).

Russian

In contrast, students in Grades 1 and 2 read 21 and 29 unfamiliar Russian words in 1 minute and 25 and 46 unfamiliar words in 2 minutes, respectively (Table 34). About one-half of Grade 1 and one-fourth of Grade 2 students attempted the entire unfamiliar word section in less than 1 minute; six students in Grade 1 and four students in Grade 2 received a score of zero. A similar pattern was also observed when comparing the reading pace of students who took more than 1 minute but less than or equal to 2 minutes to finish the entire section. Grades 1 and 2 read faster (25 - 17 = 8 vs. 17; 46 - 27 = 19 vs. 27) in the second minute than in the first minute. A mixed pattern was observed when compared by gender. Although male students in Grade 1 read a higher number of words than females students at the end of minutes 1 (22 vs. 20) and 2 (25 vs. 24), the differences were not statistically significant. On the other hand, female students in Grade 2 performed statistically significantly better than male students did at the end of minutes 1 (28 vs. 29 for Grade 2).

SCHOOL TYPE COMPARISON

Tajik

When students' unfamiliar word identification in Tajik was compared by school type (i.e., rural and urban), it was observed that students in urban schools performed statistically significantly better than students in rural schools, irrespective of their grade level. In Tajik, students in Grade 1 rural and urban schools read 23 and 29 words correctly in minute 1 and 25 and 27 words in minute 2, respectively (Table 35), whereas students in Grade 2 of rural and urban schools read at a rate of 26 and 31 words in minute 1, and 40 and 43 words in minute 2, respectively.

TABLE 35: UNFAMILIAR WORD IDENTIFICATION, BY SCHOOL TYPE

Grade	Type	Tajik			Russian		
		End of Min. 1	End of Min. 2	Between Min. 1 and 2	End of Min. 1	End of Min. 2	Between Min. 1 and 2
1	Rural	23.3	24.5	14.4			
	Urban	28.9	26.8	17.9	21.0	24.6	17.2
	Total	24.9	25.0	15.2	21.0	24.6	17.2
2	Rural	26.3	40.2	23.6	22.6	38.9	20.3
	Urban	30.5	43.1	25.6	29.1	46.7	27.6
	Total	27.1	40.7	24.0	28.8	46.2	27.1

Note: End of Min. 1 refers to number of words students read correctly in 1 minute; if the student took less than 1 minute, then the score was estimated for the entire 1 minute, and if the students took more than 1 minute, then the score at minute 1 was reported. End of Min. 2 refers to number of words students (those who finished the entire unfamiliar word section in more than 1 minute but less than or equal to 2 minutes) read correctly in 2 minutes; if the student took less than less than 2 minutes but more than 1 minute, then the score was estimated for an entire 2 minutes. Between Min. 1 and 2 refers to reading pace (of students who took more than 1 minute to finish the entire section) during the 2nd minute (i.e., number of unfamiliar words read correctly during the 2nd minute).

Russian

The pattern observed for Russian unfamiliar words was similar to that obtained for the Tajik language. Students in urban schools performed statistically significantly better than their counterparts in rural schools in Grade 2. Grade 2 students in rural and urban schools read 23 and 29 unfamiliar words at the end of minute 1, respectively. Students in Grade 2 read 39 words (rural) and 47 words (urban) (Table 35); the difference was statistically significant.

ORAL VOCABULARY KNOWLEDGE

The sixth section of the EGRA assessed oral vocabulary. The purpose of this section was to examine whether students in Grades 1 and 2 were able to understand meaning of familiar words at their grade levels. The administrator read 10 words aloud (one word at a time). In the student booklet, students were presented with a set of four pictures for each word read and were asked to identify the picture that best matched the word.

About one-fifth to two-fifths of students in Tajik and two-fifths to three-fourths of students in Russian obtained perfect scores in oral vocabulary knowledge. In Tajik, two students in Grade 1 and three in Grade 2, and none in Russian scored a zero.

TABLE 36: ORAL VOCABULARY, BY GENDER

Grade	Gender	Tajik	Russian
1	Male	8.9	9.5
	Female	9.0	9.6
	Total	8.9	9.5
2	Male	8.0	8.8
	Female	8.0	8.9
	Total	8.0	8.8

GENDER COMPARISON

Tajik

Overall, students in Grades 1 and 2 performed very similarly in both the Tajik and Russian oral vocabulary knowledge section, and no observed differences were found to be statistically significant (Table 36). In Tajik, students in Grades 1 and 2 obtained average scores of 8.9 and 8.0 respectively, out of a possible score of 10. Only two students in Grade 1 and three students in Grade 2 obtained a score of zero. In contrast, about 45 percent of students in Grade 1 and 19 percent in Grade 2 obtained a perfect score (i.e., 10 out of 10) in this oral vocabulary knowledge section.

Russian

On the other hand, students in Grades 1 and 2 received average scores of 9.5 and 8.8, respectively, in Russian oral vocabulary knowledge (Table 36). About 77 percent of the students in Grade 1 and 43 percent in Grade 2 obtained a perfect score (i.e., 10 out of 10), and no students obtained a score of zero. When compared by gender, no meaningful differences were observed between them.

TABLE 37: ORAL VOCABULARY, BY SCHOOL TYPE

Grade	Type	Tajik	Russian
1	Rural	8.9	
	Urban	9.2	9.5
	Total	8.9	9.5

Grade	Type	Tajik	Russian
2	Rural	7.9	8.1
	Urban	8.2	8.9
	Total	8.0	8.8

SCHOOL TYPE COMPARISON

Tajik

When students' performance in oral vocabulary was compared by school type (rural and urban), no statistically significant differences between them were observed. In general, students in both grades irrespective of school type obtained 7.9 or more out of 10. Students in rural and urban schools obtained average scores of 8.9 and 9.2 in Grade 1 and 7.9 and 8.2 in Grade 2, respectively (Table 37).

Russian

By contrast, students in Russian oral vocabulary knowledge obtained average scores of 7.0 to 9.5 out of a possible score of 10. The students in Grade 2 at urban schools performed statistically significantly higher than their peers in the rural schools did. Students in Grade 2 at urban schools obtained an average score of 8.9 (as compared to 8.1 in rural schools). Grade 1 could not be compared because students from rural Russian schools were not selected in the sample.

READING PASSAGE

The seventh section of the EGRA assessed reading passage. The purpose of this section was to examine whether students in Grades 1 and 2 were able to read a passage aloud with comprehension at their respective grade levels. This section includes one short paragraph (~25 words for Grade 1 and ~40 words for Grade 2; words are grade-appropriate familiar words), presented in the student booklets.

The pattern of student performance in reading passages comprised of familiar words from textbooks was similar to the one observed in familiar word identification section. A total of 19 Grade 1 and 15 Grade 2 students in Tajik and two Grade 1 and one Grade 2 students in Russian obtained a score of zero.

TABLE 38: READING PASSAGE, BY GENDER

Grade	Gender	Tajik			Russian		
		End of Min. 1	End of Min. 2	Between Min. 1 and 2	End of Min. 1	End of Min. 2	Between Min. 1 and 2
1	Male	20.0	22.5	12.3	31.1	22.4	16.4
	Female	21.6	23.0	11.9	31.1	25.5	15.0
	Total	20.8	22.7	12.1	31.1	23.8	15.8
2	Male	33.2	46.2	23.8	41.7	57.6	34.9
	Female	39.1	47.5	25.3	43.2	59.1	33.2
	Total	36.1	46.8	24.5	42.3	58.2	34.2

Note: End of Min. 1 refers to number of words students read correctly in 1 minute; if the student took less than 1 minute, then the score was estimated for the entire 1 minute, and if the students took more than 1 minute, then the score at minute 1 was reported. End of Min. 2 refers to number of words students (those who finished the entire reading passage section in more than 1 minute but less than or equal to 2 minutes) read correctly in 2 minutes; if the student took less than less than 2 minutes but more than 1 minute, then the score was estimated for the entire 2 minutes. Between Min. 1 and 2 refers to reading pace (of students who took more than 1 minute to finish the entire section) during the 2nd minute (i.e., number of words in the passage read correctly during the 2nd minute).

GENDER COMPARISON

Tajik

Perhaps because the reading passages comprised grade-appropriate, familiar words, student performance in reading passages was similar to the familiar word identification section. In general, students in Grades 1 and 2 read 21 and 36 words in 1 minute, and 23 and 47 words in 2 minutes, respectively (Table 38). Over two-fifths (42 percent) of students in Grade 1 and one-third (33 percent) in Grade 2 attempted to finish the entire reading passage section in less than 1 minute; a total of 19 Grade 1 and 15 Grade 2 students obtained a score of zero. Like other reading fluency sections in the EGRA, students who took more than 1 minute but less than or equal to 2 minutes to finish the entire reading passage section also read at a much faster rate in minute 2 than in minute 1. Grade 1 students (23 - 12 = 11 vs. 12) read at a uniform rate in both minutes, and Grade 2 students (47 - 25 = 22 vs. 25) read at a faster rate in the second minute than in the first minute. When we compared student performance in reading passage by gender, it was revealed that female students performed better than their counterparts in both grade levels. Male and female students in Grades 1 and 2 read 20 vs. 22 words and 33 vs. 39 words in minute 1, and 23 vs. 23 words and 46 vs. 48 words in minute 2, respectively. However, the differences between them were only statistically significant for Grade 2 (during 1 minute).

Russian

By contrast, in the Russian language students in Grades 1 and 2 read at a rate of 31 and 42 words in 1 minute, and 24 and 58 words in 2 minutes, respectively (Table 38). A total of 72 percent of the students in Grade 1 and 38 percent in Grade 2 attempted to finish the entire reading passage section in less than 1 minute, but only two students in Grade 1 and one in Grade 2 obtained a score of zero. A similar pattern was also observed when comparing the student reading pace at the end of minutes 1 and 2. Students tended to read faster (24 - 16 = 8 vs. 16 for Grade 1; 58 - 34 = 24 vs. 34 for Grade 2) in the second minute than in the first minute. When we compared student performance in reading passage by gender, it was revealed that female students performed better than their male counterparts did. In Grade 1, male and female students read at a rate of 31 words per minute, and 22 and 26 words in two minutes, respectively, although the difference was not statistically significant. In Grade 2 and 4, female students read at a rate of 43 words per minute (as compared to 42 for males), and 59 words in 2 minutes (as compared to 58 for males). The differences were not statistically significant

TABLE 39: READING PASSAGE, BY SCHOOL TYPE

Grade	Type	Tajik			Russian		
		End of Min. 1	End of Min. 2	Between Min. 1 and 2	End of Min. 1	End of Min. 2	Between Min. 1 and 2
1	Rural	19.8	22.4	11.8			
	Urban	23.2	23.8	13.1	31.1	23.8	15.8
	Total	20.8	22.7	12.1	31.1	23.8	15.8
2	Rural	35.3	46.5	24.5	27.8	46.9	25.1
	Urban	39.6	48.3	24.5	43.1	59.2	35.0
	Total	36.1	46.8	24.5	42.3	58.2	34.2

Note: End of Min. 1 refers to number of words students read correctly in 1 minute; if the student took less than 1 minute, then the score was estimated for the entire 1 minute, and if the students took more than 1 minute, then the score at minute 1 was reported. End of Min. 2 refers to number of words students (those who finished the entire reading passage section in more than 1 minute but less than or equal to 2 minutes) read correctly in 2 minutes; if the student took less than less than 2 minutes but more than 1 minute, the score was estimated for an entire 2 minutes. Between Min. 1 and 2 refers to reading pace (of students who took more than 1 minute to finish the entire section) during the 2nd minute (i.e., number of words in the passage read correctly during the 2nd minute).

SCHOOL TYPE COMPARISON

Tajik

When students' scores of reading passage in Tajik and Russian were compared by school type (i.e., rural and urban), the same pattern observed for other reading sections of the EGRA were also revealed for the reading passage. In Tajik, students in urban schools for Grades 1 and 2 performed significantly better than their peers in the rural schools. Students at Grade 1 in rural and urban schools read 20 and 23 words correctly in 1 minute and 22 and 24 words in 2 minutes, respectively (Table 39). Whereas, students in Grade 2 at rural and urban schools read at a rate of 35 and 40 words in 1 minute and 47 and 48 words in 2 minutes, respectively.

Russian

On the other hand, in Russian, students in Grade 2 from urban schools read 43 words (as compared 28 in rural schools) at the end of minute 1, and 59 words (as compared to 47 in rural schools) at the end of minute 2 (Table 39). The differences between rural and urban were statistically significant for Grade 2 at the end of minutes 1 and 2. The comparison for Grade 1 was not possible, as students from rural Russian schools were not selected in the sample. Again, the same pattern was observed for the reading pace of both languages in minutes 1 and 2; students tend to read faster in minute 2 than in minute 1.

READING COMPREHENSION

The seventh section assessed reading comprehension. The purpose of this section was to examine whether students in Grades 1 and 2 were able to comprehend the passage they read. After the student read the passage aloud, he or she was asked 4 to 5 questions about the passage. In Tajik, there were five questions for each of Grades 1 and 2 related to their respective reading passages. In Russian, there were four questions for Grade 1 and five questions for Grade 2 related to their respective reading passages.

Students have difficulty comprehending the text that they read. A higher percentage of students obtained a score of zero in reading comprehension. About 19 percent of the students in Grade 1 and 28 percent in Grade 2 in Tajik, and 36 percent in Grade 1 and 21 percent in Grade 2 in Russian obtained a score of zero.

TABLE 40: READING COMPREHENSION, BY GENDER

Grade	Gender	Tajik	Russian
1	Male	2.1	1.2
	Female	2.1	1.5
	Total (out of the possible score: Tajik 5; Russian 4)	2.1	1.3
2	Male	1.9	2.2
	Female	2.0	2.4
	Total (out of the possible score: Tajik 5; Russian 5)	2.0	2.3

GENDER COMPARISON

Tajik

Overall, students in Grades 1 and 2 did not perform well in the Tajik and Russian reading comprehension section. In Tajik, students in Grades 1 and 2 obtained an average score of 2.1 out of the possible score of 5 and 2.0 out of 5, respectively (Table 40). About 19 percent of Grade 1 and 28 percent of Grade 2 obtained a score of zero, whereas 6 percent of Grade 1 and 9 percent of Grade 2 students secured a perfect score (i.e., 5 out of 5). Both male and female students performed similarly; therefore, no significant difference was observed.

Russian

However, in Russian, students in Grades 1 and 2 obtained average scores of 1.3 out of the possible score of 4 and 2.3 out of 5, respectively (Table 40). A total of 36 percent of Grade 1 and 21 percent of Grade 2 students obtained a score of zero in Russian reading comprehension, whereas 7 percent of Grade 1 and 12 percent of Grade 2 students received perfect scores. When compared by gender, although female students performed relatively better than male students in both grade levels, the difference between male and female students was not significant.

TABLE 41: READING COMPREHENSION, BY SCHOOL TYPE

Grade	Type	Tajik	Russian
1	Rural	2.0	
	Urban	2.2	1.3
	Total (out of the possible score: Tajik 5; Russian 4)	2.1	1.3
2	Rural	1.9	1.1
	Urban	2.4	2.4
	Total (out of the possible score: Tajik 5; Russian 5)	2.0	2.3

SCHOOL TYPE COMPARISON

Tajik

When students' performance in reading comprehension was compared by school type, statistically significant differences between rural and urban schools were observed for Grade 2 (Table 41). In Tajik, students in urban schools (2.2 for Grade 1 and 2.4 for Grade 2 out of a possible score of 5) outperformed students in rural schools (2.0 for Grade 1 and 1.9 for Grade 2).

Russian

In Russian, regardless of school type, students obtained the highest scores of 1.3 out of 4 in urban Grade 1 (no sample from rural schools) and 2.4 out of 5 in urban Grade 2 (as compared to 1.1 for rural). The differences between rural and urban schools were found statistically significant for Grade 2 (Table 41).

LISTENING COMPREHENSION

The eighth section of the EGRA assessed listening comprehension. The purpose of this section was to examine whether students in Grades 1 and 2 were able to comprehend the passage they had just heard. This section included one short paragraph (~25 words for Grade 1 and ~40 words for Grade 2). This was a listening exercise. The test administrator read a passage aloud to the students only once, slowly (about one word per second) and then asked them 4 to 5 oral comprehension questions related to the passage. In Tajik, there were five questions each for Grades 1 and 2 related to their respective listening passages. In Russian, there were four questions for Grades 1 and 2.

Students performed relatively better in listening comprehension than in reading comprehension. About 27 percent of Grade 1 and 11 percent of Grade 2 students obtained perfect scores in Tajik listening comprehension, whereas 29 percent of Grade 1 and 15 percent of Grade 2 students obtained perfect scores in Russian listening comprehension.

TABLE 42: LISTENING COMPREHENSION, BY GENDER

Grade	Gender	Tajik	Russian
1	Male	3.2	2.2
	Female	3.3	2.3
	Total (out of the possible score: Tajik 5; Russian 4)	3.2	2.3
2	Male	2.5	2.0
	Female	2.4	2.1
	Total (out of the possible score: Tajik 5; Russian 4)	2.5	2.1

GENDER COMPARISON

Tajik

Overall, students performed relatively better in the listening comprehension than in the reading comprehension section. In Tajik, students in Grades 1 and 2 obtained an average score of 3.2 (3.2 for males, 3.3 for females) out of possible scores of 5 and 2.5 (2.5 for males, 2.4 for females) out of 5, respectively (Table 42). About 5 percent of Grade 1 and 9 percent of Grade 2 students obtained a score of zero, while 27 percent of Grade 1 and 11 percent of Grade 2 students secured perfect scores in their respective listening comprehension sections. When performance differences between male and female students were compared, statistically significant differences were only observed for Grade 2 students.

Russian

In Russian, students in Grades 1 and 2 obtained average scores of 2.3 (2.2 for males, 2.3 for females) out of a possible score of 4 and 2.1 (2.0 for males, 2.1 for females) out of 4 (Table 42). A total of 19 percent of Grade 1 and 15 percent of Grade 2 students obtained a score of zero. By contrast, 29 percent of Grade 1 and 15 percent of Grade 2 students received perfect scores in their respective listening comprehension sections. Although female students performed relatively better than male students in both grade levels did, the difference was not statistically significant.

SCHOOL TYPE COMPARISON

Tajik

In Tajik, students in Grades 1 and 2 urban schools obtained average scores of 3.7 out of 5 (as compared to 3.0 in rural schools) and 2.6 out of 5 (as compared to 2.4 in rural schools), respectively. When students' performance in listening comprehension was compared by school type, students in urban schools outperformed students in rural schools in both grade levels, and the differences were found statistically significant.

TABLE 43: LISTENING COMPREHENSION, BY SCHOOL TYPE

Grade	Type	Tajik	Russian
1	Rural	3.0	
	Urban	3.7	2.3
	Total (out of the possible score: Tajik 5; Russian 4)	3.2	2.3
2	Rural	2.4	1.3
	Urban	2.6	2.1
	Total (out of the possible score: Tajik 5; Russian 4)	2.5	2.1

Russian

In Russian, students in Grade 2 in rural and urban schools obtained average scores of 1.3 and 2.1 out of possible scores of 4, respectively (Table 43). Within Grade 2, students in urban schools performed statistically significantly better than their counterparts in rural schools. The comparison for Grade 1 was not possible because students from rural Russian schools were not selected in the sample.

DICTATION

The ninth section of the EGRA assessed dictation. The purpose of this section was to examine whether students in Grades 1 and 2 were able to write a complete sentence correctly using appropriate formation, size, signs, symbols, and spacing. The dictation sentence for Grades 1 and 2 consisted of 4 to 5 words and 5 to 6 words, respectively. In Tajik, the maximum possible scores for Grades 1 and 2 were 12 and 16, respectively. In Russian, the maximum possible scores were 16 for Grade 1 and 18 for Grade 2.

Students performed well in writing. About 31 percent of Grade 1, 34 percent of Grade 2 students in Tajik, nine percent of Grade 1, and 15 percent of Grade 2 students in Russian obtained perfect scores in their respective writing sections.

TABLE 44: DICTATION, BY GENDER

Grade	Gender	Tajik	Russian
1	Male	8.4	11.4
	Female	8.9	11.4
	Total (out of the possible score: Tajik 12; Russian 16)	8.7	11.4
2	Male	12.8	13.0
	Female	13.7	13.8
	Total (out of the possible score: Tajik 16; Russian 18)	13.3	13.4

GENDER COMPARISON

Tajik

Overall, students performed well in dictation. Students in Grades 1 and 2 obtained average scores of 8.7 (73 percent) out of a possible score of 12 and 13.3 (83 percent) out of 16, respectively (Table 44). About 9 percent of Grade 1 and 2 percent of Grade 2 students obtained a score of zero in dictation. More than one-third of Grade 1 (31 percent) and Grade 2 (34 percent) students secured perfect scores. Female students performed statistically significantly better than their male peers in both grade levels did. The female students obtained average scores of 8.9 in Grade 1 (as compared to 8.4 for males) and 13.7 in Grade 2 (as compared to 12.8 for males), respectively.

Russian

Students' performance in Russian dictation was very similar to that of students in Tajik. Students in Grades 1 and 2 obtained average scores of 11.4 (71 percent) out of a possible score of 16 and 13.4 (74 percent) out of 18. Only 1 percent of each of Grade 1 and 2 students obtained a score of zero in writing. On the other hand, 9 percent of Grade 1 and 15 percent of Grade 2 students received perfect scores in their respective writing sections. When compared by gender, a similar pattern was observed for Russian as was observed for Tajik. Female students outperformed male students in Grade 2, and differences were statistically significant. However, there was no difference between male and female in Grade 1. The male and female students obtained average scores of 11.4 for Grade 1 and 13.0 and 13.8 for Grade 2, respectively (Table 44).

TABLE 45: DICTATION BY SCHOOL TYPE

Grade	Type	Tajik	Russian
1	Rural	8.4	
	Urban	9.2	11.4
	Total (out of the possible score: Tajik 12; Russian 16)	8.7	11.4
2	Rural	13.2	8.7
	Urban	13.7	13.6
	Total (out of the possible score: Tajik 16; Russian 18)	13.3	13.4

SCHOOL TYPE COMPARISON

Tajik

Urban schools performed relatively better than their peers in rural schools across both grade levels, and statistically significant differences were observed for both Grades 1 and 2. Students in Grades 1 and 2 in urban schools obtained average scores of 9.2 out of 12 (as compared to 8.4 in rural) and 13.7 out of 16 (as compared to 13.2 in rural), respectively (Table 45).

Russian

A similar pattern was observed for writing in Russian as was observed for Tajik. Students in Grade 2 from urban schools performed statistically significantly better than their counterparts in rural schools did. However, the comparison for Grade 1 was not possible because students from rural Russian schools were not selected in the sample. In Grade 2, students in rural and urban schools obtained average scores of 8.7 and 13.6 out of a possible score of 18 (Table 45).

VI. RECOMMENDATIONS

The MOES efforts to improve the quality of education in Tajikistan and their attention to literacy improvement are significant. The baseline data collection and analysis efforts in both the Tajik and Russian EGRA are an essential first step toward ensuring that the strengths, weaknesses, and challenges of early grade reading are documented and understood so that appropriate interventions can be determined to efficiently and effectively improve student learning outcomes in literacy. The results of the Tajik and Russian EGRAs (USAID, 2012) showed that a majority of the students in Grade 4 could not read at their grade level. It was also evident that almost 41 percent of the students in both grades were not able to read at national standards related to reading fluency, the only available benchmark of reading skills. The outcomes in reading comprehension indicated that students performed better on literal questions than inferential questions, indicating difficulty with reading comprehension and critical understanding of text, an indicator of functional literacy. It is no surprise that similar results have been found in the baseline of the USAID Quality Reading Project.

Research shows that literacy skills in one's first language (e.g., visual awareness, phonemic awareness, and reading speed) support reading in any other language. The cognitive and linguistic skills attained while learning to read in the first language need not be relearned in the second language (however, the orthographic specific rules of the second or third language may need specific instruction). This suggests that MOES efforts to improve Tajik literacy will also support improvements in students' Russian language literacy. Further recommendations to improve literacy in early grades follow.

READING AND LITERACY TEACHING METHODS AND STRATEGIES: It is clear from the results that students are not learning the skills necessary to read or comprehend what they hear in the Tajik and Russian languages. Although this assessment was not designed to determine the effect IST had on student achievement, ample research demonstrates that teacher subject knowledge and participation in IST have a positive impact on student performance. Therefore, teachers at both the pre-service and in-service levels need to be introduced to proven research-based methods and strategies for teaching students the languages. The USAID Quality Reading Project has assisted the MOES in developing and implementing IST training packages for teaching students reading; the training is currently underway. Research also shows that teachers need long-term guidance and support to sufficiently understand and utilize information received in IST programs. Additional IST and CPD activities should take place to ensure that teachers continue to develop their skills in using these methodologies in the classroom on a daily basis. Furthermore, these strategies and methods should be incorporated into the pre-service curriculum of every TTI so that all graduating teachers are equipped with the methodologies they will need to be successful in teaching literacy skills to their students.

READING MATERIALS: It was revealed from the baseline EGRA that students have difficulty with reading fluency and comprehension. This suggests that students do not have an adequate opportunity to practice reading due to a lack of materials. Significant efforts need to be made to procure, develop, and distribute quality reading materials and teaching aids so that students and teachers have easy access to materials to increase teaching and learning literacy outcomes. One component of the USAID Quality Reading Project, "development and procurement of low-cost materials," supports schools in the development of low-cost reading materials.

READING ENVIRONMENTS: It is recommended that children are provided opportunities to interact with books and texts at an early age. Children begin to acquire some basic understandings of the concepts about literacy and language and its functions long before they

are able to read and write. This is the beginning of learning to read and write. Preschool and home environments can make a large difference in children's reading and writing readiness before they enter Grade 1.

STANDARDS-BASED EGRAS: The current National Strategy for Education Development (2012–2020) focuses on universal access and quality education. The three main priorities are to (i) modernize the curricula, (ii) reorganize the education system, and (iii) ensure equal access to quality education. As planned, Tajikistan has made significant progress toward modernizing the curricula (resulting in reforming a standards or competency-based education). Moreover, the USAID Quality Reading Project developed national reading standards and an aligned IST training package for the primary grades in 2014, and it was later approved by the AOE and the MOES. Additionally, the project developed EGRAs aligned with reading standards and performance benchmarks in order to measure student progress against those benchmarks. Therefore, it is expected that future EGRAs must employ a standards-based approach and make a strong connection to the baseline EGRA when constructing future EGRA tests, analyzing the data, and reporting the progress through a psychometric procedure called "test score equating." The test equating procedure would help by bringing both baseline and follow-up tests onto the same measurement scale and answering the question, "If Student X is taking the test in 2015 (follow-up), what would his or her score have been in 2014 (baseline) if he or she had taken the test in 2014?" If the baseline and future EGRA are not linked, then the improvement or decline in student performance would not be measured on the same scale and would be due to the difference in the difficulty level of the tests. A standards-based approach ensures that the results of the assessment provide relevant data on student performance according to the curriculum standards established by the MOES.

FORMATIVE ASSESSMENT LITERACY TOOLS: Summative assessments such as this EGRA are important for providing information on a systemic level. However, teachers need to be able to assess student learning on a daily basis and must be able to adapt their lesson plans and methodologies according to what provides the greatest results for their students. Simple formative assessment tools for literacy learning should be developed and incorporated into pre-service, in-service, and CPD training programs so teachers can better understand student learning in the classroom and adjust their lesson planning accordingly.

VII. REFERENCES

Baker, C. (2001). *Foundations of bilingual education and bilingualism* (3rd ed). Clevedon, UK: Multilingual Matters.

Beck, M. (2003). *Standard setting: If it is science, it's sociology and linguistics, not psychometrics.* Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, IL.

Cizek, G., & M. Bunch. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications, Inc.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3–49). Los Angeles, CA: Evaluation, Dissemination and Assessment Center, California State University.

Hambleton, R. K., & van der Linden, W. J. (1982). Advances in item response theory and applications: An introduction. *Applied Psychological Measurement, 6*, 373–378.

Hamilton, L. S., Stecher, B. M., & K. Yuan. (2008). *Standards-based reform in the United States: History, research, and future directions.* Los Angeles, CA: RAND Corporation.

Linn, R. L. (2001). *The design and evaluation of educational assessment and accountability systems* (CSE Technical Report No. 539). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Linn, R. L., & Herman, J. L. (1997a). *A policymaker's guide to standards-led assessment.* Denver, CO: The Education Commission of the States.

Linn, R. L., & Herman, J. L. (1997b). *Standards-led assessment: Technical and policy issues in measuring school and student progress* (CSE Technical Report No. 426). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Livingston, S. A., & Zieky, M. J. (1982). *Passing scores: A manual for setting standards of performance on educational occupational tests.* Princeton, NJ: Educational Testing Service.

Loomis, S. C., & Bourque, M. L. (2001). From tradition to innovation: Standard setting on the National Assessment of Educational Progress. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 175–217). Mahwah, NJ: Erlbaum.

Lord, F. M. (1950). *Notes on comparable scales for test scores* (ETS RB-50-48). Princeton, NJ: Educational Testing Service.

Perie, M. (2008). A guide to understanding and developing performance-level descriptors. *Educational Measurement: Issues and Practices,* 27(4), 15–29.

Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming and equating. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.) (pp. 221–262). New York, NY: Macmillan.

Plake, B. S., & Ferdous, A. A. (2005). *Setting multiple performance standards using the yes/no method: An alternative item mapping method*. Paper presented to the meeting of the National Council on Measurement in Education, Montreal, Canada.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. In B. D. Wright (Ed.). Chicago, IL: The University of Chicago Press.

United Nations Children's Fund. (2005). *Monitoring learning achievement: Nationwide study of the quality of education in primary schools*.

United States Agency for International Development. (2012). *Review of early grade reading: Teaching and skills*. The Kyrgyz Republic and Tajikistan.

U.S. Agency for International Development
1300 Pennsylvania Avenue, NW
Washington, DC 20523
Tel: (202) 712-0000
Fax: (202) 216-3524
www.usaid.gov