



Girls' Improved Learning Outcomes (GILO)
Quality Education for Girls

تحسين الأداء التعليمي للبنات
جودة التعليم للبنات

Improved Reading Performance in Grade 2

GILO-Supported Schools vs. Control Schools

Results of the
Early Grade Reading Assessments (EGRAs):
2009 Baseline
and
2011 Post-Intervention Assessments



USAID | **EGYPT**
FROM THE AMERICAN PEOPLE



جمهورية مصر العربية

وزارة التربية والتعليم
وإدارة التدريب والتعليم

Table of Contents

Introduction	1
Comparison of 2009 and 2011 Results for Primary Grade 2	2
Conclusion	6
Annex 1: GILO Implementation of EGRA 2009 and EGRA 2011	8

Introduction:

This *Final Report* of the 2009 and 2011 rounds of the Early Grade Reading Assessment (EGRA) conducted in Egypt compares the results of the 2009 baseline and the 2011 post-intervention assessments of Arabic reading skills for primary grade 2 students.

The 2009 EGRA baseline included students in primary grades 2, 3 and 4 who were tested using the same form. The ensuing intervention to enhance classroom instruction in Arabic phonics, developed and conducted by the GILO project and members of the Ministry of Education (MOE) Working Group for Early Grade Literacy, strengthened phonics instruction by classroom teachers in primary grades 1 and 2 and instructional support from their Arabic language supervisors. Teachers of grades 1 and 2 learned and applied the same *Early Grade Reading Package* of training, manuals, classroom instructional routines and resource materials, coaching and classroom observation, over six school months between September 2010 and April 2011. The inclusion of primary 2 teachers provided grade-overlap between the 2009 and 2011 EGRAs so that the outcomes of this intervention could be measured.

Both the 2009 baseline and 2011 post-intervention EGRAs were conducted in the same GILO-supported (*intervention*) and control schools. The implementation of both assessments and the selection of sample schools and students are described in detail in Annex 1. Both EGRAs used the same form. The only meaningful differences in the implementation of the two assessments were: i) the 2009 baseline was conducted in January-February while the 2011 EGRA was conducted in April-May, following the Egyptian revolution that began on January 25 and a full school year of intervention support; and ii) the 2011 EGRA included a narrowed selection of sub-tests.

The results of the 2009 EGRA baseline were formally presented to USAID Egypt and the Ministry of Education (MOE) in October 2009. The analyses presented then emphasized the positive relationship between students' ability to read syllables (letter sounds) and their ability to decode unknown words and read passages. The results recommended an intervention to strengthen phonics instruction in order to enhance overall reading performance. USAID and MOE subsequently approved development of the *Early Grade Reading Package* and its implementation in GILO-supported schools during the 2010/11 school year.

This *Final Report* does not repeat the October 2009 analyses. The analyses and results presented here: i) are specific to primary grade 2 *only*, and ii) compare and contrast the 2009 and 2011 results of the GILO-supported (*intervention*) and control schools on just three key sub-tests. This report does not aim to be comprehensive. Rather its chief purpose is to provide a succinct presentation of key results to MOE *muderiyyas*, *idaras* and schools for their knowledge and consideration as they embark on a nationwide effort in 2011/12 to improve reading in all Grade 1 classrooms of Egyptian public schools.



Comparison of 2009 and 2011 Results for Primary Grade 2:

The 2011 results showed a very marked improvement in the reading performance of sampled students in intervention schools from the 2009 EGRA baseline, as compared with the control schools. Table 1 below presents the respective mean scores¹ of sampled Primary 2 students in the intervention and control schools for each of three (3) priority sub-tests:

Table 1: Mean Scores of Intervention and Control School Students on EGRA Sub-Tests

EGRA Measures	Mean Scores, INTERVENTION Schools		% Change	Mean Scores, CONTROL Schools		% Change
	2009	2011		2009	2011	
Syllable Reading	9.76	28.47	+ 192%	8.55	10.10	+ 18%
Word Reading	7.35	15.50	+ 111%	5.56	7.45	+ 34%
Oral Reading Fluency	11.09	21.14	+ 91%	8.92	10.93	+ 23%

The performance differences are clear. After six months of improved phonics instruction, Primary 2 students in GILO-supported schools could read many more syllables and words correctly in one minute than similar grade students in control schools that did not receive this support:

- **Syllable reading** improved 192% in the *intervention* schools: from an average score of 9.76 syllables read correctly in one minute in 2009 to 28.47 correct syllables in 2011. In contrast, the improvement in CONTROL schools was just 18% – from an average score of 8.55 syllables read correctly in one minute in 2009 to 10.10 correct syllables in 2011. The result: Primary 2 students in *intervention* schools were, on average, correctly reading nearly 3 times as many syllables (28.47) per minute in 2011 – after 6 months of improved reading instruction – as students of the same grade in CONTROL schools (10.10).
- **Word reading** improved 111% in the *intervention* schools: from an average score of 7.35 words read correctly in one minute in 2009 to 15.50 correct words in 2011. The improvement in CONTROL schools was 34% – from an average score of 5.56 words read correctly in one minute in 2009 to 7.45 correct words in 2011. Primary 2 students in *intervention* schools were, on average, correctly reading more than twice as many words (15.50) per minute in 2011 as students of the same grade in CONTROL schools (7.45).
- **Oral reading fluency** (passage reading) improved 91% in the *intervention* schools: from an average of 11.09 passage words read correctly in one minute in 2009 to 21.14 correct passage words in 2011. The improvement in CONTROL schools was only 23% – from an average score of 8.92 passage words read correctly in one minute in 2009 to 10.93 correct passage words in 2011. Primary 2 students in *intervention* schools were, on average, correctly reading nearly

¹ Mean scores shown are the total number of syllables or words – depending on the sub-test – correctly read in 60 seconds. Consistent with international EGRA practice, students who cannot correctly read *any* one of the first ten letters or words in a specific EGRA sub-test are dropped from that sub-test sample and the calculation of sub-test mean scores.

twice as many passage words (21.14) in one minute in 2011 than students of the same grade in CONTROL schools (10.93).

These are remarkable improvements in *average* reading proficiency in the first year of the GILO intervention. Average scores do not, however, present a complete picture of reading improvement in the classrooms of intervention schools. Table 2 and Table 3 below compare the change in letter sound and word reading – from 2009 to 2011 – for each of four (4) levels of reading proficiency in intervention and control classes.

Improved Syllable Reading: The GILO *Early Grade Reading Package* significantly improved syllable reading across the majority of students in intervention classrooms. This finding is especially important: improved reading was not an achievement of a minority of students.

As Table 2 shows clearly, in 2009 only 12.6% of students in *intervention* schools – and 12.0% in CONTROL schools – could read 27 or more syllables correctly in one minute.² After six months of GILO support, more than half (51.2%) of Primary 2 students in *intervention* schools in 2011 could read at this highest level of reading proficiency. But only 13.5% of Primary 2 students in CONTROL schools in 2011 could read at this level.

Table 2

Comparing EGRA Results: 2009 and 2011						
Improved Syllable Reading						
Intervention versus Control Schools: Primary Grade 2						
	Intervention Schools			Control Schools		
	2009	2011	Change	2009	2011	Change
27+ Correct Syllables	12.6%	51.2%	+ 38.6%	12.0%	13.5%	+ 1.5%
16-26 Correct Syllables	14.0%	15.9%	+ 1.9%	8.0%	10.7%	+ 2.7%
1-15 Correct Syllables	25.0%	21.8%	- 3.2%	28.4%	30.7%	+ 2.3%
<u>NO</u> Correct Syllables	48.4%	11.1%	- 37.3%	51.6%	45.0%	- 6.6%
Mean Number of Correct Letters	9.76	28.47		8.55	10.10	

Similarly marked change also occurred at the lowest level of reading proficiency, i.e. those students who read none of the first 10 syllables correctly. In 2009, 48.4% of Primary 2 students in *intervention* schools – and a similar percentage (51.6%) of students in CONTROL schools – read none of the first 10 syllables correctly. After the GILO intervention (2011), however, just 11.1% of students in intervention

² Reading 27+ letter sounds correctly in one minute is the desired level of reading proficiency for Primary 2 students.

schools read none of the first 10 syllables correctly. By comparison, 45.0% of students in CONTROL schools still could not correctly read any syllables in 2011.

This is an key result. The percentage of students who are unable to read dropped sharply – from nearly half (48.4%) of all Primary 2 students in *intervention* schools (2009) to only 1 in 9 students in (11.1%) in 2011. In sharp contrast, nearly half (45.0%) of Primary 2 students in CONTROL schools could still not read syllables at the end of the 2011 school year.

Improved Word Reading: Improvement in *word* reading was also widespread in GILO-supported schools. The percentage of students who could not correctly read a single word dropped sharply in these *intervention* schools between 2009 and 2011. In marked contrast, the word reading performance of Primary 2 students in CONTROL schools was nearly unchanged in this period.

Table 3

Comparing EGRA Results: 2009 and 2011						
Improved Word Reading						
<i>Intervention versus Control Schools: Primary Grade 2</i>						
	Intervention Schools			Control Schools		
	2009	2011	Change	2009	2011	Change
25+ Correct Words	8.8%	24.2%	+ 15.4%	7.1%	10.6%	+ 3.5%
16-24 Correct Words	9.7%	12.2%	+ 2.5%	5.8%	6.6%	+ 0.8%
1-15 Correct Words	32.9%	42.2%	+ 9.3%	31.8%	31.0%	- 0.8%
<u>NO</u> Correct Words	48.6%	21.4%	- 27.2%	55.3%	51.8%	- 2.5%
Mean Number of Correct Words	7.35	15.50		5.56	7.45	

Table 3 compares the change in word reading performance of Primary 2 students in *intervention* and CONTROL schools between 2009 and 2011. The performance improvement in GILO-supported schools is extensive; but virtually no change in CONTROL schools.

In the *intervention* schools, the percentage of students reading at the desired level of proficiency (25+ correct words per minute), jumped from 8.8% of sampled students in 2009 to 24.2% of sampled students in 2011 – an increase of 15.4 percentage points. In these same schools, the percentage of students unable to read any correct words dropped more than half: from 48.6% of all sampled students in 2009 to 21.4% of students in 2011 – a drop of 27.2%.

In CONTROL schools, there was no meaningful change in word reading performance among sampled students.

Improved Oral Fluency (Passage Reading): Lastly, the results of the oral fluency test mirror the change in word reading performance described above: broad and significant improvement in oral fluency in *intervention* schools; no meaningful change in CONTROL schools.

Table 4

Comparing EGRA Results: 2009 and 2011						
Improved Oral Fluency (Passage Reading)						
Intervention versus Control Schools: Primary Grade 2						
	Intervention Schools			Control Schools		
	2009	2011	Change	2009	2011	Change
45+ Correct Words	4.7%	19.7%	+ 15.0%	4.1%	7.6%	+ 3.5%
21-44 Correct Words	15.1%	21.1%	+ 6.0%	11.4%	12.6%	+ 1.2%
1-20 Correct Words	35.8%	38.5%	+ 2.7%	33.8%	33.4%	- 0.4%
<u>NO</u> Correct Words	44.4%	20.7%	- 23.7%	50.8%	46.5%	- 4.3%
Mean Number of Correct Words	11.09	21.14		8.92	10.93	

As Table 4 shows, in the *intervention* schools the percentage of students reading the passage at the desired level of proficiency (45+ correct words per minute), rose sharply from just 4.7% of sampled students in 2009 to 19.7% of sampled students in 2011 – an increase of 15.0 percentage points. In these same schools, the percentage of students unable to read any correct words dropped more than half: from 44.4% of all sampled students in 2009 to 20.7% of students in 2011 – a drop of 23.7%.

In marked contrast, there was no meaningful change in oral fluency in CONTROL schools between 2009 and 2011. In control schools, the oral fluency of Primary 2 students in 2011 remained at the same level shown for *intervention* schools in 2009: almost half of all students were unable to read a single word correctly in the passage's first line.

Conclusion:

The GILO *Early Grade Reading Package* strengthened phonics instruction of the Arabic alphabet and significantly improved reading performance on all three key measures: syllable reading, word reading and oral fluency. Many more Primary 2 students in *intervention* schools are now reading with greater facility and correctness after their teachers were trained and supported to improve reading instruction. Anecdotal observation suggests that the impacts of improved reading extend far beyond reading: students are more satisfied with their school experience, less disruptive in class, and better learners in all subjects. Schools and teachers observe more contented and enriched students and the school environment is enhanced for all.

Annex 1

GILO Implementation of EGRA 2009 and EGRA 2011

Background:

The GILO project began development and testing of the first Arabic version of the Early Grade Reading Assessment in summer 2008 (see *EGRA Timeline in Egypt* on next page). At the same time, the project, launched in March 2008, selected its first cohort of primary and preparatory schools for project support, applying school selection criteria agreed with USAID Egypt. The project planned training and professional development for teachers and instructional supervisors, school administrators and school Boards of Trustees, educational supervisors and school-based trainers in varied topics of support for improved learning outcomes of girls and boys. Importantly, the project design did not provide for improved instruction in early grade reading – only the development and application of an Arabic version of the Early Grade Reading Assessment. Nor did the selection of project schools anticipate the later implementation of project support for improved reading in early grades.

Implementing EGRA 2009:

Using the Arabic assessment tool developed and tested in summer 2008 (see *Annex 2*), GILO conducted its first implementation of EGRA in January-February 2009. For this first EGRA, a stratified random sample of nearly 2900 students was selected from primary grades 2, 3 and 4 in each of 59 participating schools. These 59 schools included: i) all 29 primary schools in the first cohort of project-supported schools in eleven districts of al-Fayoum, al-Minia and Qena governorates in Upper Egypt – the *intervention schools*, and ii) 10 control schools selected in a single district outside the geographic scope of GILO intervention in each of these same three governorates – for a total of 30 *control schools*. The control schools were selected as a stratified random sample that mirrored, as closely as possible, the rural/urban distribution and school population characteristics of the intervention schools (project-supported schools) in the same governorate. Neither the intervention nor control schools were selected in a fully random manner. This first implementation of EGRA did not aim to provide generalizable findings for Egypt or the three Upper Egyptian governorates where the assessment was implemented. The aim was to measure baseline reading performance of students in 3 grades in a defined set of intervention and control schools, applying the same assessment tool in all three grades. From each school, nearly equal numbers of students were tested in each grade (total 48-55 students per school), with equal numbers of boys and girls selected from each grade.

For EGRA 2009, GILO trained forty enumerators plus 7-10 assessment team leaders selected from volunteers and staff of leading NGOs active in education in al-Minia governorate and recent university graduates. GILO staff did not conduct any assessments. Enumerators received four (4) days of practical training. All EGRA 2009 assessments in all schools and governorates were conducted by these same enumerators over three weeks. All 20 Minia schools (intervention and control schools) were tested in four days (5 schools tested each day), then all 19 Fayoum schools in four days and then all 20 Qena schools tested in four days with the same team of enumerators moving from one governorate to the next. The samples of students tested in each school were randomly selected by GILO's M&E Director from complete student lists, with similar numbers of students (half boys, half girls) selected from all classrooms in the same grade. School principals were provided the list of selected students in advance of the test. If selected students were absent on the day of EGRA's implementation in a specific

school, the next child of the same gender on the student list was substituted. Substitutions were few and close attention was given to assuring the random selection of students in each school for testing. Each assessment was conducted individually between an enumerator and one student, with each complete test requiring 20-25 minutes.

Implementing EGRA 2011:

Owing to the Egyptian revolution that began 25 January 2011, the second implementation of EGRA was conducted in April-May 2011, some three months later in the school year than the January-February implementation of EGRA 2009. EGRA 2011 followed six months of project support that included an average of 6-8 days total of teacher training to apply specific classroom instructional routines and use specific learning aids, close observation of classroom instruction and follow-up support for improved early grade instruction in Arabic reading from GILO staff and professionals of the Ministry of Education's Working Group for Early Grade Literacy in primary grades 1 and 2 from September 2010. Improved instruction focused chiefly on phonics instruction in the Arabic alphabet. The majority of early grade teachers had previously received GILO training in student-centered active learning, classroom management, teaching resources, and other topics of classroom professional development.

For EGRA 2011, GILO trained enumerators plus assessment team leaders for three (3) days. Half of the enumerators had previously implemented EGRA 2009 and were well experienced (60-80 EGRA tests previously implemented by each enumerator). The other half of enumerators were Ministry of Education (MOE) staff from each of the three participating governorates. These MOE enumerators joined the experienced EGRA enumerators in implementing assessments in the selected intervention and control schools of their governorates. Each school was assessed by a mix of experienced and MOE enumerators. The intervention and control schools for EGRA 2011 were the same schools tested for EGRA 2009. The sample of tested schools for EGRA 2011 was three less than in EGRA 2009: two control schools in al-Fayoum governorate were dropped and not tested in EGRA 2011 in order to complete the tests in two intervention schools with low attendance on the previous, scheduled test day. The selection of the two control schools dropped was random. A third school – an intervention school in Qena – was dropped after it was discovered on the day of the EGRA that the test form had been leaked to the students the previous day. This school was subsequently dropped from the EGRA 2009 sample. Leakage of the EGRA test form was not found in any other school. Close attention was given by GILO to assuring the integrity and comparability of the test results. The test form used for EGRA 2011 was the same form used for EGRA 2009. The form had been closely guarded in EGRA 2009 and no copies were made available to the schools between the two assessments.

Consistent with the scope of project support from September 2010 to April 2011 for improved reading instruction in primary grades 1 and 2, EGRA 2011 was conducted only with students in primary grade 2 to overlap with the EGRA 2009 baseline. (Primary 1 students were not tested in EGRA 2009.) The training and support given to primary grade 2 students in the intervention schools was the same given to primary grade 1 students. The final EGRA 2011 sample included 574 primary grade 2 students from the same intervention schools and 635 from the same control schools.³ Approximately equal numbers of boys and girls were tested in each school. Tested students were selected randomly by GILO M&E staff from complete student lists provided by each school. Substitutions for sample students absent on the day of the test were selected systematically and were few. Where substitutions could not be confirmed to have been made systematically, the substituted students were dropped from the sample.

³ In EGRA 2009, 444 primary 2 students were tested in intervention (GILO-supported) schools and 465 students in control schools.

In sum, the EGRA 2009 and EGRA 2011 results reported here used a randomized controlled design with non-overlapping cohorts of primary grade 2 students.