

All Children Learning

Early Grade Learning Brief: EGYPT



I Findings

The Egypt Early Grade Reading Program (EGRP) began with Grade 1 in the 2011–12 school year. Four assessment studies were conducted from 2009 to 2013 to investigate the state of early grade reading in the country. The first, in 2009, established an illustrative baseline for Grade 2, Grade 3, and Grade 4 students across three governorates (El-Fayoum, El-Minya, and Qena). The second, April to May 2011, when combined with the 2009 Grade 2 baseline, offered an analysis of the impact on children’s reading of a specific phonics-based intervention in EGRP Grade 2 classrooms. A third assessment, conducted with Grade 1 students in Beheira and Cairo governorates in October 2011, tested the usefulness and sensitivity of the Early Grade Reading Assessment (EGRA) for Grade 1 learners in mainly urban settings. The first national-level baseline was conducted for Grade 3 in March 2013.

Summary of EGRA studies and their samples

Type of Information	2009 Baseline	2011 Follow-up	2011 Grade 1	2013 Baseline
Period of assessment	January – February 2009	April – May 2011	October 2011	March 2013
Grade levels	Grades 2, 3, 4	Grade 2 only	Grade 1 only	Grade 3 only
Coverage	Regionally representative for El-Fayoum, El-Minya, and Qena governorates	Regionally representative for El-Fayoum, El-Minya, and Qena governorates	Non-representative sample, drawn from El-Beheira and Cairo governorates	Nationally representative sample from all 5 regions, 25 of 27 governorates
Schools in effective sample	59 schools (19–20 schools from each governorate)	56 of same schools as in 2009; half treatment, half control schools in each governorate	60 schools (20 in Cairo; 40 in El-Beheira) from selected idaras	200 schools randomly selected across 25 governorates
Students in sample	909 Grade 2 students (and 2000 Grade 3 & 4)	1209 Grade 2 students	1098 Grade 1 students	1992 Grade 3 students
Other	Baseline/follow-up with Grade 2 results permits impact evaluation of early grade reading package provided to 28 treatment schools between September 2010 and April 2011		Purposive, illustrative baseline sample only	First nationally representative EGRA baseline

This brief has been made possible by the generous support of the American people through the United States Agency for International Development (USAID) for the All Children Learning Workshop (Rabat, Morocco 2013). The contents are the responsibility of RTI International and do not necessarily reflect the views of USAID or the United States government.



The results of the 2009–2011 Grade 2 impact study, the 2011 Grade 1 assessment, and the 2013 Grade 3 baseline are presented here as a basis for further reflection and discussion.

Baseline sample results confirm grade progression

NOTE: Because only the Grade 3 (2013) sample is nationally representative, and assessments occurred at different times in the school year for each group, these grade level results are not directly comparable to each other. However, they offer illustrative confirmation that students' reading skills appear to be progressing by grade level.

Average baseline scores achieved

EGRA Subtask	Grade 1 (Oct 2011)	Grade 2 (Jan–Feb 2009)	Grade 3 (Mar 2013)
Letter-name knowledge (clpm)*	12.8 clpm	34.9 clpm	- - -
Letter-sound knowledge (clpm)*	3.4 clpm	9.1 clpm	18.8 clpm
Isolated familiar word reading (cwpm)*	0.9 cwpm	6.4 cwpm	- - -
Decoding of new (invented) words (cwpm)*	0.9 cwpm	5.5 cwpm	5.9 cwpm
Oral reading fluency (cwpm)*	1.2 cwpm	11.0 cwpm	21.9 cwpm
Reading comprehension (% correct of 6)	0.0 / 6	0.7 / 6	32% (1.9 / 6)

* Abbreviations used: “clpm” = Correct letters per minute; “cwpm” = Correct words per minute.

Placed side by side, the results show a natural progression in average reading proficiency from Grade 1 to Grade 2 to Grade 3. As one example, in oral reading fluency average scores began at 1.2 correct words per minute in Grade 1 and rose to 11 correct words in Grade 2 (2009), and to 21.9 correct words in Grade 3.

- For both Grade 1 and Grade 2 students, letter-sound knowledge was much lower than letter-name knowledge, but was more strongly correlated with decoding of new words and with oral reading fluency, suggesting that understanding the relationship of sounds to letters is a specific and important skill for later reading development.
- Decoding skills, which showed some progression between Grade 1 and Grade 2, were still relatively weak among Grade 3 students, despite improvements in letter-sound and in oral reading fluency. Decoding permits the reader to decipher and recognize less-familiar words.
- The leveling off of decoding skill between Grade 2 and Grade 3 suggests that Grade 3 classrooms no longer attended to development of this skill. Indeed, decoding and letter recognition are skills that the Egyptian curriculum expects students to master by Grade 2.

Proportion of students at baseline with zero scores

Baseline results by grade show the opposite trend: zero scores were reduced as grade level increased. The data also revealed some striking gaps in skill for many students.

EGRA subtask	Grade 1 (Oct 2011)	Grade 2 (Jan–Feb 2009)	Grade 3 (Mar 2013)
Letter-name knowledge	31.5%	7.6%	---
Letter-sound knowledge	72.4%	50.1%	18.3%
Isolated familiar word reading	85.6%	52.1%	---
Decoding of new (invented) words	87.1%	57.5%	27.4%
Oral reading fluency	86.9%	47.6%	21.6%
Reading comprehension	98%	70%	35.4%

* Abbreviations used: “clpm” = Correct letters per minute; “cwpm” = Correct words per minute.

- Fully half of Grade 2 students in the 2009 baseline were unable to demonstrate any knowledge of letter sounds, and even more were unable to read a single isolated word or to decode a new (invented) word.
- Almost half of Grade 2 students were unable to read a single word of connected text, and 70% were unable to demonstrate any comprehension of the text.
- While Grade 3 students showed better performance than Grade 2 students at baseline, a substantial proportion of third graders were still unable, toward the end of their second term, to show any skill, even in areas that should have been mastered in Grade 2.

Key results by EGRA subtask for Egypt’s 2013 Grade 3 national baseline

The Grade 3 national baseline data provide important information.

EGRA subtask	Percentage of 2013 Grade 3 students with zero scores	Average score for 2013 Grade 3 sample	Illustrative benchmarks for Grade 3 *	Percentage of 2013 Grade 3 students at or above illustrative benchmark
Letter-sound knowledge (clpm)	18.3 %	18.8 clpm	27 clpm	30 %
Decoding of new (invented) words (cwpm)	27.4 %	5.9 cwpm	14 cwpm	11 %
Oral reading fluency (cwpm)	21.6 %	21.9 cwpm	45 cwpm	16 %
Reading comprehension (% correct of 6)	35.4 %	32% (1.9 / 6)	83% (5 / 6)	9%
MAZE* task comprehension (% correct of 14)	35.4%	26% (3.6 / 14)	86% (12 / 14)	7 %
Listening comprehension in MSA (% correct of 7)	13.3%	46% (3.2 / 7)	86% (6 / 7)	18 %

*MAZE is a task in which a student silently reads a passage and selects an appropriate missing word (multiple choices are provided in the case of maze).

The “illustrative benchmarks” presented in this table are heuristic measures used to indicate only approximate shares of Grade 3 students reading at or near a desired level of proficiency. These benchmarks are not definitive, but facilitate understanding and utilization of the results found and illustrates the utility of setting proficiency benchmarks.

Further discussion and analysis among Egyptian educators and decision makers would be needed to establish formal research- and consensus-based benchmarks.

Key findings of the national Grade 3 baseline include:

Low reading skills across all EGRA measures. Most Grade 3 students had limited pre-reading skills. Too few students could read with sufficient fluency to comprehend the texts. Fewer than 20% of students performed at or above proposed benchmarks for all subtasks. The sole exception: the pre-reading skill of letter sound knowledge (30% at or above benchmark). The average student scores on nearly all subtasks were less than half the proposed benchmark. Students whom teachers observed reading quickly and smoothly were “mechanical” readers or “word-callers” who did not comprehend well what they were reading.

Many Grade 3 students were nonreaders. Even in Grade 3, more than 18% of students could not provide the sound of a single letter; 27% could not demonstrate the ability to decode new words, and nearly 22% were unable to read a single word in connected text. Even more students (over 35%) could not demonstrate any reading comprehension ability. Soon to enter Grade 4, these nonreaders could be unlikely ever to learn to read without remedial instruction. Reducing the percentage of nonreaders should be as essential a measure of improved reading performance as increasing average grade-level scores on specific subtasks.

Both EGRA subtasks on reading comprehension showed poor development of this skill for most Grade 3 students, even though reading comprehension is a priority reading skill to be learned in Grade 3. 35.4% of sampled students could not correctly answer a single reading comprehension question about the short text they had read aloud. Moreover, this percentage does not include the 21.6% of sampled students who could not correctly read a single word on the first line of the reading passage. These zero-score students were not tested for reading comprehension as they did not read the passage.

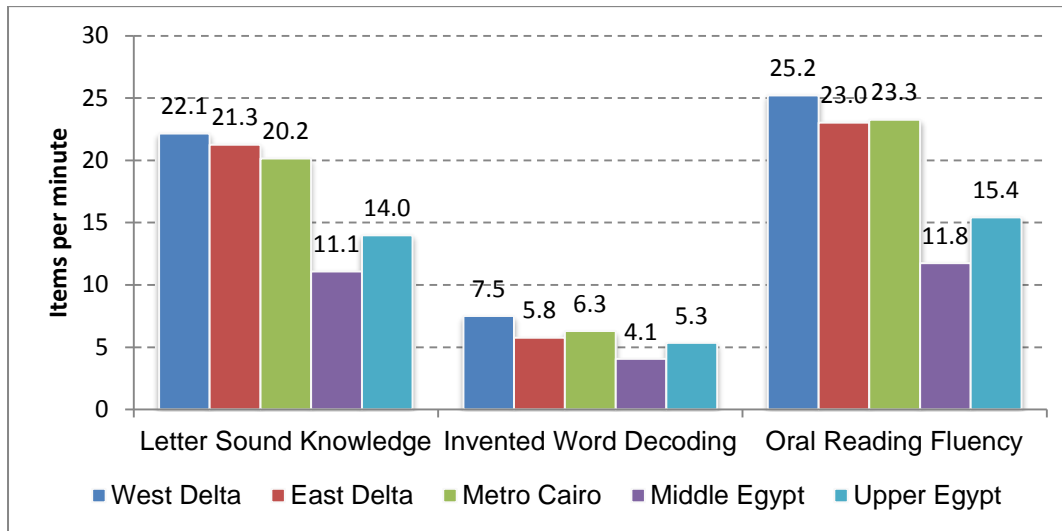
Average performance on the Maze subtask was the lowest of all comprehension tests. This is the most difficult of the comprehension tests. But it is important to note that nonreaders were *not* given the Maze subtask. The low scores would be significantly lower if nonreaders had also been tested on Maze. Nevertheless, more than one-third of students (35.4%) who took the Maze test could not correctly answer a single item. The average score was just 3.6 of 14 total items (26%) answered correctly.

Listening comprehension results were markedly better. Just 13% could not correctly answer a single listening comprehension question. On average, nearly half of the questions were answered correctly (3.2 correct of 7 questions). All sampled students, even nonreaders, were given the listening comprehension test. In an oral culture such as Egypt, it is not surprising that listening skills are stronger than reading comprehension skills. Greater concentration and understanding in listening is, however, desired.



Urban/rural differences and gender differences in reading proficiency were not great. On average, urban students in Grade 3 out-performed their rural peers. But the difference was not great and contributing factors were multiple. There was a wide range of reading performance in most primary schools, both urban and rural, with both strong readers and nonreaders in nearly all schools, often in similar proportion. Similarly, average scores for girls exceeded those for boys on pre-reading skills and oral reading fluency. But there was little difference in comprehension skills.

By region, average performance was similar in West and East Delta and Metro Cairo, but substantially lower in Upper Egypt and particularly Middle Egypt, on letter-sound knowledge and oral reading fluency. All regions performed poorly on decoding new (invented) words.



Overall the 2013 EGRA national baseline results suggest that even by the end of Grade 3, most students had still not yet acquired sufficient foundational skills to read fluently with comprehension in Modern Standard Arabic. The majority of third graders had not mastered recognizing letters in their different forms or applying this knowledge to sound out unfamiliar words. Without these skills, students’ oral reading fluency and reading comprehension also suffered.

Summary of results on impact of GILO/EGRP instruction in Grade 2

In part as a result of the 2009 EGRA findings, Egypt, through the Girls’ Improved Learning Outcomes (GILO) project, embarked on an experimental six-month reading instructional intervention (the EGRP) in a sample of primary schools in three governorates in 2010. Inputs centered on a phonics approach to early grade reading instruction and included teacher training in the method, explicit strategies for learning, vocabulary and comprehension development, highly structured lesson plans, varied student reading materials and extended opportunities for students’ reading practice, and continuous assessment of student performance to inform subsequent teaching and learning. Half of the schools that participated in the 2009 Grade 2 EGRA were randomly assigned to receive the intervention, while the other half were not, and Grade 2 students in both sets of schools were again tested after the intervention was completed. The data thus collected produced an opportunity to examine the impact of EGRP in a scientific manner.

EGRA subtask	Group	2009 baseline scores	2011 post-intervention scores	% change	POST-INTERVENTION	
					Percentage with zero scores	Percentage at or above illustrative benchmark*
Letter-sound knowledge (clpm)	Control	8.6	10.1	+18 %	45.0 %	13.5 %
	GILO / EGRP	9.8	28.5	+192 %	11.1 %	51.2 %
Decoding of new (invented) words (cwpm)	Control	5.6	7.5	+34 %	51.8 %	10.6 %
	GILO / EGRP	7.4	15.5	+111 %	21.4 %	24.2 %
Oral reading fluency (cwpm)	Control	8.9	10.9	+23 %	46.5 %	7.6 %
	GILO / EGRP	11.1	21.1	+91 %	20.7 %	19.7 %

* “Illustrative benchmarks” applied here are: 27 or more syllables per minute on letter-sound knowledge; 25 or more words per minute on decoding of new (invented) words; and 45 or more words per minute on oral reading fluency.

These results indicate a **strong, positive effect of the GILO/EGRP intervention** on second graders’ reading skills in all areas.

Letter-sound knowledge (syllables): Control schools, which did not receive the EGRP intervention, showed modest improvement in average student scores (18% over baseline level), but GILO/EGRP schools showed 192% improvement over baseline scores.

- In addition, by 2011 45% of control school Grade 2 students were unable to provide a single correct answer on this task, and only 13.5% were able to read 27 or more letters correctly in one minute.
- In sharp contrast, only 11.1% of Grade 2 students in GILO/EGRP schools could not provide a single correct answer, and over 51% were able to read at this illustrative benchmark level.

Decoding of new (invented) words: Similarly, control schools showed only modest improvement from 2009 to 2011 (34% gain over baseline level) on Grade 2 students’ average scores, but GILO schools gained 111% over their baseline performance level. By 2011, more than 51% of control school students could not provide a single correct answer, while only 21.4% of GILO / EGRP school students fell in this group.

Considering an illustrative benchmark for this task of 25 or more words read correctly in one minute, by 2011 only 10.6% of students in control schools reached this benchmark, while GILO/EGRP schools more than doubled the proportion (24.2%) of students able to reach the benchmark.

Oral reading fluency: Average student scores in control schools again improved modestly (23% gain over baseline), while scores in GILO/EGRP schools improved substantially (91%). By 2011, the proportion of students unable to show any skill on this task in GILO/EGRP schools (20.7%) was less than half that in control schools (46.5%). Fewer than 8% of students in control schools were able to reach the 45 words-per-minute illustrative benchmark on this task in 2011, against nearly 20% of students in GILO/EGRP schools.



Grade 2 students in GILO/EGRP schools equaled or outperformed students in Grade 3, further supporting the impact of improved reading instruction. GILO/EGRP Grade 2 students’ average scores on letter naming and on decoding of new words markedly exceeded, and their oral reading fluency equaled, the Grade 3 baseline scores of students who had an additional year of schooling. In addition, the percentage of nonreaders dropped more than 50%. This comparison hints at the scale of reading improvement that can be expected from the planned extension of the Egypt Ministry of Education’s EGRP to Grade 3.

II Key Issues for Discussion

The results of the series of EGRA studies presented here raise a number of issues for consideration in efforts to improve early grade learning for students in Egypt. Among the most important of these are the following.

Need for and effectiveness of phonics-rich reading instruction in Grade 1 and Grade 2

The results of the baseline and impact studies presented here underscore the effectiveness of the EGRP to help schools meet the early grade reading challenge that Egypt faces. On the one hand, the Grade 3 baseline confirms that many third graders have not developed or mastered foundational reading skills through the typical instructional practices and materials available in Grade 1 and Grade 2 classrooms.

At the same time, the Grade 2 impact study provides compelling evidence that programs such as GILO/EGRP can help children develop these foundational skills by Grade 2, and that Egypt's teachers and schools are able to effectively implement these programs. The program involves several essential elements: time set aside during the school day to implement the package; explicit, structured lessons; resources for teaching and practicing phonics and other innovative instructional practices; a gender-sensitive environment to ensure positive learning experiences for both boys and girls; intensive initial training of teachers as well as coaching, supervision, and on-site mentoring and additional training opportunities; and other opportunities for reading outside the school and classroom, such as short Summer Reading Programs.

Continued assessment to monitor EGRP implementation progress in Grade 1 and Grade 2

Egypt is currently expanding EGRP to all Grade 1 and Grade 2 classrooms in public primary schools across the country, through an ambitious cascade program of capacity development at all levels of the system. To ensure the same high quality and impact on students' learning, a regular cycle of student learning assessment and analysis, with adjustment and corrective action where needed, is called for.

Further development of higher-order reading comprehension skills in Grade 3 and beyond

EGRP's effectiveness to boost and sustain learners' reading fluency and comprehension well into Grade 3 and beyond has yet to be demonstrated, although the Grade 2 impact analysis results hold promise for such an outcome. Monitoring learning of Grade 3 students who have participated in EGRP in earlier grades will be important to understand EGRP's longer term effects, as well as to inform changes needed in Grade 3 instruction.



Using benchmarks constructively to encourage, measure, and communicate progress

The process of setting benchmarks, whose utility was illustrated in the 2013 Grade 3 baseline analysis, can nonetheless be contentious and politically charged. Performance targets can be threatening to certain interests anchored in the status quo, and genuine differences of opinion exist among dedicated experts and specialists. Having empirical data directly relevant to the task can be extremely useful in **establishing realistic and attainable student learning benchmarks**.

Using learning benchmarks in a constructive manner can help teachers orient and adjust their instruction through continuous assessment, foster and measure meaningful system-wide efforts to improve outcomes, and more effectively communicate the results of these efforts to broader audiences and the general public. Is Egypt on track in its setting and application of meaningful student learning benchmarks?

REFERENCES

- RTI International (2011). *Improved reading performance in Grade 2, GILO-supported schools vs. Control schools*. Cairo: Ministry of Education and USAID/Egypt (GILO Project).
- RTI International (June 2012). *A Reading Snapshot: Results of the Early Grade Reading Assessment in El-Beheira and Cairo Governorates*. Cairo: Ministry of Education and USAID/Egypt (GILO Project).
- LaTowsky, R., C. Cumiskey, and P. Collins (2013). *Egypt Grade 3 Early Grade Reading Assessment Baseline* (Report for EdData II Task Order No. 15). Washington, DC: USAID (EdData II Project).