

Baseline Report:

Interactive Online Platform for Arabic Early Grade Literacy

Little Thinking Minds, Jordan

Prepared by:
School-to-School International and Integrated Services Indigenous Solutions
For All Children Reading: A Grand Challenge for Development

February 2016

Table of Contents

I. Executive Summary	3
Key Findings.....	3
II. Project Description	5
III. Purpose.....	5
IV. Evaluation Design and Methodology	5
Instrument Development.....	6
Sampling Population.....	7
V. Fieldwork Preparation and Data Collection.....	7
Assessor Training.....	7
Inter-rater Reliability (IRR) Test	8
Institutional Review Board for Human Participants (IRB).....	8
Data Analysis	8
VI. Summary of Findings.....	9
VII. Results by Group and Gender.....	11
Letter Sound Knowledge	11
Syllable Reading	13
Non-word Reading.....	15
Oral Reading Fluency	17
Reading Comprehension.....	19
Listening Comprehension	21
VIII. Contextual Factors	22
IX. Conclusions.....	24
X. Recommendations	24
Recommendations for Program Implementation	24
Recommendations for the Education Sector at Large.....	25
XI. Annex 1: Instruments.....	26
Student Survey	26
XII. Annex 2: List of Schools	28
XIII. Annex 3: Student Characteristics	29
XIV. Annex 4: Descriptive Statistics and Scores by Group and Gender.....	35
XV. Item Level Reliability	49

I. Executive Summary

Recognizing that literacy is fundamental to learning, skill acquisition, and success in primary school and beyond, education stakeholders are increasing their focus on the assessment of early grade reading skills. The Early Grade Reading Assessment (EGRA) is an oral student assessment designed to measure the most basic foundational skills for literacy acquisition in the early grades: recognizing letters of the alphabet, reading simple words, understanding sentences and paragraphs, and listening with comprehension.¹ The EGRA methodology was developed under EdData II, and has been applied in more than 30 countries and 60 languages.²

All Children Reading (ACR): A Grand Challenge for Development, a joint partnership between the United States Agency for International Development (USAID), World Vision, and the Australian Government, Department of Foreign Affairs and Trade (DFAT), has adopted the standard EGRA to systematically assess reading skills across all Round 2 grantees. The instrument is adapted according to each grantee's project context.

Little Thinking Minds, an ACR round 2 grantee conducted an EGRA baseline assessment in 20 public schools in collaboration with their local partner, Integrated Services Indigenous Solutions, and School-to-School International (STS). These schools represent the Jordan Education Initiative (JEI) discovery schools that are participating in the ACR-funded Interactive Online Platform for Arabic Early Grade Literacy project in Jordan.

The results of the baseline data collection, conclusions, and recommendations are presented in this report. Below is a summary of the key findings.

Key Findings

- Overall, students' *knowledge of letter sound* correspondence was low. The average score for the letter sound subtask was 23.2 letter sounds correctly identified in one minute in experimental schools and 22.2 letter sounds correct per minute in control schools. Furthermore, almost a quarter (24.3%) of students could not identify a single letter sound correctly. Such low scores indicate a low-level of skill mastery, which is reinforced by a recent national survey in Jordan³ in which the average score was 26.4 letter sounds per minute and approximately one quarter of Jordanian students scored zero on this subtask. Compared with other countries in the region, the control and

¹ RTI International and International Rescue Committee. (2011). *Guidance Notes for Planning and Implementing Early Grade Reading Assessments*.

² USAID EdData II. Available at: <https://www.eddataglobal.org/reading/>

³ RTI International (2013). Student performance in reading and mathematics, pedagogic practice, and school management in Jordan (EdData II Task Order No. 16). USAID / Jordan: Amman.

treatment schools scored lower than the West Bank and on par with students in Morocco.⁴

- **Syllable identification** was a challenge for students. In one minute, the average experimental school student correctly identified 15.8 syllables and the average control school student identified 12.6. Nearly a third of students were unable to read a single syllable.
- Decoding words proved most challenging for the students as evidenced by the scores on the **non-word reading subtask**. On average, students correctly decoded 4.4 and 4.1 respectively, words per minute in the experimental and control schools. Up to half of the students in the control group and one-third in the experimental group scored zero on this subtask. In Jordan's national EGRA,⁵ 47% of students scored zero on this subtask, and correctly decoded an average of 5.7 words per minute, indicating a low level of decoding skills. For this subtask the control and treatment school scores were lower than the West Bank, Morocco and Egypt and were on par with students in Iraq.⁶
- Approximately one quarter of students recorded zero-scores in the **oral reading fluency** subtask. Students were able to correctly read 8.3 words per minute in experimental schools versus 7.4 words per minute in control schools. Treatment and control students scored lower than Jordan's national mean (15.2 words) and lower than all other countries reporting Arabic EGRA scores (i.e. Egypt, Morocco, Iraq, Yemen and the West Bank).
- On average, 76.4% and 77.3% of students in experimental and control schools were unable to answer any question correctly in the **reading comprehension** subtask. This is substantially higher than the results from Jordan's national survey⁷ which shows that 25% of students were unable to respond correctly to a single comprehension question.
- Lastly, students' **listening comprehension** of spoken modern standard Arabic was quite low. On average, students correctly responded to two of five listening comprehension questions based on the short story. Only 10.2% of students in

⁴ EdData, <https://www.eddataglobal.org/countries/index.cfm>

⁵ RTI International (2013). Student performance in reading and mathematics, pedagogic practice, and school management in Jordan (EdData II Task Order No. 16). USAID / Jordan: Amman.

⁶ EdData, <https://www.eddataglobal.org/countries/index.cfm>

⁷ RTI International (2013). Student performance in reading and mathematics, pedagogic practice, and school management in Jordan (EdData II Task Order No. 16). USAID / Jordan: Amman.

experimental schools and 11.1% in control schools were able to answer all five questions correctly. Meanwhile, 11.8% of students in experimental schools and 14.7% in control schools could not answer a single comprehension question correctly. In Jordan's national EGRA survey⁸ 12% of the students were unable to answer any question.⁹ The schools scored lower than West Bank and Iraq in this subtask.

II. Project Description

The Interactive Online Platform for Arabic Early Grade Literacy program in Jordan is a one-year program funded by the All Children Reading Round 2 grant. The program aims to increase early grade literacy skills in Arabic by offering self-paced, interactive, online reading materials that supplement ongoing classroom instruction. In order to assess the impact of the digital reading program on students' reading scores, the program is conducting an impact evaluation in 20 schools. EGRA data will be collected at the baseline and end line from 10 treatment schools and 10 control schools. Students in the treatment group will receive an online reading program supported by ICT interns, while the control group will proceed with traditional Arabic literacy instruction. The online reading program will enable students in the treatment group to access an array of digital library materials appropriate to their reading level.

III. Purpose

The purpose of the EGRA baseline assessment is to measure the level of reading skills for students in both the treatment and control schools prior to project inception. The baseline and end line EGRA aim to respond to the following research questions:

- How efficient is the program in improving literacy for Grade 2 students?
- What percentage of students are able to decode new words and improve their reading fluency?
- What proportion of students were able to enhance their literacy through the platform (compared with control schools)?

IV. Evaluation Design and Methodology

To measure results of the program, an Early Grade Reading Assessment (EGRA) is conducted in two phases: a baseline assessment and an end line assessment. The baseline was conducted at the beginning of the academic year in September 2015; therefore, it measured what students learned at the end of Grade 1 prior to entering Grade 2. The end

⁸ RTI International (2013). Student performance in reading and mathematics, pedagogic practice, and school management in Jordan (EdData II Task Order No. 16). USAID / Jordan: Amman.

⁹ The baseline EGRA survey consisted of five listening comprehension questions while the national survey consisted of six listening comprehension question.

line assessment is expected to take place at the end of the academic year, in May 2016, and will measure gains at the end of Grade 2.

The impact evaluation was designed to assess a cross-section of Grade 2 students in 20 schools. Twenty-seven primary schools were selected from a pool of 55 JEI discovery schools that met the following selection criteria:

- 1) Schools in similar socio-economic areas in Amman;
- 2) Schools in which no simultaneous or similar interventions were conducted in the last school year (Grade 1 or Grade 2); and
- 3) A 50% ratio of mixed schools and single-sex schools.

Prior to the EGRA assessment, 20 schools were randomly selected from the 27 identified schools. Upon completion of the EGRA baseline data collection, the 20 schools were randomized and classified as control and experimental schools. Although 20 schools agreed to implement the program if selected as an experimental school, two schools were unable to participate in the program for the following reasons:

- An afternoon shift targeting Syrian students was introduced to the school by the Ministry of Education (MOE), and thus there would be no capacity to implement the program as the teachers and classrooms will be occupied for the afternoon shift;
- The second school is currently facing changes in the school structure (a new principal and teacher), and the number of Grade 2 students is very low. In addition, those students were not all enrolled in the school the year before. They advised that they would like to work with their students before introducing a project to them.

Therefore, the final sample comprised 20 schools. The students that were assessed at baseline will be assessed at the end of the project to evaluate the success of the interventions, by comparing the results of treatment and control groups.

Instrument Development

A student demographics survey and EGRA instrument were developed for the baseline assessment. The EGRA tool covers the following six subtasks:

1. Letter sound knowledge
2. Syllable reading
3. Non-word reading
4. Oral reading fluency
5. Reading comprehension
6. Listening comprehension

These subtasks were chosen for a variety of reasons. First, to ensure that “core” reading skills are captured across all ACR projects, STS, in consultation with a literacy expert, determined that a minimum of four subtasks should be included across projects: letter name and/or letter sound knowledge non-word reading, oral reading fluency and reading comprehension. ACR grantees were encouraged to include other EGRA subtasks as well, depending on the nature of their intervention and language components. In the case of Jordan, the instrument was developed by RTI International (RTI) and adapted to the linguistic context. As a result the EGRA included both letter sound knowledge and syllable reading.

A demographics survey was developed to capture the characteristics of the students sampled. The survey collected information on students’ gender, age, education, reading habits, and parental involvement, in addition to any extracurricular activities attended. A copy of the demographic survey is available in Annex 1.

Sampling Population

The final sample size was 803 students in 20 public schools in Amman. The total sample population disaggregated by gender and group type is as follows:

Table 1: Total number of students assessed by Group and Gender

Group	No. boys	No. girls	Total
Control	256	166	422
Treatment	126	255	381
Total	382	421	803

At each school where the EGRA was administered, 50 students were randomly selected from all Grade 2 streams. If the school had fewer than 50 Grade 2 students, the assessment was administered to all students present in Grade 2. In order to capture gender differences in reading performance, an equal number of boys and girls were selected whenever possible.

V. Fieldwork Preparation and Data Collection

Assessor Training

EGRA supervisors and assessors participated in a five-day training hosted by RTI from August 9-13, 2015 in Jordan. During the training, the participants:

- Reviewed the EGRA principles and gained a comprehensive understanding of the EGRA instrument components;
- Practiced EGRA administration and scoring procedures;

- Practiced conducting the EGRA assessment on tablets;
- Became familiar with the roles and responsibilities of both supervisors and assessors in the field;
- Underwent an Inter-rater Reliability (IRR) test administration and scoring.

Each supervisor and assessor was provided with a tablet and stimulus sheet to utilize during the data collection phase.

Inter-rater Reliability (IRR) Test

Inter-rater reliability is a measure of reliability used to assess the degree to which different assessors agree in their assessment decisions. Inter-rater reliability tests ensured that the different assessors interpreted answers in the same way; assessors may disagree within an acceptable level (10%) and it will have minimal effect on the EGRA score for each student. These tests were done prior to the baseline data collection. The scores from the first test and the second were correlated in order to determine the degree of consistency in administration. All assessors, except for one who was excluded from fieldwork, met the 90% threshold for inter-rater reliability in EGRA administration.

Institutional Review Board for Human Participants (IRB)

The Institutional Review Board (IRB) is responsible for ascertaining the acceptability of proposed research in terms of institutional commitments and regulations, applicable laws, standards of professional conduct and practice, and ethical and societal norms. The IRB examines subject recruitment procedures, proposed remuneration, and the informed consent process. The Board also evaluates the potential risks and benefits to participants outlined in each protocol. For the purposes of the baseline assessment, the team relied on RTI's approval from RTI's Committee for the Protection of Human Subjects.

Data Analysis

The data was analyzed using STATA and Excel, which resulted in graphs and frequency tables. The EGRA results were analyzed as follows:

Table 2 Subtask and data analysis method

Subtask	Analysis
Letter-Sound Knowledge	The score for this subtask is the number of letter sounds a student reads correctly in one minute, a measure known as

	Correct Letter Sounds per Minute” (CLSPM). There are a total of 100 letters presented on the stimulus.
Syllable Reading	The score of this subtask is the number of syllables read correctly in one minute, a measure known as Correct Syllables per Minute (CSPM). There are a total of 100 syllables presented on the stimulus.
Non-Word Reading	The score for this subtest is a measure of the number of Correct Non-Words Read per Minute (CNWPM). There are a total of 50 words presented on the stimulus.
Oral Reading Fluency (ORF)	The score of this subtest is a measure of the number of Correct Words read per Minute (CWPM). There are a total of 52 words presented on the stimulus.
Reading Comprehension	This score is a measure of the number of questions answered correctly based on the passage read in the ORF subtask. The student is asked questions corresponding with the number of sentences in the passage s/he was able to read within one minute. Therefore, this subtask score reports total number of questions answered correctly out of total number attempted. There are a maximum of five questions on this subtask.
Listening Comprehension	This score is a measure of the number of questions answered correctly out of a total of five questions asked. The questions are based on a short passage read aloud to the student.

VI. Summary of Findings

This section presents the overall summary of findings from the EGRA baseline assessment for the total sample population (803 students). Figures 1, 2 and 3 display summary results for timed subtasks, untimed subtasks, and zero scores across all subtasks. Four out of the six subtasks were timed: letter sound, syllable reading, non-word reading, and oral reading fluency. All students’ scores on the timed subtasks were recorded at the one-minute mark. The reading comprehension and listening comprehension subtasks were untimed, and are reported in terms of average number of questions answered correctly out of a total of five questions (The number correct out of number attempted is reported in subsequent sections).

Figure 1: Summary of Results for Timed Subtasks

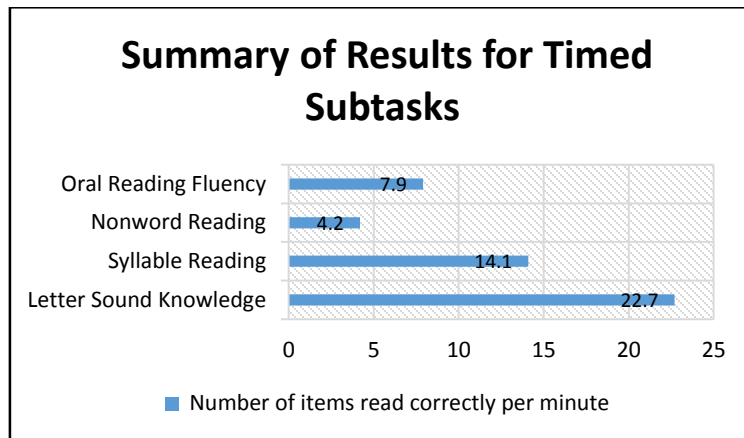
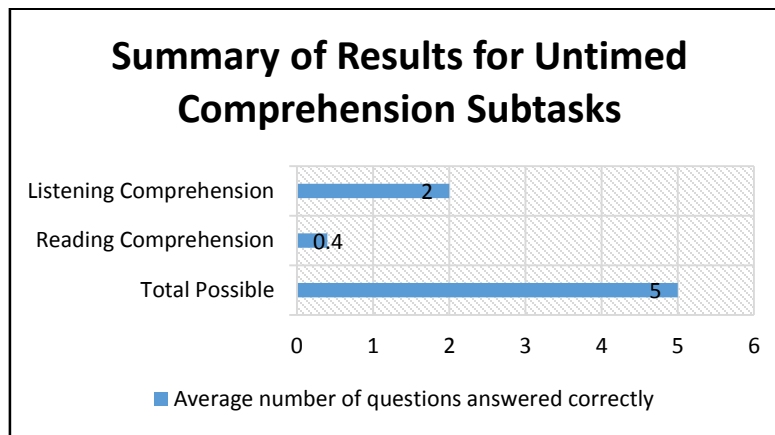


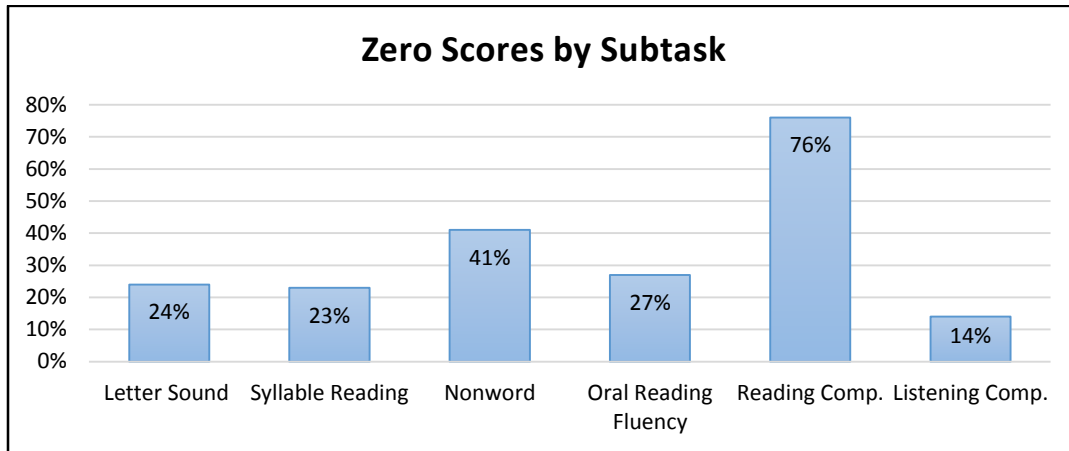
Figure 2: Summary of Results for Untimed Subtasks



Overall, Grade 2 students in both the experimental and control groups, totaling 803 students, scored relatively low across all subtasks. On the timed reading subtasks, students scored highest on letter sound knowledge subtask reading 22.7 letters correctly per minute. Students scored lowest on the non-word reading subtask with a mean score of 4.2 correct non-words read per minute indicating a lack of phonics and decoding skills. On the untimed subtasks, students, on average, correctly responded to two questions correctly on listening comprehension subtask and 0.4 questions correctly on reading comprehension subtask out of a total possible of five questions for each subtask. One-third of the students assessed could not read one word of the oral passage and two-thirds (615 out of 805 students) could not respond accurately to one comprehension question correctly mostly due to very few reading enough of the passage to be asked any

questions. Therefore, these findings indicate that the majority of students are unable to read and comprehend grade-level text.

Figure 3: Zero Scores by Subtask



VII. Results by Group and Gender

This section illustrates baseline results by subtask, group type, and gender. In each subsection, there is a description of the subtask followed by the mean scores, percentage of zero scores and distribution of responses. The results are displayed for males and females in both the experimental and control groups.

Letter Sound Knowledge

The letter sound subtask measures students' understanding of the "alphabetic principle" which states that each letter of the alphabet corresponds to a specific sound. Letter sound knowledge is defined as identifying appropriate sounds for each letter symbol. Recent studies suggest that the Arabic literacy orthographic system poses challenges since the Arabic script can represent multiple sounds depending on the position of the character within the text. For instance, letters can have three variations of sound depending on word placement.¹⁰ The letters are also differentiated by dots, or diacritical markings, that can change the letter sound into eight variations. These letter sound distinctions complicate reading acquisition. The ability to match letters with correct sounds is critical to reading fluency and comprehension. Thus, the first test of the EGRA examines the child's knowledge of letter sounds.

¹⁰ Maamouri, Mohammed. "Arabic Literacy," <http://papers.ldc.upenn.edu/EALL/ArabicLiteracy.pdf> (University of Pennsylvania, 1999), 3.

For this subtask, each student was presented with a stimulus of 100 letters written in the various positions in the word placement and were asked to read as many of the sounds as they could in one minute. After one minute, the student was asked to stop. There is an auto stop rule in all the timed EGRA subtasks. In this case the test was discontinued if a student was unable to correctly name any the first 10 letters on the stimulus.

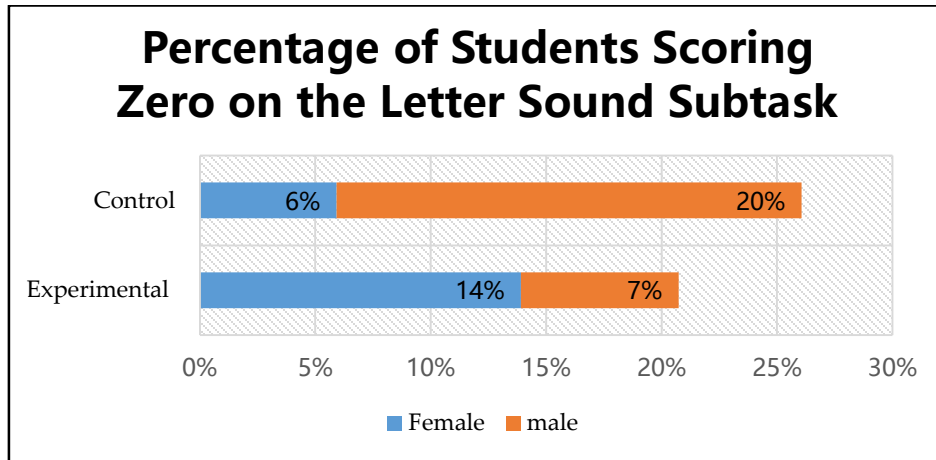
The mean scores for the letter sound subtask are presented in Table 4. The minimum score was 0 for both the experimental and control group, while the maximum was 83 in experimental schools and 86 in control schools. In experimental schools, Grade 2 students could correctly identify on average 23.2 letter sounds while students in the control schools were able to identify correctly 22.2 letter sounds. No significant differences in terms of gender were detected in the experimental schools; but in control schools, the mean score for male students was 17.7 which was significantly lower than that of their female classmates who has a mean of 29.1.

Table 4: Letter Sound Fluency by Treatment Group and Gender

Group	Gender	N	Mean (CLPM)	SD	Zero scores
Experimental	Boys	126	24.7	19.0	26
	Girls	255	22.4	18.8	53
Total		381	23.2	18.9	79
Control	Boys	256	17.7	19.7	85
	Girls	166	29.1	21.2	25
Total		422	22.2	21.0	110
Sample Total		803	22.7	20.0	189

Figure 4 illustrates that 20.7% of students in the experimental group were unable to identify a single letter sound compared to 26.1% of students in the control group. In the experimental group, there was a negligible difference between the percent of girls and boys who couldn't say one letter correctly. In the experimental group, 20.8% of girls and 20.6% of boys have a zero score. In contrast, the control group had a great disparity between the rate of zero scores with 33.2% of boys recording a zero score compared with only 15.1% of girls.

Figure 4 Percentage of students that could not identify a single letter sound



Syllable Reading

Syllables are the intermediate orthographic unit of language between letters and words. In terms of difficulty, the skill falls between recognizing phonemes and the ability to decode words. Testing children’s ability to read syllables identifies if the children have the foundational skills needed to read words. These abilities are important in language acquisition and provide the foundation for reading fluency and comprehension.

The syllable reading subtask consisted of 100 common syllables comprised of two letters or one letter and a diacritical marking or vowel which were randomly distributed in 10 rows. The syllables were derived from frequently used words for this age group. Syllables that were also words on their own in Arabic were omitted from the test. As with all the timed subtasks, the student was asked to read as many syllables as s/he could in one minute. After one minute, the student was asked to stop. The subtask was discontinued if a student was unable to correctly name any of the first 10 syllables.

Table 5 represents the results of the syllable reading subtask. The overall performance was similar across both the experimental and control group. On average, students in the experimental group attempted 23 syllables and were able to correctly identify 16, while those in the control group attempted an average of 21 syllables and correctly identified 13. There were no significant gender differences in the experimental group. However, boys performed significantly lower than girls in the control group.

Table 5: Syllable Reading Fluency by Treatment Group

Group	Gender	N	Mean (CSPM)	SD	Zero scores
Experimental	Boys	126	14.8	13.4	22
	Girls	255	16.3	12.1	35
Total		381	15.8	12.5	57
Control	Boys	256	9.5	12.2	99
	Girls	166	17.4	15.3	27
Total		422	12.6	14.1	126
Sample Total		803	14.1	13.4	183

The percentage of students that could not identify a single syllable is shown in Figure 5. Per the results, the control group had more than twice as many zero scores as the experimental group (30% compared with 15%). Moreover, the males in the control group performed significantly lower than males in the experimental group, with 24% scoring zero versus 6%, respectively.

Figure 5 Percentage of students that could not identify any syllable

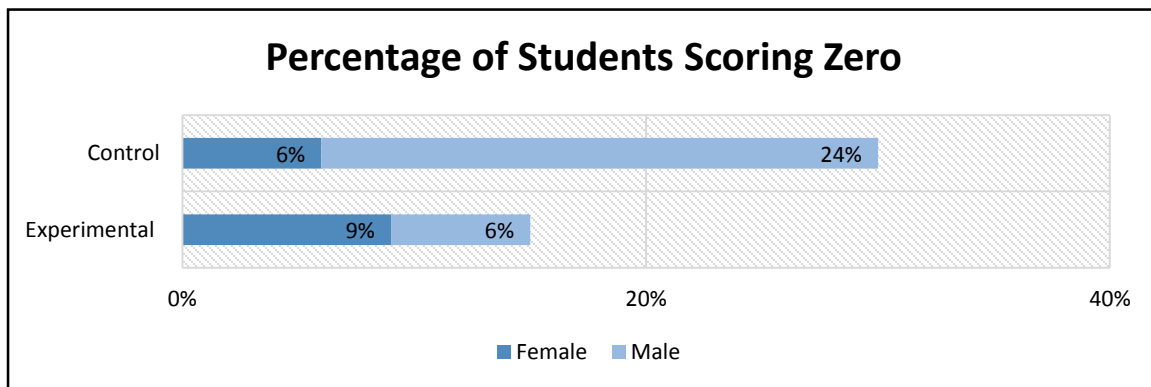
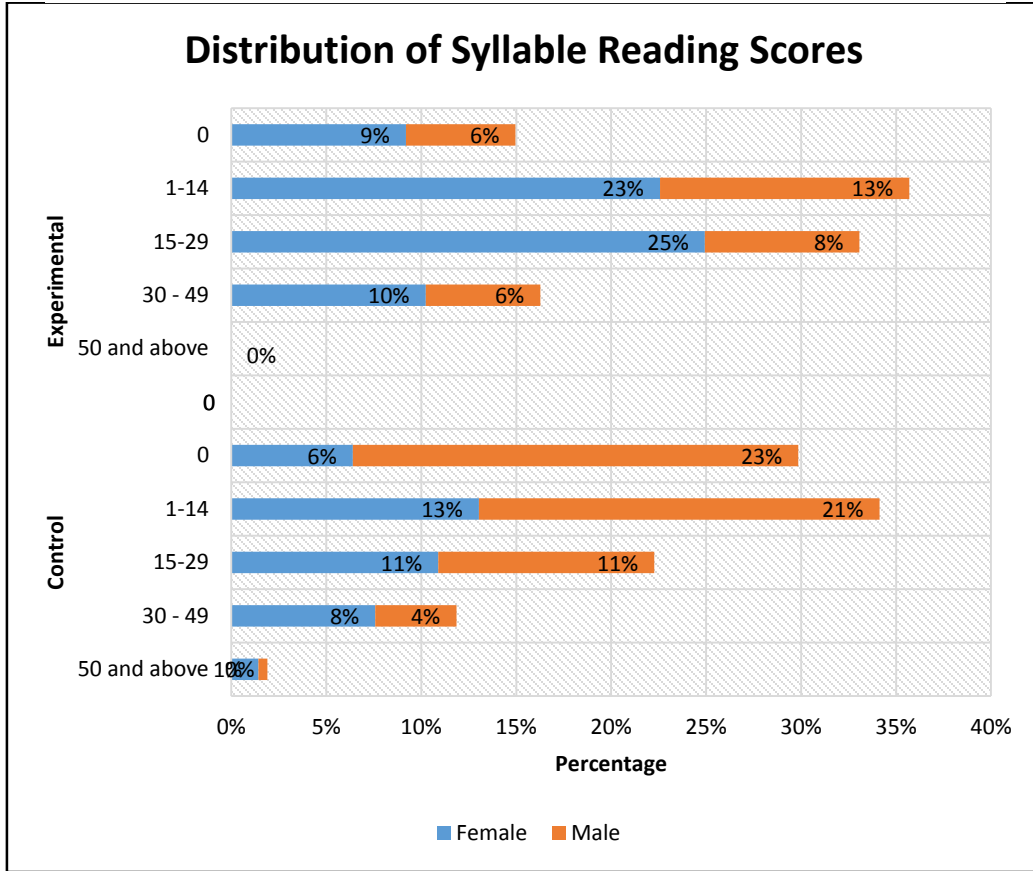


Figure 6 shows the range of student scores from 0 to 50 on the syllable reading subtask. The results are disaggregated by gender and school type. The majority of students in both groups were able to read between one to 14 syllables per minute. There was a greater percentage of experimental group students in the 15-29 range compared to the control group. The highest score for the experimental group was 49 syllables, while one percent of students in the control group read 50 or more syllables.

Figure 6 Distribution of Syllable Reading Scores



Non-word Reading

The non-word reading subtask is a measure of decoding ability and is designed to present children with words that they would not be able to recognize on sight through familiarity. Many children in the early grades learn to memorize or recognize a range of familiar words by sight alone. Thus, to assess children’s decoding skills, they are presented with invented (nonsense) words, which require them to sound out each letter and syllable to decode a word.

During this subtask, a child was presented with 50 non-words and asked to read as many as possible in one minute. After one minute, the student was asked to stop. (The subtask was discontinued if a student was unable to correctly read any the first 10 non-words).

Overall, the results were very low. In experimental and control schools, children could correctly decode an average of 4.2 words per minute. The highest score in the experimental group was a child who was able to correctly read 20 non-words per minute,

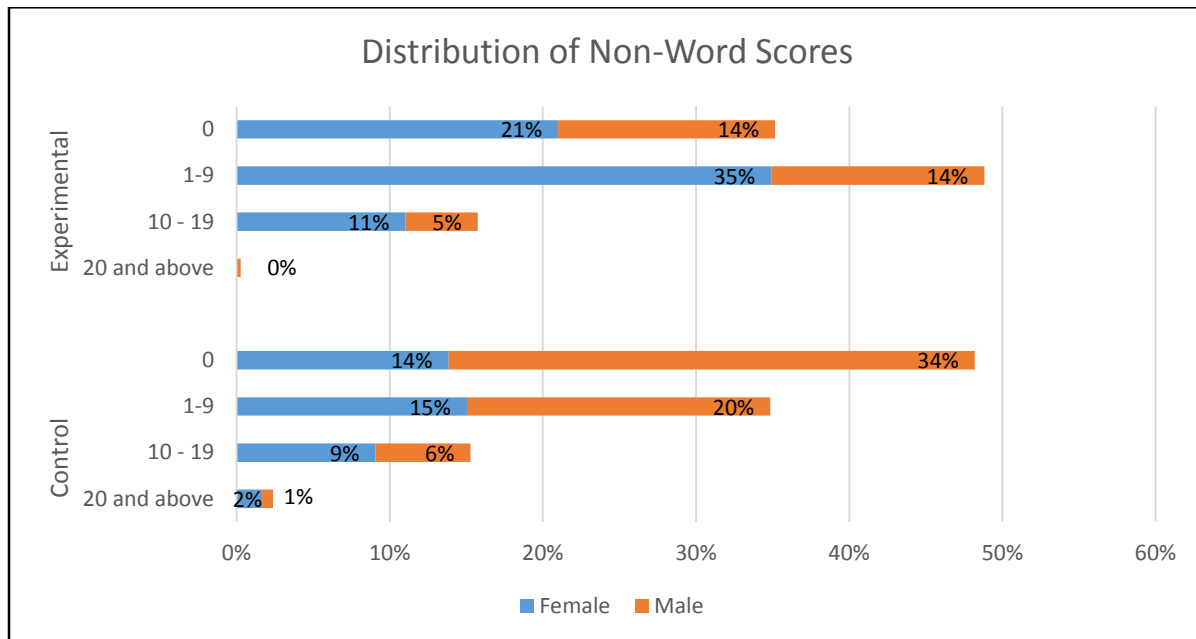
compared to a child from the control group who was able to correctly read 29 non-words per minute (See Table 6.) The gender findings on this subtask reflect previous subtasks. There was essentially no gender difference in the experimental group; while within the control group, boys performed 50% lower than girls.

Table 6: Non-word Fluency by Treatment Group and Gender

Group	Gender	N	Mean (CNWPM)	SD	Zero scores
Experimental	Boys	126	3.8	5.2	54
	Girls	255	4.7	4.8	80
Total		381	4.4	4.9	134
Control	Boys	256	3.1	5.0	144
	Girls	166	5.7	6.5	58
Total		422	4.1	5.8	202
Sample Total		803	4.2	5.4	336

The distribution of mean scores, ranging from 0 to 29, is shown by gender and school type in Figure 7. The results indicate a high proportion of zero scores for both groups on this subtask. In control schools, 47.9% of the students and in experimental schools 35.2% were not able to decode a single non-word. This implies that these students are not able to decode words that they have not previously encountered. Among the students who are able to decode a non-word, the majority could read between only one and nine non-words. Only 15.5% of students in both groups could read between 10 and 19 non-words.

Figure 7 Distribution of Non-Word Scores



Oral Reading Fluency

Oral reading fluency (ORF) is a measure of overall reading competence: the ability to translate letters into sounds, unify sounds into words, process connections, relate text to meaning, and make inferences to fill in missing information.¹¹ A child’s ORF score is dependent on the skills in previous subtasks, since children need to have some mastery of letter sounds, phonics and decoding of non-words in order to read fluently. Furthermore, because Arabic script employs diacritical markings (e.g., vowels, shaddah, hamza), the reader must be able to guess the word from both memory and sentence context. An empirical research study undertaken in Abu Dhabi and Palestine on the role of diacritics for beginning readers reveals that “diacritical markings were found to significantly influence the reading of both poor and skilled readers.”¹² The research study also found that “both skilled and poor readers improved their reading accuracy when they read vowels correctly [diacritical markings].” The diacritical markings assist students with identifying the sound of the letter; thus, being able to read diacritical markings largely affects the reader’s ability to correctly read the letter sound. As literacy progresses and students begin to read with automaticity, students tend to overlook vowels. Thus, students’ reading performances can falter until they are able to retain the

¹¹ Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *International Reading Association*, 636–644.

¹² Salim Abu-Rabia. (1999). “The effect of vowels on the reading comprehension of second and sixth-grade native Arab children.” *Journal of Psycholinguistics Research*. 28, 999993-101.

diacritical markings in their working memory long enough to read with both speed and accuracy.

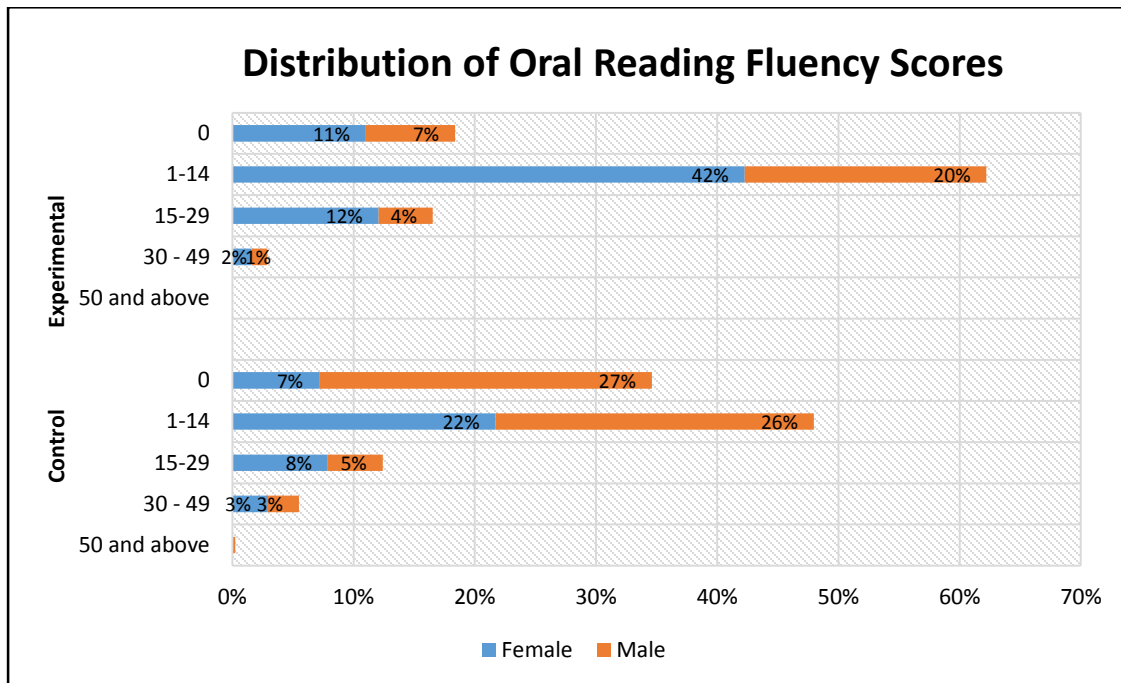
Table 7: Oral Reading Fluency by Treatment Group and Gender

Group	Gender	N	Mean (CWPM)	SD	Zero scores
Experimental	Boys	126	7.7	10.1	28
	Girls	255	8.8	9.3	42
Total		381	8.4	9.6	70
Control	Boys	256	5.6	9.9	115
	Girls	166	10.4	11.7	30
Total		422	7.5	10.9	145
Sample Total		803	7.9	10.3	215

In this EGRA subtask, students were asked to read aloud a 52-word passage with full diacritical markings. The results illustrated in Table 7 indicate that students are generally not reading with fluency in Grade 2. The minimum score in both groups was zero, while the maximum was 47 words in the experimental group and 50 words in the control group. On average, students were able to read eight correct words per minute in experimental schools and seven words per minute in control schools. Similar to other subtasks, boys in the control group read fewer words than girls, while there was virtually no difference among boys' and girls' reading ability in the experimental group.

The majority of students in both experimental and control schools were able to read between one and 14 words correctly per minute. Further analysis revealed that these students were able to read conjunctions in the first line of text but were not able to read any other words. Therefore, their vocabulary is limited to connective words such as and, in and on. As illustrated in Figure 8, there were 34.4% of students in the control group who could not read a single word. Meanwhile, only 18.4% of students in the experimental group could not read a single word.

Figure 8 Distribution of Oral Reading Fluency Scores



Reading Comprehension

The reading comprehension subtask identifies how well students understood the oral reading fluency passage. Upon completion of the ORF subtask, students were asked to answer up to five basic comprehension questions, which were read aloud by the assessor. Students were asked to attempt questions pertaining to the section of text they had read. For instance, if a student read the first line of text (10 words), s/he would be asked the first comprehension question. Similarly, if a student read all 52 words, s/he would be asked all five questions. Thus, for this subtask, the sample size is based on the number of students who were asked the questions, not all students. The zero scores in Table 8, however, reflect two types of students: students who read too little or nothing at all of the passage and those who read enough to be asked as least one comprehension question, but answered incorrectly.

The mean scores for all students are reflected in Table 8. On average, students attempted one comprehension question and the majority could not respond to any questions attempted. Overall, 76.4% of students in the experimental group and 77.3% in the control group did not answer any questions correctly. This percentage includes students who scored zero on the ORF subtask, representing 34.4% in the control group and 18.4% in the experimental group.

Table 8: Reading Comprehensions Questions Correct

Group	Number of Questions Correct	Girls		Boys		N Total	% Total
		N	%	N	%		
Experimental	0	186	72.9	103	81.8	289	75.9
	1	52	20.4	17	13.5	69	18.1
	2	13	5.4	2	1.6	15	3.9
	3	4	1.6	3	2.4	7	1.8
	4	0	0.0	1	0.8	1	0.3
	5	0	0.0	0	0.0	0	0.0
Total		255	100%	126	100%	381	100%
Control	0	105	63.3	221	86.3	326	77.3
	1	33	19.9	17	6.6	50	11.9
	2	21	12.7	12	4.7	33	7.8
	3	3	1.8	32	0.8	5	1.2
	4	3	1.8	24	1.6	7	1.7
	5	1	0.6	0	0.0	1	0.2
Total		166	100%	256	100.00%	422	100%

Table 8 shows the percentage of students in the sample who attempted each comprehension question and the percentage of correct responses. The sample size is less than 803 due to the high number of zero scores. A total of 756 students attempted the questions, and 188 were able to answer at least one question correctly. Based on the results below, the majority of students attempted one question, but were unable to correctly respond. Over 70% of students did not respond to any comprehension question correctly (n=615), and over 90% of students did not attempt questions 3 through 5.

Table 9 reports results for students who read one or more words of the passage correctly. When the students who scored zero are excluded from the analysis, the average number of questions attempted increases to two questions; however, the majority of students were still unable to respond to at least one question correctly. In fact, 70% of students in the experimental group and 65% in the control group could not accurately respond to one comprehension question. The high percentage of zero scores, in general across all students, indicates that the majority of students lack reading comprehension skills. There was no significant difference in terms of performance by gender or group type for this subtask.

Table 9: Mean reading comprehension questions answered correctly for children who read one or more words

Group	Gender	Mean Number of questions attempted	Mean Questions answered correctly	Percentage Students who did not answer any question correctly
Experimental	Boys	2	0.3	77
	Girls	1	0.4	68
Total		2	0.4	70
Control	Boys	2	0.4	75
	Girls	2	0.7	55
Total		2	0.6	65

Listening Comprehension

On this subtask, students listened to a short story read aloud by the assessor. They were asked to respond to five comprehension questions of varying difficulty related to the story. This subtask was a measure of listening comprehension, which is a pre-reading skill. As this subtask was untimed and all students heard the entire passage, they were asked all five questions. Thus, the results in Table 10 show the mean questions answered correctly for all five questions attempted.

On average, students in both experimental and control groups were able to answer two comprehension questions correctly. There was a low percentage of zero scores (12.1% in the experimental and 14.7% in the control group), which can be expected since this subtask does not require interaction with text and thus is fairly easy compared with the other subtasks.

Table 10: Listening Comprehensions Questions Correct

Group	Number of Questions Correct	Girls		Boys		N Total	% Total
		N	%	N	%		
Experimental	0	25	9.8	21	16.7	46	12.1
	1	51	20.0	19	15.1	70	18.4
	2	56	21.9	38	30.2	94	24.7
	3	48	18.8	26	20.6	74	19.4
	4	47	18.4	12	9.5	59	15.5
	5	28	11.9	10	7.9	38	10.0
Total		255	100	126	100%	381	100%

Control	0	21	12.7	41	16.0	62	14.7
	1	29	17.5	63	24.6	92	21.8
	2	24	14.5	54	21.1	78	18.5
	3	36	21.7	48	18.8	84	19.9
	4	27	16.3	32	12.5	59	13.9
	5	29	17.5	18	7.9	47	11.1
	Total	166	100	256	100%	422	100%

Table 10 also shows the percentage of responses correct per listening comprehension question. The range of responses across all five questions indicates student’s listening comprehension abilities were much higher than their reading comprehension abilities. Twenty percent of students could accurately respond to questions 1, 2 and 3; 15% responded to question 4; and 11.1% of students could answer all five comprehension questions correctly. However, the average number of questions answered correctly overall is still quite low given this is a pre-reading skill that should be mastered by the end of Grade 1. This indicates that while students have some listening comprehension abilities, and they understand more of what they hear than what they read, their listening comprehension skills are still not fully developed.

VIII. Contextual Factors

To better understand the student population participating in the study, the team conducted a demographic survey including background information regarding students’ use of libraries, current reading practices, and use of home language. This was the survey previously mentioned in the tool development section. These contextual factors were collected to better understand the sample population. Some of the findings might help explain the causes of low reading achievement:

Libraries: The survey found that all 18 schools had a library on the premises. However, only 24% of students borrowed books from the library.

Reading time: According to best practices, teachers should spend a minimum of 40 minutes per day teaching reading.¹³ This helps build reading fluency and decoding skills. Within the 18 sample schools, the majority of students (83%) are encouraged by a family member to read at home and 85% of all students surveyed reported to set aside time to read at home. However, only 31% of these same students report that they read on a daily basis. Of the students that do allocate some time to read at home, the role of the adult

¹³ Maamouri, Mohammed. “Arabic Literacy,” <http://papers.ldc.upenn.edu/EALL/ArabicLiteracy.pdf> (University of Pennsylvania, 1999), 3.

proved very important with 94% of student reporting they read aloud to an adult, and 78% of the students said they have an adult read aloud to them, as shown in Table 11.

Table 11 Frequency of reading at home

Frequency of reading	Percentage of students that read aloud to an adult at home	Percentage of students that have an adult read aloud to them at home
Percentage of students that allocate time to read at home	94%	78%
Number of times an adult is reading with the child at home:		
O– sometimes/rarely	38%	29%
Once a week	8%	10%
Twice/Three times a week	15%	10%
Daily	31%	28%
Don't Know/No response	2%	1%
Never	6%	22%

Extracurricular activities: Students were also asked to specify if they were currently participating in any extracurricular activities that assisted them with Arabic literacy. Twenty percent stated that they were attending Quran courses in teaching centers or Mosques, which helped them practice reading in Arabic at the time of the survey.

Type of Arabic spoken at home: Only 2% of the students surveyed spoke Arabic-Fusha at home, while 96% of students spoke non-standard (colloquial) Arabic at home and 1% used a mix of both. The remaining one percent refused to answer. This gap between use of Fusha (formal Arabic used in formal education) and the colloquial Arabic spoken at home may be a factor contributing to low learning achievement in schools, as all textbooks and stories are written in Arabic-Fusha. Thus, practicing formal Arabic-Fusha at home through speaking to an adult, online platforms and reading materials may help strengthen students' reading fluency skills¹⁴.

¹⁴ Z. Eviater and R. Ibrahim (edited by Mark Leikin, Mila Schwartz, Yishai Tobi (2011)) *Current Issues in Bilingualism: Cognitive and Socio-linguistic Perspectives*, Springer.

Based on the above findings, it can be concluded that promising reading practices (teaching reading at least 40 minutes a day, using standard Arabic, borrowing library books, etc.) are lacking within the sample population.

IX. Conclusions

The results of this EGRA baseline reveal that the majority of students at the beginning of Grade 2 do not have the foundational skills necessary to read fluently with comprehension. Students in both experimental and control groups scored low across all subtasks. While scores were highest on the letter sound subtask, they were particularly low on non-word and oral reading fluency-timed subtasks, indicating students' lack the ability to decode unfamiliar words. Students also performed very low on the reading comprehension subtasks. About one-quarter of students could not read a single word of the oral reading passage and three-quarters could not correctly respond to one reading comprehension question.

Mean scores were very similar between students in experimental and control schools with a difference of one to two points. There was, however, a statistically significant difference in performance by gender in the control group in which males scored consistently lower than females on the timed subtasks.

With regards to contextual factors, the majority of students were not engaged in activities that promote reading, such as borrowing library books, reading at least 40 minutes per day, or using standard Arabic at home.

X. Recommendations

The results of this EGRA baseline raise a number of issues worth considering in the effort to improve students' early grade reading abilities. The key considerations and recommendations are divided into two sections: a) recommendations at the program level, and b) recommendations for the education sector at large.

Recommendations for Program Implementation

- Introduce **letter sound recognition, syllable recognition, and decoding games** into the QYSAS literacy platform, in addition to a diagnostics test, which will help the teacher and students identify specific challenges students face with the Arabic language. The results indicate that students need more practice with letter sound and phonics so that they can translate those skills into reading fluently and ultimately comprehension.

- **Provide access to supplemental reading materials at home** that are both engaging and interactive. In doing so, it will ultimately lead to improved literacy because students will regularly and consistently attend literacy clubs and engage with the platform through a one-on-one interface.
- **Train teachers to teach reading:** Reading is a foundational learning skill and needs to be taught from the early grades. It is recommended that teachers be trained to teach components of reading which include: letter sounds, phonics, decoding, and reading and listen comprehension strategies from Grade 1.
- **Encourage parents to support/motivate their children's to read at least 40 minutes per day:** The home environment matters with regards to improving reading outcomes. Children who practice reading with an adult are able to read the passage clearly with only some mistakes. Parents can support and encourage their children to read by reading to them every day for at least 40 minutes. This provides an opportunity for students to practice with space and time to read at their own pace.

Recommendations for the Education Sector at Large

- **Establish national benchmarks for reading:** Encourage the MOE to create national benchmarks for each reading skill for the early grades. These benchmarks are critical for teachers to identify which students are progressing appropriately and which students may need additional home or school support. This understanding would be based on classroom activities designed to help students reach the level required.
- **Train teachers to assess the reading abilities of the students:** Teachers should be trained to assess their students reading skills against national benchmarks. Teachers' assessment practices should include continuous assessment rather than only focusing on end-of-term grades.
- **Target drop-out primary students:** Students who have dropped out of primary schools are often not equipped with early grade reading skills, thus organizations should collaborate to identify these students and enroll them in extracurricular literacy programs such as QYSAS.

XI. Annex 1: Instruments

Student Survey

Demographics	
Name	
EGRA student number	
School	
Area	
Gender	(<input type="checkbox"/>) Female (<input type="checkbox"/>) Male
Age	
How many brothers and sisters do you have?	

Section 1	
What grade were you in last year?	(<input type="checkbox"/>) Grade 1 (<input type="checkbox"/>) Grade 2 (<input type="checkbox"/>) Refused to answer
Did you go to Kindergarten before coming school?	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No (<input type="checkbox"/>) Refused to answer
What Language do you speak at home?	(<input type="checkbox"/>) Arabic – <i>fusha</i> (<input type="checkbox"/>) Arabic (<input type="checkbox"/>) Mix of both
How do you go to school?	(<input type="checkbox"/>) Walk <i>alone</i> to school (<input type="checkbox"/>) Walk with <i>siblings</i> to school (<input type="checkbox"/>) Walk with <i>colleagues</i> to school (<input type="checkbox"/>) Walk with an <i>adult family member</i> to school (<input type="checkbox"/>) Public transportation (<input type="checkbox"/>) Other _____ (<input type="checkbox"/>) Refused to answer

Section 2	
Do you have a library or reading class at school? (also ask if there is reading time)	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No (<input type="checkbox"/>) Don't know/refused to answer
Do you allocate time to read at home?	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No (<input type="checkbox"/>) Don't know/ refused to answer
How many times a week do you read out loud to an adult at home?	(<input type="checkbox"/>) Yes - <i>record number</i> (<input type="checkbox"/>) No (<input type="checkbox"/>) Don't know/refused to answer
Does anyone from home encourage you to read?	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
How many books a month do you read?	Please state number _____

	(<input type="checkbox"/>) Doesn't read (<input type="checkbox"/>) Don't know/refused to answer
Are you borrowing books from the library?	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
Do you take any extracurricular activities for reading?	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
<i>If yes please specify</i>	-----
Do you take any after school Arabic classes?	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No

Do you own any of the following at home?	
Radio	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
Television	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
Electricity	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
Computer	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
Is it connected to the internet	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
Tablet/Smart phone	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
Is it connected to the internet	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
Do you read books on the tablet / computer / smartphone?	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
Have you ever downloaded any educational material on the tablet / computer / smartphone?	(<input type="checkbox"/>) Yes (<input type="checkbox"/>) No
Thank them and move to EGRA	

XII. Annex 2: List of Schools

EXPERIMENTAL						
School	Governorate	Gender	Amman Area	Number of "grade 2" students per school	Number of EGRA tests implemented	Number of sections
Experimental 1	Amman	Female	East	39	35	1
Experimental 2	Amman	Mixed	East	50	48	3
Experimental 3	Amman	Female	West	26	25	1
Experimental 4	Amman	Mixed	West	350	53	7
Experimental 5	Amman	Male	West	48	46	2
Experimental 6	Amman	Female	West	90	51	2
Experimental 7	Amman	Mixed	West	38	38	1
Experimental 8	Amman	Mixed	West	105	55	2
Experimental 9	Amman	Mixed	East	60	41	2
Experimental 10	Amman	Female	East	60	41	2
CONTROL						
	Governorate	Gender	Amman Area	Number of "grade 2" students per school	Number of EGRA tests implemented	Number of sections
Control 1	Amman	Mixed	West	40	36	1
Control 2	Amman	Male	East	22	20	1
Control 3	Amman	Mixed	West	60	49	2
Control 4	Amman	Female	East	100	49	3
Control 5	Amman	Male	West	50	37	1
Control 6	Amman	Mixed	West	152	52	2
Control 7	Amman	Mixed	West	93	47	2
Control 8	Amman	Mixed	West	80	47	2
Control 9	Amman	Male	West	80	39	2
Control 10	Amman	Male	East	90	46	3

XIII. Annex 3: Student Characteristics

Control Schools

What grade were you in last year?	Percentage
Grade 1	98%
Grade 2	1%
Refused to answer	0%
Total	100%

Did you go to Kindergarten before coming school?	Percentage
Yes	77%
No	23%
Refused to answer	0%
Total	100%

What type of Arabic Language do you use at home?	Percentage
Arabic – fusha	3%
Arabic - non standard	96%
Mix of both	0%
Refused to answer	0%
Total	100%

How do you go to school?	Percentage
Walk alone to school	19%

Experimental schools

What grade were you in last year?	Percentage
Grade 1	98%
Grade 2	1%
Refused to answer	1%
Total	100%

Did you go to Kindergarten before coming school?	Percentage
Yes	83%
No	17%
Refused to answer	0%
Total	100%

What type of Arabic Language do you use at home?	Percentage
Arabic – fusha	2%
Arabic – non standard	97%
Mix of both	1%
Refused to answer	0%
Total	100%

How do you go to school?	Percentage
Walk alone to school	15%

Walk with siblings to school	27%
Walk with colleagues to school	9%
Dropped or walk with an adult family member to school	23%
Public transportation	9%
Other	13%
Refused to answer	0%
Total	100%

Walk with siblings to school	24%
Walk with colleagues to school	5%
Dropped or walk with an adult family member to school	37%
Public transportation	6%
Other	13%
Refused to answer	0%
Total	100%

Do you have a library or reading class at school?	Percentage
Yes	66%
No	31%
Don't Know/ Refused to answer	3%
Total	100%

Do you have a library or reading class at school?	Percentage
Yes	67%
No	28%
Don't Know/ Refused to answer	5%
Total	100%

Do you allocate time to read at home?	Percentage
Yes	84%
No	15%
Don't Know/ Refused to answer	0%
Total	100%

Do you allocate time to read at home?	Percentage
Yes	87%
No	13%
Don't Know/ Refused to answer	0%
Total	100%

How many times a week do you read out loud to an adult at home?	Percentage
Often	36.6%
Once a week	6.9%
Twice/Three times a week	11.4%

How many times a week do you read out loud to an adult at home?	Percentage
Often	28.6%
Once a week	7.7%
Twice/Three times a week	13.6%

Daily	24.0%
Don't Know/ Refused to answer	1.7%
Never	19.5%
Total	100%

Daily	29.7%
Don't Know/ Refused to answer	1.1%
Never	19.3%
Total	100%

How many times a week do you read out loud to an adult at home?	Percentage
Often	27%
Once a week	10%
Twice/Three times a week	9%
Daily	25%
Don't Know/ Refused to answer	1%
Never	29%
Total	100%

How many times a week do you read out loud to an adult at home?	Percentage
Often	23%
Once a week	7%
Twice/Three times a week	8%
Daily	24%
Don't Know/ Refused to answer	1%
Never	37%
Total	100%

Does anyone from home encourage you to read?	Percentage
Yes	84%
No	13%
Don't Know/ Refused to answer	3%
Total	100%

Does anyone from home encourage you to read?	Percentage
Yes	83%
No	17%
Don't Know/ Refused to answer	1%
Total	100%

Are you borrowing books from the library?	Percentage
Yes	20%
No	80%
Don't Know/ Refused to answer	0%
Total	100%

Are you borrowing books from the library?	Percentage
Yes	27%
No	73%
Don't Know/ Refused to answer	0%
Total	100%

Do you take any extracurricular activities for reading?	Percentage
Yes	21%
No	78%
Don't Know/ Refused to answer	1%
Total	100%

Do you take any extracurricular activities for reading?	Percentage
Yes	39%
No	60%
Don't Know/ Refused to answer	1%
Total	100%

Specify	Percentage
Quran	16%
Reads at home	1%
None	83%
Total	100%

Specify	Percentage
Quran	25%
Summer Camp	1%
None	74%
Total	100%

If yes please specify Do you take any Reading classes?	Percentage
Yes	9%
No	91%
Don't Know/ Refused to answer	0%
Total	100%

If yes please specify Do you take any Reading classes?	Percentage
Yes	10%
No	89%
Don't Know/ Refused to answer	0%
Total	100%

Do you own a Radio?	Percentage
Yes	29%
No	69%
Don't Know/ Refused to answer	2%
Total	100%

Do you own a Radio?	Percentage
Yes	36%
No	63%
Don't Know/ Refused to answer	1%
Total	100%

Do you own a TV?	Percentage
Yes	99%
No	1%
Don't Know/ Refused to answer	0%
Total	100%

Do you own a TV?	Percentage
Yes	99%
No	1%
Don't Know/ Refused to answer	0%
Total	100%

Do you have electricity at home?	Percentage
Yes	99%
No	1%
Don't Know/ Refused to answer	0%
Total	100%

Do you have electricity at home?	Percentage
Yes	100%
No	0%
Don't Know/ Refused to answer	0%
Total	100%

Do you own a computer?	Percentage
Yes	52%
No	48%
Don't Know/ Refused to answer	0%
Total	100%

Do you own a computer?	Percentage
Yes	57%
No	43%
Don't Know/ Refused to answer	0%
Total	100%

Is the computer connected to the internet	Percentage
Yes	58%
No	35%
Don't Know/ Refused to answer	8%
Total	100%

Is the computer connected to the internet	Percentage
Yes	62%
No	28%
Don't Know/ Refused to answer	11%
Total	100%

Do you own a tablet or a smartphone?	Percentage
Yes	76%
No	24%
Don't Know/ Refused to answer	1%
Total	100%

Do you own a tablet or a smartphone?	Percentage
Yes	83%
No	17%
Don't Know/ Refused to answer	0%
Total	100%

Is the tablet or smartphone connected to the internet?	Percentage
Yes	71%
No	23%
Don't Know/ Refused to answer	6%
Total	100%

Is the tablet or smartphone connected to the internet?	Percentage
Yes	69%
No	22%
Don't Know/ Refused to answer	9%
Total	100%

Do you read any books on the technological devices?	Percentage
Yes	21%
No	78%
Don't Know/ Refused to answer	1%
Total	100%

Do you read any books on the technological devices?	Percentage
Yes	31%
No	69%
Don't Know/ Refused to answer	0%
Total	100%

S3.9: Have you downloaded any educational materials on the technological devices?	Percentage
Yes	40%
No	59%
Don't Know/ Refused to answer	1%
Total	100%

Have you downloaded any educational materials on the technological devices?	Percentage
Yes	44%
No	55%
Don't Know/ Refused to answer	1%
Total	100%

XIV. Annex 4: Descriptive Statistics and Scores by Group and Gender

Sample

	Female	Male	Total
Control	166 (39.3%)	256 (60.7%)	422
Experimental	255 (66.9%)	126 (33.1%)	381
Total	421 (52.4%)	382 (47.6%)	803

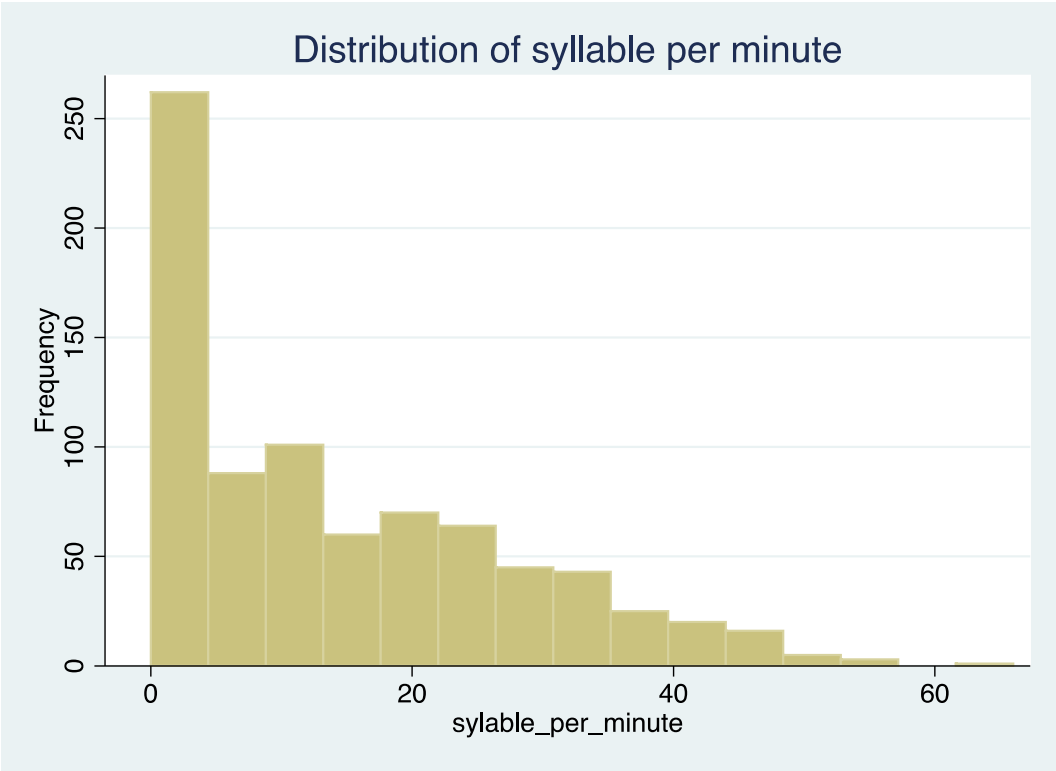
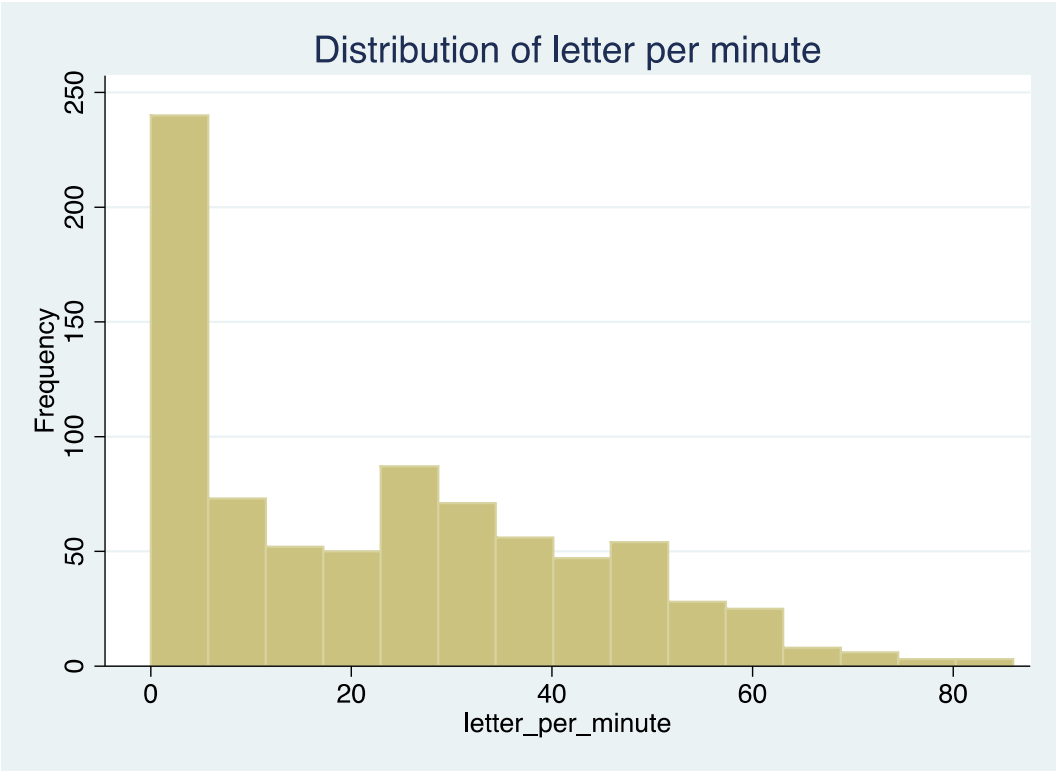
Timed subtask

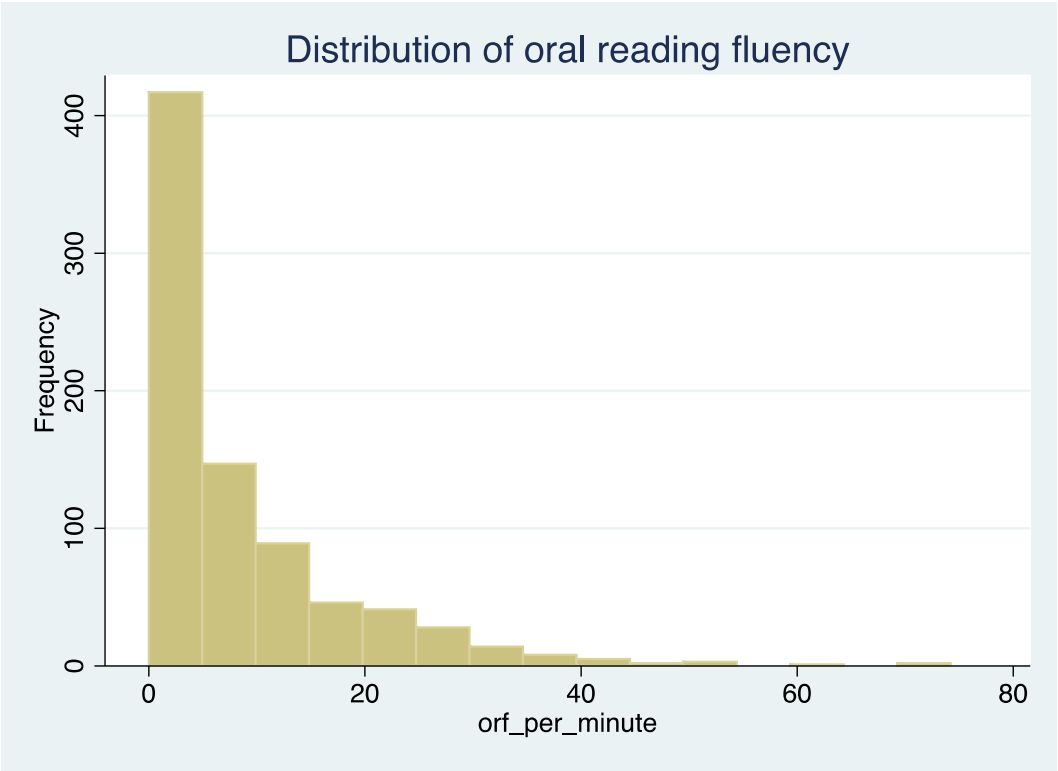
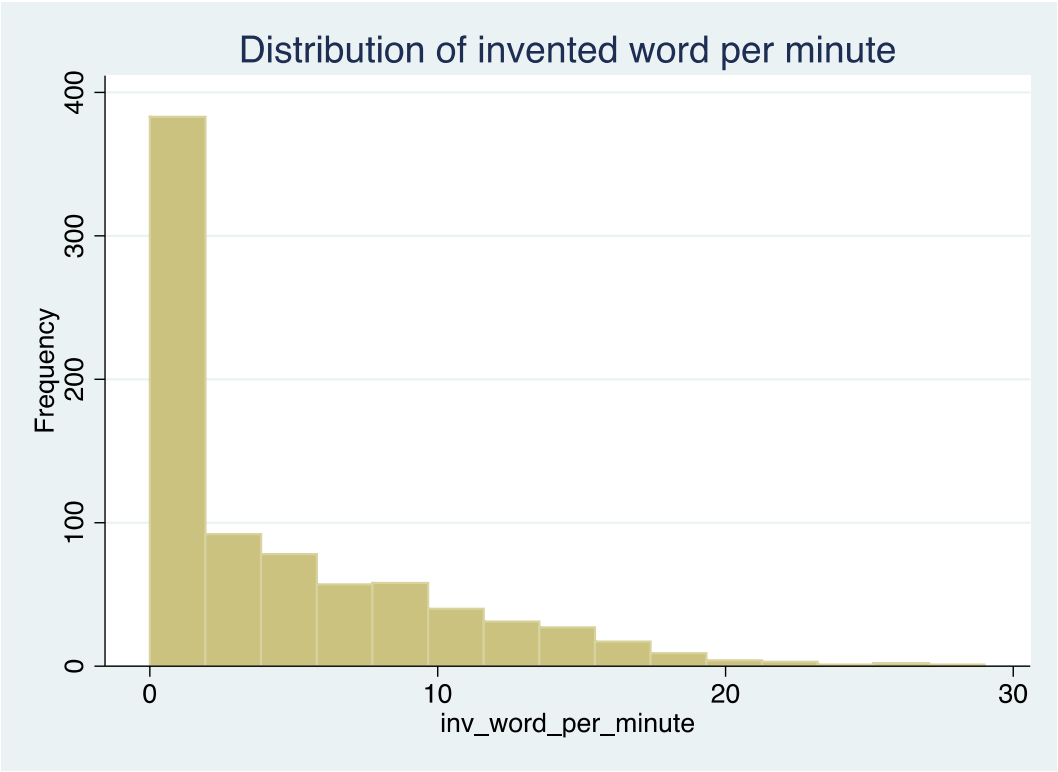
Mean and standard deviation for timed subtask

	Number of items attempted	Number of correct response per minute
Letter fluency	30.9 (17.6)	22.7 (20.0)
Syllable fluency	21.5 (11.6)	14.1 (13.4)
Invented word fluency	9.6 (6.0)	4.2 (5.4)
Oral reading fluency	14.2 (10.1)	7.9 (10.3)

Proportion of zero score for timed subtask

	Zero-score
Letter fluency	23.5%
Syllable fluency	22.8%
Invented word fluency	41.8%
Oral reading fluency	26.7%





Untimed subtasks

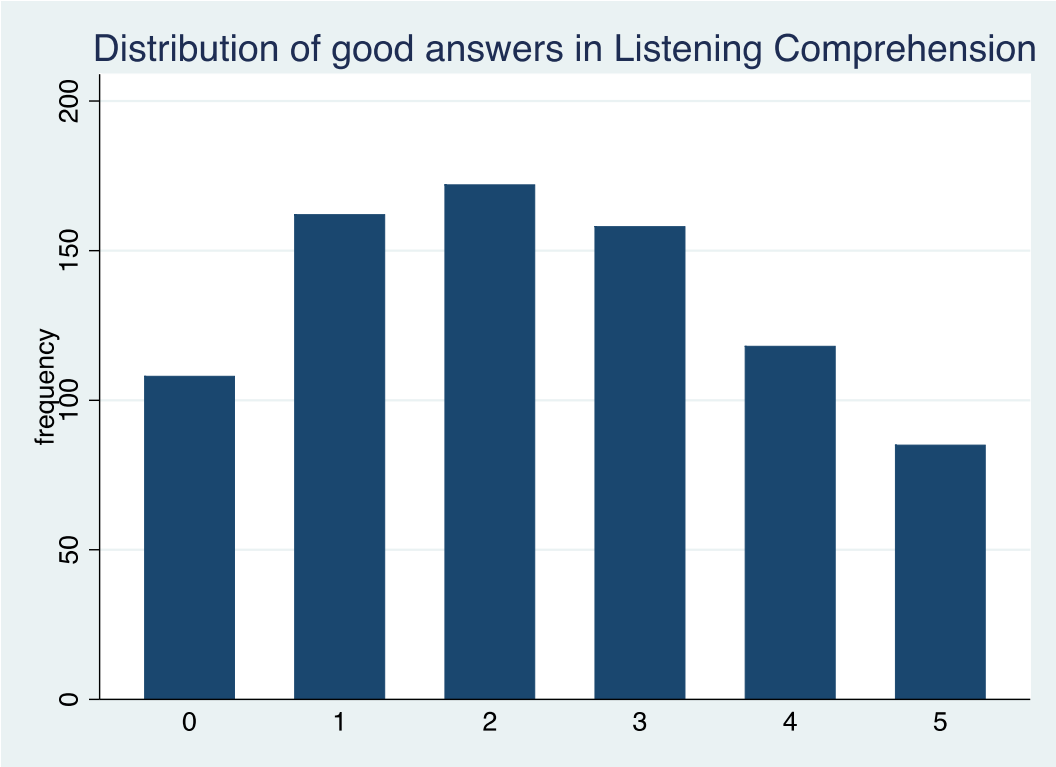
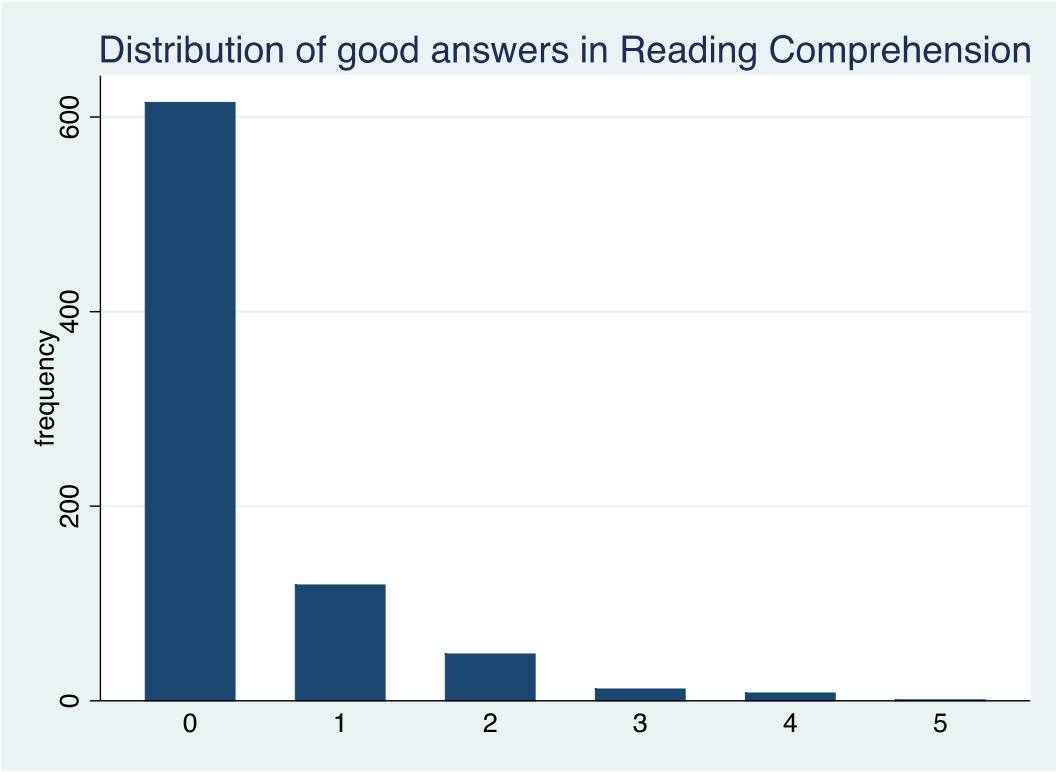
Proportion of attempted answers for reading comprehension*

Number of items	Reading comprehension	
	n	%
0	47	5.9%
1	542	67.5%
2	131	16.3%
3	47	5.9%
4	19	2.4%
5	17	2.1%
Total	803	100%

*All listening comprehension were attempted by the students

Proportion of numbers of good answers for reading and listening comprehension

Number of items	Reading comprehension		Listening comprehension	
	n	%	n	%
0	615	76.6%	108	13.5%
1	119	14.8%	162	20.2%
2	48	5.9%	172	21.4%
3	12	1.5%	158	19.7%
4	8	1.0%	118	14.7%
5	1	0.1%	85	10.6%
Total	803	100%	863	100%



Disaggregation by group and gender

Timed subtasks

Mean and Standard Deviation for Letter per minute by group and gender

	n	mean	sd
Experimental			
Female	255	22.4	18.8
Male	126	24.7	19.0
Total	381	23.2	18.9
Control			
Female	166	29.1	21.2
Male	256	17.7	19.7
Total	422	22.2	21.0

ANOVA results for Letter per minute by group and gender

Effect	F	p	Effect size
Group	0.01	0.9099	.000
Gender	9.86	0.0017	.012
Group X Gender	22.50	<0.0000	.027

Mean and Standard Deviation for Syllable per minute by group and gender

	n	mean	sd
Experimental			
Female	255	16.3	12.1
Male	126	14.8	13.4
Total	381	15.8	12.5
Control			
Female	166	17.4	15.3
Male	256	9.5	12.2
Total	422	12.6	14.1

ANOVA results for Syllable per minute by group and gender

Effect	F	p	Effect size
Group	4.71	0.0303	.006
Gender	23.29	<0.0000	.028
Group X Gender	10.78	0.0011	.013

Mean and Standard Deviation for Invented Words per minute by group and gender

	n	mean	sd
Experimental			
Female	255	4.7	4.8
Male	126	3.8	5.2
Total	381	4.4	4.9
Control			
Female	166	5.7	6.5
Male	256	3.1	5.0
Total	422	4.1	5.8

ANOVA results for Invented Words per minute by group and gender

Effect	F	p	Effect size
Group	0.15	0.7025	.000
Gender	20.54	<0.0000	.025
Group X Gender	5.27	0.0219	.007

Mean and Standard Deviation for Oral Reading per minute by group and gender

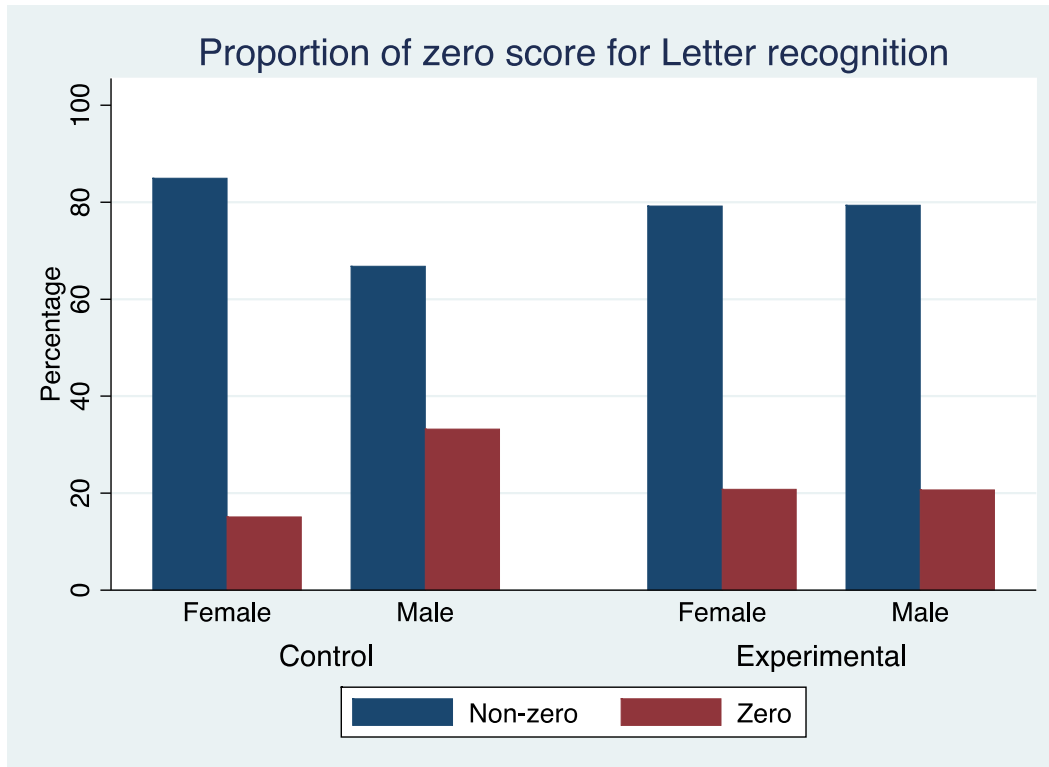
	n	mean	sd
Experimental			
Female	255	8.8	9.3
Male	126	7.7	10.1
Total	381	8.4	9.6
Control			
Female	166	10.4	11.7
Male	256	5.6	9.9
Total	422	7.5	10.9

ANOVA results for Oral Reading per minute by group and gender

Effect	F	p	Effect size
Group	0.10	0.7562	.000
Gender	15.35	0.0001	.019
Group X Gender	6.28	0.0124	.008

Proportion of zero-score for Letter recognition by Group and Gender

	n	%
Experimental		
Female	53	20.8%
Male	26	20.6%
Total	79	20.7%
Control		
Female	25	15.1%
Male	85	33.2%
Total	110	26.1%



Chi-square test for differences in proportion for zero score in Letter recognition

Overall	Chi2	p
Group	3.16	.075
Gender	12.34	<0.000

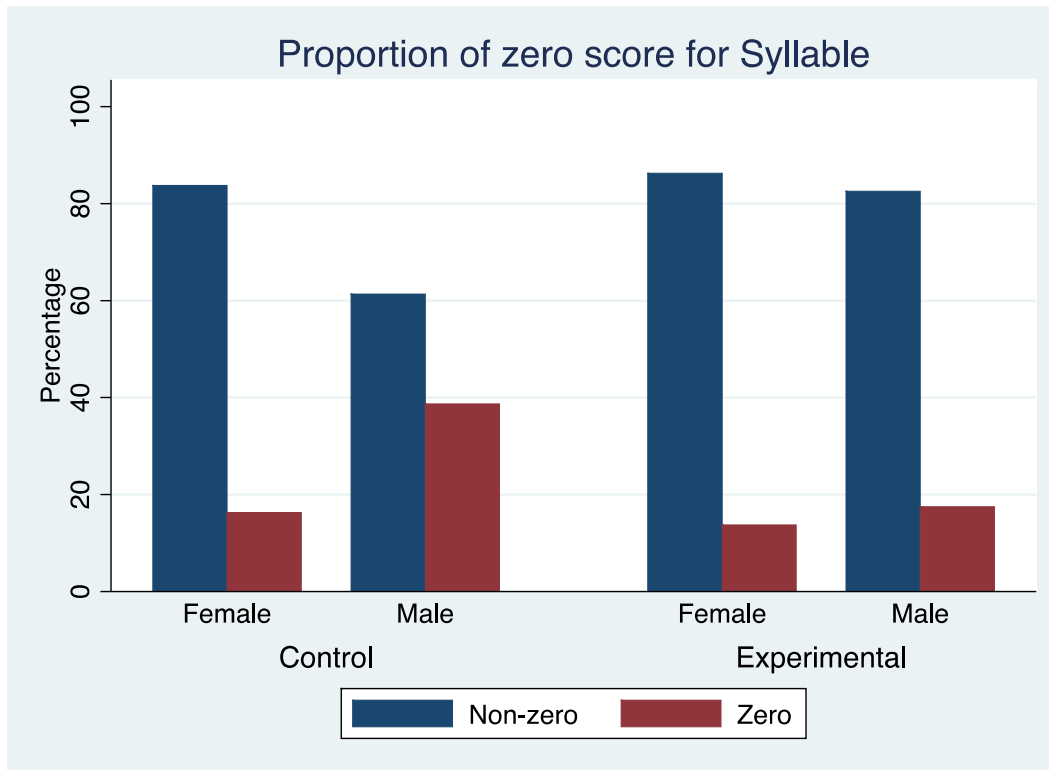
Group by gender		
Female	2.18	0.140
Male	6.47	0.011

Gender by group

Experimental	0.00	0.973
Control	17.19	<0.000

Proportion of zero-score for Syllable recognition by Group and Gender

	n	%
Experimental		
Female	35	13.7%
Male	22	17.5%
Total	57	14.9%
Control		
Female	27	16.3%
Male	99	38.7%
Total	126	29.9%

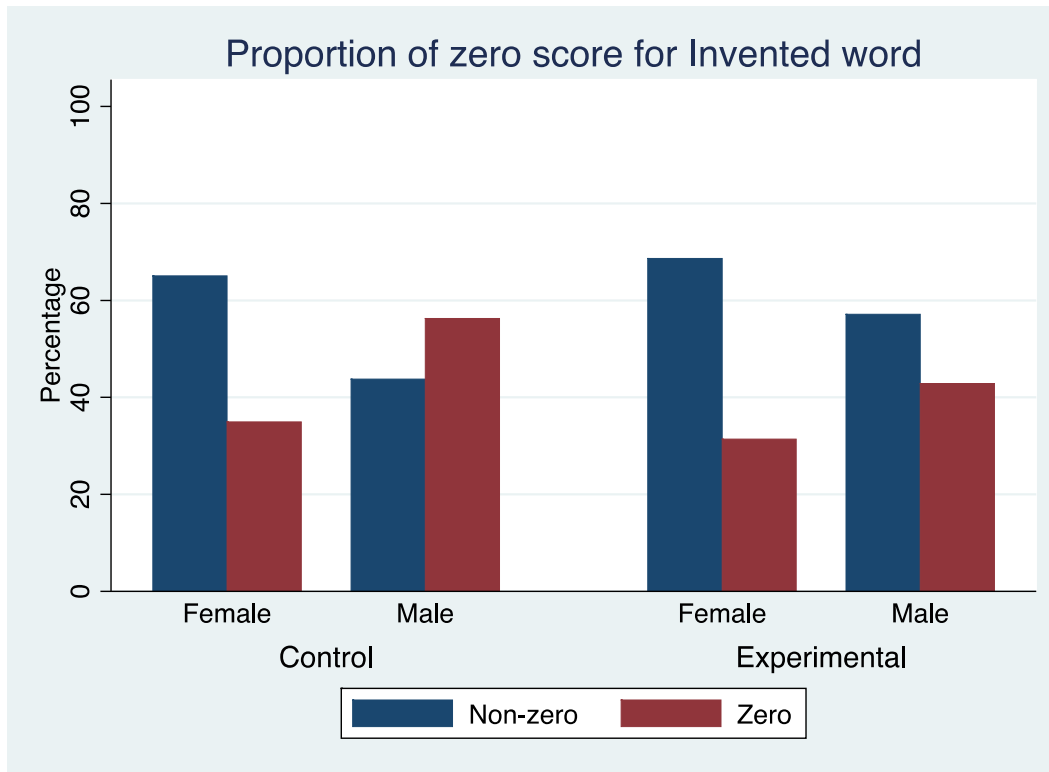


Chi-square test for differences in proportion for zero score in Syllable

Overall	Chi2	p
Group	25.25	<0.000
Gender	32.70	<0.000
Group by gender		
Female	0.52	0.472
Male	17.55	<0.000
Gender by group		
Experimental	0.92	0.336
Control	24.14	<0.000

Proportion of zero-score for Invented Words reading by Group and Gender

	n	%
Experimental		
Female	80	31.4%
Male	54	42.9%
Total	134	35.2%
Control		
Female	58	34.9%
Male	144	56.3%
Total	202	47.9%



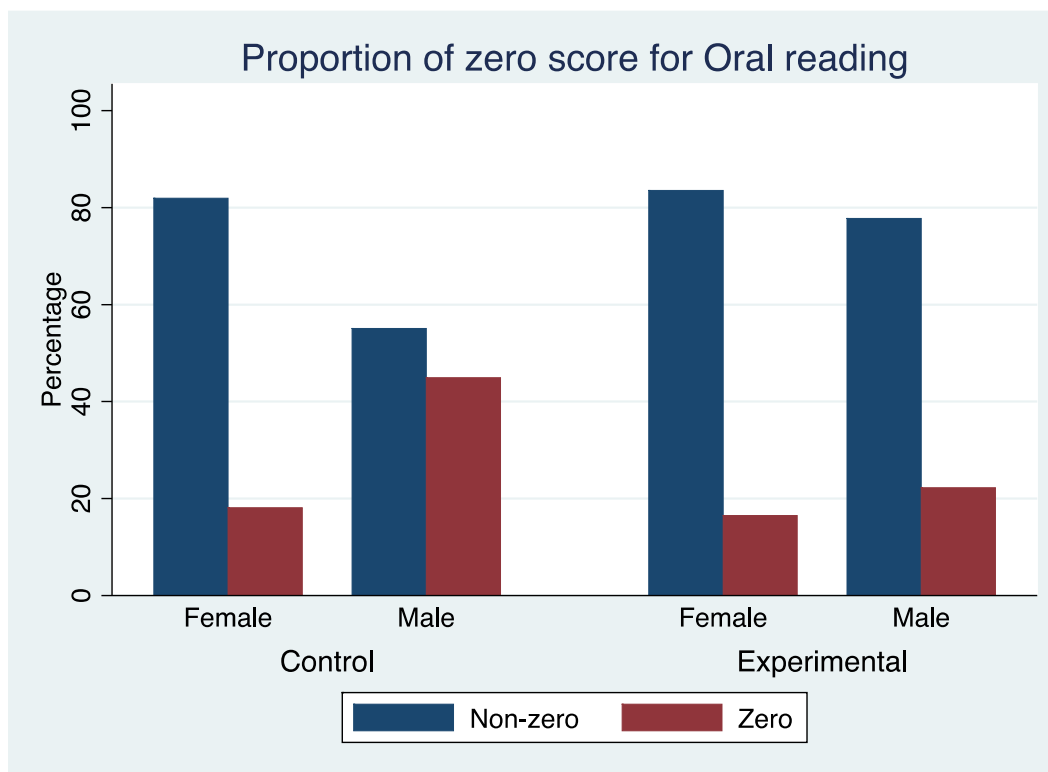
Chi-square test for differences in proportion for zero score in Invented words

Overall	Chi2	p
Group	13.26	<0.000
Gender	29.88	<0.000
Group by gender		
Female	0.58	0.446
Male	6.07	0.014
Gender by group		
Experimental	4.88	0.027
Control	18.33	<0.000

Proportion of zero-score for Oral Reading by Group and Gender

	n	%
Experimental		
Female	42	16.5%
Male	28	22.2%
Total	70	18.4%
Control		

Female	30	18.1%
Male	115	44.9%
Total	145	34.4%



Chi-square test for differences in proportion for zero score in Oral reading

Overall	Chi2	p
Group	26.10	<0.000
Gender	42.23	<0.000
Group by gender		
Female	0.18	0.670
Male	18.58	<0.000
Gender by group		
Experimental	1.86	0.173
Control	32.19	<0.000

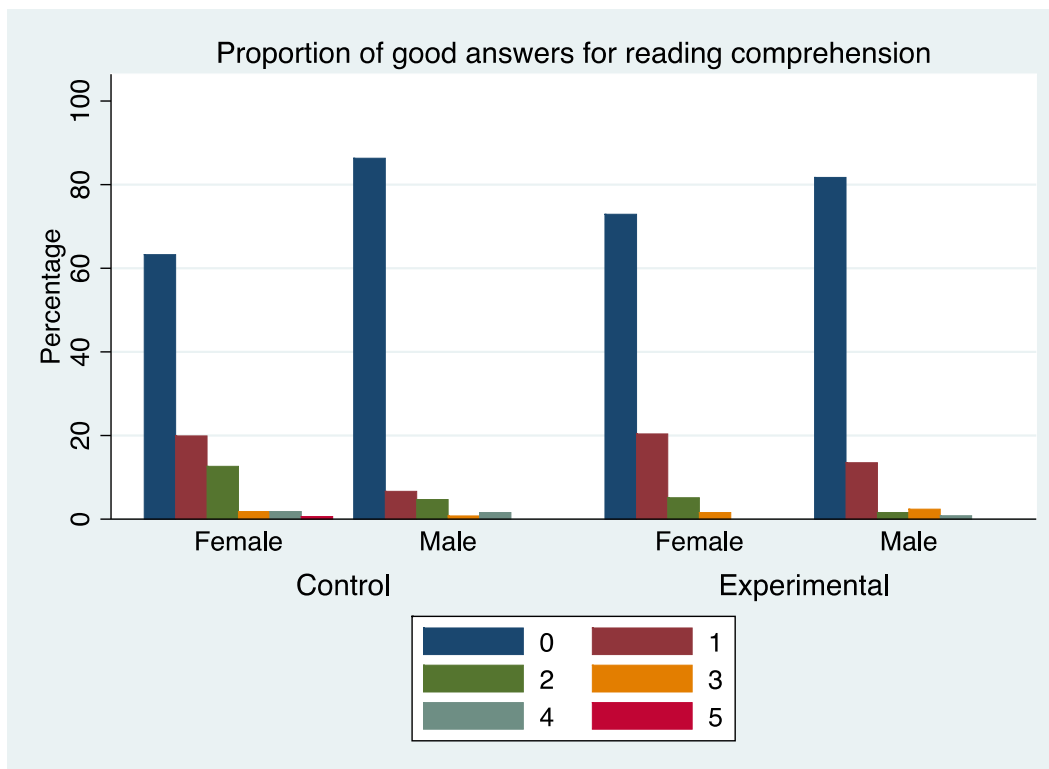
Untimed subtasks

Proportion of numbers of good answers for reading comprehension by group and gender

Experimental

Number of items	Female		Male		Total	
	n	%	n	%	n	%
0	186	72.9%	103	81.8%	289	75.9%
1	52	20.4%	17	13.5%	69	18.1%
2	13	5.1%	2	1.6%	15	3.9%
3	4	1.6%	3	2.4%	7	1.8%
4	0	0.0%	1	0.8%	1	0.3%
5	0	0.0%	0	0.0%	0	0.0%
Total	255	100%	126	100%	381	100%

Number of items	Control Female		Male		Total	
	n	%	n	%	n	%
0	105	63.3%	221	86.3%	326	77.3%
1	33	19.9%	17	6.6%	50	11.9%
2	21	12.7%	12	4.7%	33	7.8%
3	3	1.8%	2	0.8%	5	1.2%
4	3	1.8%	4	1.6%	7	1.7%
5	1	0.6%	0	0.0%	1	0.2%
Total	166	100%	256	100%	422	100%



Chi-square test for differences in proportion for reading comprehension

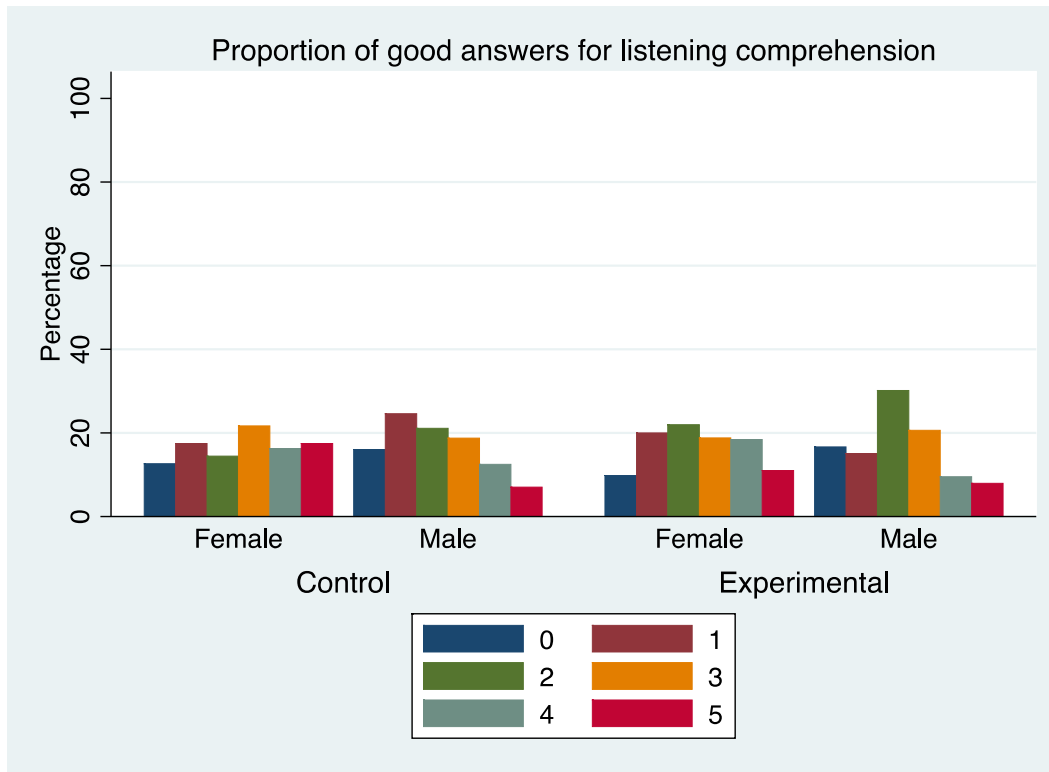
Overall Chi2 p

Group	15.79	0.007
Gender	31.98	<0.000
Group by gender		
Female	14.66	0.012
Male	8.91	0.063
Gender by group		
Experimental	8.05	0.090
Control	32.48	<0.000

Proportion of numbers of good answers for listening comprehension by group and gender

Number of items	Experimental		Male		Total	
	n	%	n	%	n	%
0	25	9.8%	21	16.7%	46	12.1%
1	51	20.0%	19	15.1%	70	18.4%
2	56	21.9%	38	30.2%	94	24.7%
3	48	18.8%	26	20.6%	74	19.4%
4	47	18.4%	12	9.5%	59	15.5%
5	28	11.0%	10	7.9%	38	10.0%
Total	255	100%	126	100%	381	100%

Number of items	Control		Male		Total	
	n	%	n	%	n	%
0	21	12.7%	41	16.0%	62	14.7%
1	29	17.5%	63	24.6%	92	21.8%
2	24	14.5%	54	21.1%	78	18.5%
3	36	21.7%	48	18.8%	84	19.9%
4	27	16.3%	32	12.5%	59	13.9%
5	29	17.5%	18	7.0%	47	11.1%
Total	166	100%	256	100%	422	100%



Chi-square test for differences in proportion for reading comprehension

Overall	Chi2	p
Group	6.36	0.273
Gender	19.54	0.002
Group by gender		
Female	7.87	0.163
Male	7.37	0.194
Gender by group		
Experimental	11.95	0.036
Control	16.84	0.005

XV. Item Level Reliability

Letter fluency

Item Statistics

	Mean	Std. Deviation	Corrected Item- Total Correlation
rletter1	,70	,456	,633
rletter2	,67	,469	,668
rletter3	,69	,464	,683
rletter4	,62	,486	,730
rletter5	,50	,500	,623
rletter6	,59	,492	,666
rletter7	,49	,500	,610
rletter8	,64	,480	,686
rletter9	,50	,500	,735
rletter10	,60	,491	,725
rletter11	,64	,479	,684
rletter12	,51	,500	,636
rletter13	,63	,482	,718
rletter14	,57	,495	,760
rletter15	,54	,499	,758
rletter16	,61	,487	,732
rletter17	,59	,493	,725
rletter18	,54	,498	,764
rletter19	,53	,499	,709
rletter20	,28	,451	,454
rletter21	,48	,500	,751
rletter22	,43	,495	,664
rletter23	,49	,500	,781
rletter24	,51	,500	,774
rletter25	,47	,499	,759
rletter26	,47	,499	,787
rletter27	,40	,489	,773
rletter28	,46	,499	,808
rletter29	,42	,494	,802
rletter30	,42	,494	,804
rletter31	,41	,492	,809
rletter32	,34	,476	,773
rletter33	,37	,484	,811
rletter34	,33	,472	,794
rletter35	,34	,475	,809
rletter36	,31	,464	,791
rletter37	,32	,465	,796
rletter38	,32	,466	,782
rletter39	,25	,436	,708
rletter40	,25	,433	,732
rletter41	,23	,424	,759

rletter42	,24	,428	,764
rletter43	,23	,423	,754
rletter44	,21	,410	,736
rletter45	,22	,412	,723
rletter46	,20	,402	,707
rletter47	,19	,390	,701
rletter48	,17	,380	,697
rletter49	,16	,370	,680
rletter50	,13	,333	,586
rletter51	,13	,340	,642
rletter52	,12	,325	,622
rletter53	,12	,323	,627
rletter54	,11	,317	,618
rletter55	,10	,301	,591
rletter56	,09	,286	,565
rletter57	,08	,279	,554
rletter58	,08	,269	,538
rletter59	,07	,263	,525
rletter60	,07	,255	,521
rletter61	,05	,225	,482
rletter62	,05	,212	,463
rletter63	,04	,196	,426
rletter64	,04	,190	,414
rletter65	,03	,174	,389
rletter66	,02	,148	,355
rletter67	,02	,152	,362
rletter68	,02	,140	,351
rletter69	,02	,144	,347
rletter70	,02	,140	,351
rletter71	,01	,116	,308
rletter72	,01	,105	,285
rletter73	,01	,099	,275
rletter74	,01	,093	,264
rletter75	,01	,086	,249
rletter76	,01	,086	,249
rletter77	,01	,086	,249
rletter78	,01	,086	,249
rletter79	,01	,086	,249
rletter80	,01	,086	,249
rletter81	,01	,079	,232
rletter82	,01	,079	,232
rletter83	,01	,079	,232
rletter84	,00	,061	,182

rletter85	,00	,050	,152
rletter86	,00	,035	,110
rletter87	,00	,035	,110
rletter88	,00	,035	,110
rletter89	,00	,035	,110
rletter90	,00	,035	,110
rletter91	0,00	0,000	0,000
rletter92	0,00	0,000	0,000
rletter93	0,00	0,000	0,000
rletter94	0,00	0,000	0,000
rletter95	0,00	0,000	0,000
rletter96	0,00	0,000	0,000
rletter97	0,00	0,000	0,000
rletter98	0,00	0,000	0,000
rletter99	0,00	0,000	0,000
rletter100	0,00	0,000	0,000

Reliability Statistics

Cronbach's Alpha	N of Items
,979	100

Syllable fluency

Item Statistics

	Mean	Std. Deviation	Corrected Item-Total Correlation
rSyllable1	,54	,498	,183
rSyllable2	,58	,494	,223
rSyllable3	,39	,489	,187
rSyllable4	,50	,500	,193
rSyllable5	,52	,500	,205
rSyllable6	,60	,489	,168
rSyllable7	,54	,499	,208
rSyllable8	,56	,497	,197
rSyllable9	,62	,486	,177
rSyllable10	,62	,486	,256
rSyllable11	,45	3,514	,942
rSyllable12	,59	3,513	,948
rSyllable13	,64	3,511	,947
rSyllable14	,60	3,513	,952
rSyllable15	,58	3,513	,953

rSyllable16	,56	3,513	,949
rSyllable17	,58	3,513	,952
rSyllable18	,54	3,514	,953
rSyllable19	,55	3,514	,955
rSyllable20	,47	3,514	,953
rSyllable21	,34	,474	,245
rSyllable22	,32	,467	,238
rSyllable23	,33	,471	,254
rSyllable24	,12	,329	,151
rSyllable25	,31	,463	,253
rSyllable26	,27	,446	,243
rSyllable27	,25	,431	,253
rSyllable28	,25	,436	,254
rSyllable29	,21	,411	,251
rSyllable30	,19	,395	,238
rSyllable31	,15	,360	,227
rSyllable32	,17	,372	,241
rSyllable33	,14	,349	,228
rSyllable34	,14	,348	,234
rSyllable35	,12	,322	,227
rSyllable36	,11	,313	,219
rSyllable37	,11	,314	,226
rSyllable38	,09	,286	,208
rSyllable39	,08	,277	,204
rSyllable40	,08	,279	,211
rSyllable41	,07	,261	,205
rSyllable42	,06	,244	,194
rSyllable43	,03	,163	,129
rSyllable44	,05	,218	,178
rSyllable45	,04	,187	,153
rSyllable46	,04	,187	,156
rSyllable47	,03	,177	,150
rSyllable48	,03	,184	,159
rSyllable49	,03	,163	,145
rSyllable50	,03	,160	,142
rSyllable51	,02	,126	,115
rSyllable52	,02	,131	,123
rSyllable53	,02	,131	,123
rSyllable54	,01	,111	,112
rSyllable55	,01	,116	,112
rSyllable56	,01	,093	,083
rSyllable57	,01	,093	,092
rSyllable58	,01	,086	,085

rSyllable59	,00	,070	,065
rSyllable60	,00	,061	,061
rSyllable61	,00	,035	,047
rSyllable62	,00	,035	,047
rSyllable63	,00	,035	,047
rSyllable64	,00	,035	,047
rSyllable65	,00	,035	,047
rSyllable66	,00	,035	,047
rSyllable67	,00	,035	,047
rSyllable68	,00	,035	,047
rSyllable69	0,00	0,000	0,000
rSyllable70	,00	,035	,047
rSyllable71	,00	,035	,047
rSyllable72	0,00	0,000	0,000
rSyllable73	0,00	0,000	0,000
rSyllable74	0,00	0,000	0,000
rSyllable75	0,00	0,000	0,000
rSyllable76	0,00	0,000	0,000
rSyllable77	0,00	0,000	0,000
rSyllable78	0,00	0,000	0,000
rSyllable79	0,00	0,000	0,000
rSyllable80	0,00	0,000	0,000
rSyllable81	0,00	0,000	0,000
rSyllable82	0,00	0,000	0,000
rSyllable83	0,00	0,000	0,000
rSyllable84	0,00	0,000	0,000
rSyllable85	0,00	0,000	0,000
rSyllable86	0,00	0,000	0,000
rSyllable87	0,00	0,000	0,000
rSyllable88	0,00	0,000	0,000
rSyllable89	0,00	0,000	0,000
rSyllable90	0,00	0,000	0,000
rSyllable91	0,00	0,000	0,000
rSyllable92	0,00	0,000	0,000
rSyllable93	0,00	0,000	0,000
rSyllable94	0,00	0,000	0,000
rSyllable95	0,00	0,000	0,000
rSyllable96	0,00	0,000	0,000
rSyllable97	0,00	0,000	0,000
rSyllable98	0,00	0,000	0,000
rSyllable99	0,00	0,000	0,000
rSyllable100	0,00	0,000	0,000

Reliability Statistics

Cronbach's Alpha	N of Items
,915	100

Invented word fluency

Item Statistics

	Mean	Std. Deviation	Corrected Item-Total Correlation
rinv_word1	,20	,400	,619
rinv_word2	,33	,470	,642
rinv_word3	,41	,492	,632
rinv_word4	,41	,492	,514
rinv_word5	,36	,480	,622
rinv_word6	,19	,394	,520
rinv_word7	,29	,453	,699
rinv_word8	,26	,436	,687
rinv_word9	,32	,465	,656
rinv_word10	,17	,380	,592
rinv_word11	,29	,452	,737
rinv_word12	,18	,384	,678
rinv_word13	,18	,381	,656
rinv_word14	,07	,255	,536
rinv_word15	,11	,317	,626
rinv_word16	,10	,300	,627
rinv_word17	,07	,263	,614
rinv_word18	,05	,223	,494
rinv_word19	,06	,230	,512
rinv_word20	,04	,199	,483
rinv_word21	,03	,184	,436
rinv_word22	,02	,156	,374
rinv_word23	,01	,116	,331
rinv_word24	,03	,174	,397
rinv_word25	,01	,121	,311
rinv_word26	,01	,093	,215
rinv_word27	,01	,111	,272
rinv_word28	,00	,070	,240
rinv_word29	,00	,035	,156
rinv_word30	,00	,035	,156
rinv_word31	,00	,061	,176

rinv_word32	0,00	0,000	0,000
rinv_word33	,00	,035	,143
rinv_word34	,00	,035	,143
rinv_word35	,00	,035	,143
rinv_word36	,00	,035	,143
rinv_word37	,00	,035	,143
rinv_word38	,00	,035	,143
rinv_word39	,00	,035	,143
rinv_word40	,00	,035	,143
rinv_word41	0,00	0,000	0,000
rinv_word42	,00	,035	,143
rinv_word43	,00	,035	,143
rinv_word44	,00	,035	,143
rinv_word45	,00	,035	,143
rinv_word46	0,00	0,000	0,000
rinv_word47	0,00	0,000	0,000
rinv_word48	0,00	0,000	0,000
rinv_word49	0,00	0,000	0,000
rinv_word50	0,00	0,000	0,000

Reliability Statistics

Cronbach's Alpha	N of Items
,911	50

Reading comprehension

Item Statistics

	Mean	Std. Deviation	Corrected Item-Total Correlation
rread_comp01	,42	,494	,019
rread_comp02	,04	,201	,126
rread_comp03	,01	,111	,204
rread_comp04	,01	,086	,205
rread_comp05	,00	,070	,217

Reliability Statistics

Cronbach's Alpha	N of Items
,169	5

Listening comprehension

Item Statistics

	Mean	Std. Deviation	Corrected Item- Total Correlation
rlist_comp01	,23	,422	,514
rlist_comp02	,22	,414	,576
rlist_comp03	,22	,412	,523
rlist_comp04	,30	,459	,530
rlist_comp05	,10	,295	,465

Reliability Statistics

Cronbach's Alpha	N of Items
,751	5