



USAID
FROM THE AMERICAN PEOPLE

EdData II

Early Grade Reading Assessment Grade 2 Baseline, West Bank



**EdData II Technical and Managerial Assistance, Task Order
Task Order 15
September 2014**

This publication was produced for review by the United States Agency for International Development. It was prepared by RTI International.

Early Grade Reading Assessment Grade Two Baseline, West Bank

EdData II
Task Order No. 15

Prepared for
USAID/Washington
Data for Education Programming in Asia and the Middle East (DEP-AME) contract

Prepared by
Robert J. LaTowsky and Susan Edwards
RTI International
3040 Cornwallis Road
Post Office Box 12194
Research Triangle Park, NC 27709-2194

RTI International is a trade name of Research Triangle Institute.

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Table of Contents

Section	Page
List of Figures	iii
List of Tables	iii
Abbreviations	iv
Acknowledgments.....	v
Executive Summary	vii
Background.....	1
Context of Reading Outcomes in the Primary Grades in West Bank	1
USAID Support for Early Grade Reading in the West Bank.....	1
Why Test Early Grade Reading?	2
Purposes and Uses of this West Bank Baseline for Grade 2	3
What EGRA Measures	3
The Electronic and Paper EGRA Instruments.....	8
Implementing this West Bank Baseline for Grade 2.....	8
Developing and Testing the Instrument	8
Training the Assessors	8
The Sample Schools and Students.....	9
Field Implementation.....	9
West Bank EGRA Baseline Findings for Grade 2	10
Summary Scores and Levels of Student Performance	10
Letter Sound Knowledge.....	14
Familiar Word Reading	16
Nonword / Invented Word Reading	17
Oral Reading Fluency – with and without diacritics.....	18
Reading Comprehension – 60 seconds with diacritics.....	20
Reading Comprehension – <u>90 seconds without diacritics</u>	21
Listening Comprehension	23
Gender Differences in Reading Performance	25
School Gender Differences in Reading Performance	26
Summary Conclusions	29
Priority Recommendations and Next Steps	31
Annex A: The West Bank EGRA Instrument for Grade 2	33
List of Participants: MOEHE Tool Development Team for the EGRA Instrument ...	33

Annex B: Sample Design and Weighting	44
Annex C: The Benchmarking Workshop – Agenda, Key Presentation, Proposed Benchmarks, and List of Participants	49

List of Figures

Figure 1:	Percentages of Grade 2 Students with Zero Scores on the EGRA Subtasks	12
Figure 2:	Percentages of Grade 2 Students reading at or above Average Score Benchmarks on the EGRA Subtasks.....	13
Figure 3:	Distribution of Grade 2 Students on the Letter Sound Identification Subtask	15
Figure 4:	Distribution of Grade 2 Students on the Familiar Words Subtask.....	16
Figure 5:	Distribution of Grade 2 Students on the Nonword / Invented Word Reading Subtask.....	17
Figure 6:	Distribution of Grade 2 Students on the Oral Reading Fluency Subtask – <u>with</u> diacritics	19
Figure 7:	Distribution of Grade 2 Students on the Oral Reading Fluency Subtask – <u>without</u> diacritics	20
Figure 8:	Distribution of Grade 2 Students on the Reading Comprehension Subtask -- <i>60 seconds with diacritics</i>	21
Figure 9:	Distribution of Grade 2 Students on the Second Reading Comprehension Subtask -- <i>90 seconds without diacritics</i>	22
Figure 10:	Distribution of Grade 2 Students on the Listening Comprehension Subtask	24
Figure 11:	Reading Proficiency of Grade 2 Boys and Girls	25
Figure 12:	Selected Average Subtask Scores by School Gender.....	26
Figure 13:	Variability across MOEHE Schools in Grade 2 Reading Performance ..	28

List of Tables

Table ES1:	Summary Scores for All EGRA Subtasks	viii
Table 1:	Summary of West Bank EGRA Average Scores for Grade 2	11
Table 2:	Summary of EGRA Scores for the Number of Items Attempted	14
Table 3:	Average Subtask Scores by Gender in Mixed and Single-Gender Schools	26
Table 4:	Average Subtask Scores by Class Size	27

Abbreviations

AED	Assesment and Evaluation Department, MOEHE
clspm	Correct Letter Sounds Per Minute
cnonwpm	Correct Nonwords Per Minute
cwpm	Correct Words Per Minute
EGRA	Early Grade Reading Assessment
MOEHE	Ministry of Education and Higher Education
MSA	Modern Standard Arabic
NGO	Non-Governmental Organization
ORF	Oral Reading Fluency
RTI	Research Triangle Institute, a trademark of RTI International
USAID	United States Agency for International Development

Acknowledgments

The authors wish to acknowledge with appreciation and respect the contributions and support of education staff from USAID/West Bank and Gaza (WBG) Mission and the Palestinian Authority's Ministry of Education and Higher Education (MOEHE) in the implementation of this Early Grade Reading Assessment (EGRA) baseline for Grade 2 in the West Bank. We thank Robert Davidson, Bassam Kort, Micheline Sleibi, and Fouad Salman at the USAID/WBG Mission for their task leadership, administrative guidance and effective liaison with concerned parties in that unique context of the West Bank; and Mitch Kirby of USAID/Washington for his direction under the Data for Education Programming in Asia and the Middle East (DEP-AME) contract. At the MOEHE, the EGRA Steering Committee led by Dr. Basri Saleh, Assistant Deputy Minister for Planning and Development, together with Dr. Tharwat Zeid, Director General of the MOEHE Department of Qualification and Supervision (QSD), and Dr. Mohammad Matar, Director of the Assessment and Evaluation Department (AED), assured effective Ministry mobilization and support of participants, planning and field implementation at all levels, from start to finish. The success of this EGRA owes a debt to the more than fifty MOEHE supervisors and candidate teachers who directly implemented the assessment as trained assessors and assessor team leaders and the consultant Field Coordinators—Dr. Khalid al-Jawarish, Najeh Abu Shamsiya, Nariman Daoud Emfarah, and Lina Al-Khateeb—who capably managed and directed them as 16 field teams. Scores of school, management, district, and central ministry staff facilitated the field assessments in schools and the random pre-selection of sample Grade 2 students by providing complete enrollment lists for each school. The Arabic instrument for this assessment was developed by the EGRA Tool Adaptation Team comprised of eight Ministry representatives together with Dr. Wael Salah, RTI consultant, and Dr. LaTowsky. Adnan Al-Harazi, Osama M. Al-Fagih, and Hamdi Al-Weshah of Prodigy Systems, Inc., expertly supported the development and adaptation of the electronic instrument on iPads and final compilation of the database. Dr. Robert J. LaTowsky, Managing Director of Infonex, was Field Director for this EGRA and lead author of this report. Dr. Simon King supported and Susan Edwards of RTI led the data quality control, calculation of scores, and statistical analysis of the data. Steven Keller, Country Director, and Hani Bakri, Finance and Administration Manager, of AMIDEAST/WBG provided valuable assistance and agency support services to this EGRA. At RTI International, Medina Korda expertly served as the overall task leader and Michelle Ward-Brent backstopped and guided this assessment, especially at its initiation; David Harbin did the same for its financial and contractual management. Dr. Luis Crouch, Vice-President and Chief Technical Officer of RTI's International Development Group, co-led with Dr. LaTowsky the presentation of EGRA results to the USAID/WBG Mission and the MOEHE in late May 2014 and met with senior leadership of the USAID Mission and the Consul General to share thoughts on the strategic significance of these EGRA results for future Mission programming in education. We sincerely thank them all for their contributions to the success of this assessment.

At MOEHE request, Drs. Luis Crouch and Robert LaTowsky returned to Ramallah in September 20-25, 2014 to lead two workshops of technical and policy support to the ministry. The first workshop (September 21-22) supported the MOEHE in proposing benchmarks for Grade 2 reading skills. These benchmarks, prepared and proposed by the workshop's participants with expert inputs, are applied here in the analysis and presentation of EGRA results in this Final Report. The second workshop (September 24-25) strengthened MOEHE capacities to analyze and present future results of EGRAs that they may conduct themselves. Both workshops significantly deepened MOEHE engagement with early grade reading.

Executive Summary

This report presents key findings of the first baseline of early grade reading skills implemented in the West Bank. This EGRA for Grade 2 was conducted in March 2014 in a representative sample of 150 Ministry of Education and Higher Education (MOEHE) primary schools, stratified by school gender and selected randomly from 16 districts in the West Bank. The 2,953 tested students were randomly selected from Grade 2 enrollment lists prior to each school visit. The results are representative of MOEHE Grade 2 students and primary schools in West Bank.

The findings of this national baseline inform the strategic design and development of MOEHE enhancements to the curriculum, resources and teaching of early grade reading. Implemented by MOEHE supervisors and candidate teachers, this EGRA strengthened ministry capacities and deepened its knowledge base of early grade reading and reading assessment.

This assessment of Grade 2 reading skills in the formal language of primary school instruction, Modern Standard Arabic, comprised 6 subtasks: the pre-reading skills of letter sound identification, familiar and nonword reading learned in Grades 1 and 2 plus oral reading fluency, and two comprehension subtasks: reading comprehension and listening comprehension. Of particular note, this Arabic EGRA was the first to include two (2) oral reading passages – one *with* full diacritics and the other *without* diacritics. Comparing student reading fluency and comprehension on these two exercises is an MOEHE priority and a significant contribution of this EGRA.

The summary scores for all subtasks of this West Bank EGRA are shown in **Table ESI** below. Key findings of this Grade 2 baseline include:

1. ***Good foundation of pre-reading skills:*** The sample Grade 2 students demonstrated good proficiency in the basic pre-reading skills of letter sounds knowledge and nonword / invented word decoding. Average scores on these two subtasks are, by far, the highest of any Arabic EGRA in Grades 2 or 3; and the percentage of zero scores are the lowest. **There is a strong foundation of reading proficiency in these pre-reading skills among MOEHE Grade 2 students. This is a strong and positive finding.**
2. ***Students struggle to read connected text in passages:*** Most Grade 2 students are struggling to read familiar words in connected text (passages, stories), indicative of low reading fluency. Average scores on the oral reading fluency subtask (with diacritics) are low and the percentage of students scoring zero is high.
3. ***Low reading comprehension:*** Most Grade 2 students do not comprehend what they are reading. Average scores on reading comprehension (60 seconds reading with diacritics) are low and the percentage of students with zero scores is high. The generally good results on the listening comprehension subtask, however, suggest that most Grade 2 children capably understand formal Arabic. Hence **the low scores on reading comprehension in Grade 2 chiefly result from low reading fluency and undeveloped reading comprehension**

skills. Comprehension skills must be learned, practiced and mastered – they do not come automatically to students who mechanically read well.

Table ES1: Summary Scores for all EGRA Subtasks

Subtask	Percentage of Grade 2 students with zero scores	Grade 2 average score	Proposed benchmark (Sept 2014)	Percentage of students performing at or above benchmark
Letter sound identification (clpm)	6.8%	36.0	45	41%
Familiar Word reading (cwpm)	10.0%	17.5	na ¹	na
Nonword reading (cnonwpm)	8.7%	13.9	20	29%
Oral reading fluency (ORF) – with full diacritics (cwpm)	22.1%	16.7	30	16%
Oral reading fluency (ORF) – without diacritics (cwpm)	10.9%	24.9	35	27%
Reading comprehension – ORF with diacritics (max. 6)	35.7%	1.6	3.0 (50% correct)	26%
Reading comprehension – ORF without diacritics (max. 6)	25.9%	2.2	na	na
Listening comprehension (max. 6)	7.1%	3.3	3.0 (50% correct)	68%

4. *Few illiterates in Grade 2: The percentage of illiterate Grade 2 students in MOEHE schools is low: less than 5%. This is an important and positive finding.* Few students are now being “left behind” in early grade instruction in Arabic. The number of nonreaders can and should, however, be reduced. For all reading subtasks, targets were set at the benchmarking workshop for lower percentage zero scores to be achieved by 2018. *It is important that these targets for fewer nonreaders be a top priority for benchmark achievement.*
5. *Low variability among MOEHE schools in Grade 2 reading proficiency:* There is low variability (low divergence) among MOEHE primary schools in the reading proficiency of Grade 2 students. This finding has important social equity implications. The large majority of MOEHE schools are performing comparably in Arabic reading instruction of their Grade 2 students. There are not significant numbers of “strong” schools and “weak” schools; nor a large divide between schools in the reading performance of their Grade 2 students. The overall EGRA results are “true” and general indicators of Grade 2 reading outcomes in most MOEHE schools. Most Grade 2 classes include a full range of student reading abilities, from weak to strong readers.

¹ At the benchmarking workshop held with the MOEHE in September 2014, benchmarks were not proposed for two subtasks included in this West Bank EGRA for Grade 2: the familiar words and reading comprehension for 90 seconds without diacritics. Nor was a benchmark proposed for listening comprehension. The benchmark shown in this Table for listening comprehension is illustrative only. On this point, see the text box on page 10.

6. ***Significant gender gap in Grade 2 reading proficiency:*** Girls typically outperform boys in most early reading subtasks, especially the timed tests of pre-reading skills and oral reading fluency. This result is common to the large majority of languages and countries in which EGRAs have been implemented. It is true of all Arabic EGRAs. The gender gap in reading proficiency for this sample of West Bank Grade 2 students is, however, the largest of all Arabic EGRAs. The results cannot be extrapolated to later grades. The gender gap may diminish among older students. These Grade 2 results do recommend that [schools, teachers and parents be aware of this gender gap and ensure that boys have equal or greater opportunity for reading practice in class.](#)

Background

Context of Reading Outcomes in the Primary Grades in West Bank

Diminished proficiency in early grade reading of formal Arabic is a critical challenge across the Middle East and North Africa (MENA) region. Standardized tests administered by the MOEHE and UNRWA over the past decade in the schools they respectively administer across the West Bank, have showed declining learning outcomes and learning quality in the primary grades. In its Annual Report of 2012, the MOEHE reported that less than 50% of 7th grade boys in its schools and just 56% of 7th grade girls were reading Arabic at grade level. Only 60% of 7th grade students in UNRWA passed the 7th grade Arabic exam that year. The average scores of students in the 2012 National Arabic Assessment were only slightly better than the low results reported in 2008. Altogether, these findings recommended specific attention by the MOEHE to improved reading outcomes in the early grades. The MOEHE's new strategic five-year plan (2014-2019) emphasizes improved teaching and learning of Arabic as a foundational learning skill.

USAID Support for Early Grade Reading in the West Bank

USAID support for this EGRA seeks to generate regional and country-specific education data, and analyses of those data, that can be used by the MOEHE to prioritize education needs and future investments. It also aims to strengthen local skills in the design, evaluation and management of education programs, and quality data capture and analysis to support them. This activity resulted in a representative EGRA of Grade 2 students in West Bank public schools. The results will be used to inform policy decisions in the areas of reading assessment and teacher training methodologies.

Proposed MOEHE Benchmarks for Grade 2 Reading Skills

The first presentation of baseline EGRA results for Grade 2 to the MOEHE and USAID Mission for West Bank and Gaza in late May 2014 included *illustrative* “benchmarks” for specific reading skills. These heuristic benchmarks were included to answer the expected question: “What percentage of Grade 2 students were able to read *proficiently*, i.e. at the standard that might be desired for Grade 2 reading performance?” These *illustrative* benchmarks were set at the average score at or above which one-third (33%) of Grade 2 students performed on the timed tests of this EGRA baseline.

Subsequent to this presentation, the MOEHE requested technical and policy support from RTI in establishing Palestinian benchmarks for Grade 2 reading skills based on the actual results of the EGRA baseline and informed by international experience in early grade reading and scientific understandings of early reading cognition. The workshop, conducted September 21-22, 2014 in Ramallah with MOEHE and selected NGO representatives, collaboratively prepared benchmarks for Grade 2 reading skills

to be achieved in three years – by spring 2018 (2017/2018 school year). The workshop agenda, presentations and participants are reported here in *Annex C*. These proposed benchmarks have now been presented by the workshop’s participants to senior ministry officials for their policy discussion and possible adoption as official MOEHE benchmarks for Grade 2 reading.

This Final Report applies these latest, proposed benchmarks in presenting the results of the West Bank EGRA baseline for Grade 2. The benchmarks are specified in the subtask analyses of results on each subtask. The first presentation of West Bank EGRA results in May 2014 that applied illustrative benchmarks is also included here in *Annex D* as documentation of this baseline EGRA.

Purposes and Design of the West Bank EGRA Baseline for Grade 2

Why Test Early Grade Reading?

The ability to read and understand a simple text is one of the most fundamental skills a child can learn. Without basic literacy there is little chance that a child can escape the intergenerational cycle of poverty. Yet in many countries, students enrolled in school for as many as six years are unable to read and understand a simple text. Recent evidence indicates that learning to read both *early* and at a sufficient *rate* is essential for learning to read well. A substantial body of research documents the fact that children can learn to read by the end of Grade 3, and indeed need to be able to read to be successful in school. Acquiring literacy becomes more difficult as students grow older; children who do not learn to read in the early grades (Grades 1–3) are likely to fall behind in reading and other subjects, likely to repeat grades, and eventually to drop out of school.²

When children are first learning to read in Arabic, they must learn the letters and their forms, learn the sounds associated with each letter and diacritical marks, and apply this knowledge to decode (or “sound out”) new words that they can recognize instantly.³ By the end of this first phase, children develop sufficient speed and accuracy in decoding and word recognition that they can read with fluency. When children read with fluency, they can read orally with speed and expression similar to what they use in speech. Furthermore, reading with fluency is critical for reading comprehension, because children can concentrate on the meaning of what they read rather than having to focus on decoding.^{4,5}

² RTI. (2009). *Early Grade Reading Assessment Toolkit*. Research Triangle Park, NC: World Bank Office of Human Development, 1.

³ See E. Saiegh-Haddad. (2005). Correlates of reading fluency in Arabic: Diglossic and orthographic factors. *Reading and Writing: An Interdisciplinary Journal*, 18, 559–582. See also M. Taouk & M. Coltheart. (2004). The cognitive processes involved in learning to read in Arabic. *Reading and Writing: An Interdisciplinary Journal*, 17, 27–57.

⁴ S. Abu-Rabia. (2007). The role of morphology and short vowelization in reading Arabic among normal and dyslexic readers in Grades 3, 6, 9, and 12. *Journal of Psycholinguistic Research*, 36, 89–106.

Purposes and Uses of this West Bank Baseline for Grade 2

This baseline EGRA for Grade 2 establishes the capacities and deepens the knowledge base of MOEHE staff to implement early grade reading assessments. All EGRA assessors and assessor team leaders were MOEHE supervisors or candidate teachers and all planning, training, implementation, and dissemination was conducted in close technical collaboration with the Ministry’s Assessment and Evaluation Department (AED) and EGRA Steering Committee. Central ministry and MOEHE districts actively supported the field implementation with school liaison, orientation, and enrollment lists for sample selection.

The results inform policy decisions and planning by the Palestinian Authority (PA) for improved reading instruction and student learning outcomes in the early grades. The longer term objective is enhanced teacher training and learning resources for improved reading proficiency by primary students. The findings of this EGRA will infuse the design and development of MOEHE curricula and teaching resources for enhanced reading instruction in the earliest grades.

Thirdly, this EGRA establishes a baseline for MOEHE use in measuring and reporting future progress in enhanced reading performance in the early grades. Finally, the results of this EGRA contribute to raising the awareness of parents, teachers and school leaders to the importance and challenges of early grade reading.

What EGRA Measures

The EGRA instrument is composed of various subtasks designed to assess foundational reading skills that are crucial to becoming a fluent reader. EGRA is a method-independent approach to assessment—that is, the instrument does not reflect a particular method of reading instruction (i.e., “whole language” or “phonics-based” approach). Rather, EGRA measures basic skills that a child must have to eventually be able to read fluently and with comprehension—the ultimate goal of reading. The EGRA subtasks are based on research for a comprehensive approach to reading acquisition across languages. These foundational reading skills are described below:

- The **alphabetic principle** is considered essential for learning to read an alphabetic language. The alphabetic principle refers to the recognition and understanding that speech sounds (phonemes) are represented by units of print such as letters and diacritics (graphemes). Thus, mastery of the alphabetic principle is the understanding that there are predictable relationships between sounds and the symbols that represent them. It is necessary for mastering spelling patterns and their relationship with oral language through the letter-sound (grapheme-phoneme) correspondences.
- **Oral reading fluency** is often defined as the ability to orally read connected text with speed, accuracy, and proper expression. Reading fluency is considered critical for comprehension, because rapid, effortless word-

⁵ G. Elbeheri, J. Everatt, A. Mahfoudhi, M. A. Al-Diyar, & N. Taibah, (2011). Orthographic processing and reading comprehension among Arabic speaking mainstream and LD children. *Dyslexia*, 17(2): 123–142.

identification processes enable the reader to focus on the text and its meaning rather than decoding, or sounding out the words.⁶

- **Reading comprehension**, considered the goal of reading, refers to the ability to actively engage with, and construct meaning from, the texts that are read.
- **Listening comprehension** refers to one’s ability to make sense of oral language in the absence of print. Listening comprehension taps many skills and sources of knowledge, such as vocabulary knowledge, facility with grammar, and general background knowledge. Assessing listening comprehension is particularly important for a diglossic language such as Arabic, because children are often not introduced to the formal dialect until after they begin formal schooling. Thus, listening comprehension assesses children’s proficiency with the formal dialect of Arabic.

EGRA measures each of the above abilities/components to assess foundational reading skills. These skills are tested in individual subtasks and presented in order of increased level of difficulty (i.e., letter sound identification, then familiar word reading, then nonword reading, etc.). In general, initial subtasks are easier than later subtasks. The listening comprehension subtask is the exception: it best follows the reading comprehension subtask in implementation but is typically easier for students. EGRA thus effectively measures a wide range of reading abilities for beginning readers. The specific subtasks included in the EGRA instrument developed for this Grade 2 baseline are described below.

EGRA Measures for the West Bank Grade 2 Baseline

The EGRA instrument is implemented one-on-one—an assessor with a single student—and requires 15-20 minutes to complete. The Arabic instrument developed for this Grade 2 assessment is presented in *Annex A*. This electronic instrument (iPads) used for this assessment included the following subtasks (subtasks) implemented in this order:

1. **Letter sound identification** assessed children’s automaticity in their knowledge of the sounds associated with each letter. This was a timed subtask, in which children were shown a chart containing 80 letters with diacritics arranged in 8 rows each with 10 letters. All letters were in initial word or independent letter form with a diacritic. Students were asked to produce the sounds associated with each letter as quickly and accurately as they could within one minute, yielding a score of correct letter sounds per minute (clspm).
2. **Familiar Word reading** assessed children’s knowledge and automaticity in reading familiar words. All 50 words in this timed subtask requested by the

⁶ National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (National Institutes of Health Publication No. 00-4769). Washington, DC: U.S. Government Printing Office. See also C.A. Perfetti. (1992). The representation problem in reading acquisition. In P.B. Gough, L.C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Erlbaum.

MOEHE were selected from the current Grade 2 Arabic language textbooks for terms 1 and 2. Arranged in 10 rows of 5 words each, all words included diacritics on all but the final syllable. The subtask comprised approximately equal numbers of 2-, 3- and 4-syllable words plus a small number of 1-syllable words. Students were asked to correctly pronounce the words *with* their diacritics, reading as quickly and accurately as possible for one minute, yielding a score of correct words per minute (cwpm).

3. **Nonword reading** assessed children's skill at applying letter-sound correspondence rules to decode (i.e., sound out) unfamiliar words. To ensure that children were applying their knowledge of the relationships between sounds and symbols rather than reading words from memory, children were asked to read a chart of 50 pronounceable made-up words (invented or nonsense words) arranged in 10 rows of 5 words each. All items included diacritics on all but the final syllable and were constructed in accordance with Arabic orthography. The nonwords created for this subtask were largely unique. However, they replicated the structure and placement of nonwords in that subtask of the Egypt baseline EGRA for Grade 3: the same number of 2-letter, 3-letter, 4-letter, 5-letter, and 6-letter words *arranged in identical placement* on the chart. Before starting the subtask, each student was instructed by the assessor to correctly read aloud three practice nonwords. Children were then asked to correctly sound out as many nonwords *with* their diacritics as they could within one minute, yielding a score of correct nonwords per minute (cnonwpm).
4. **Oral passage reading** assessed children's fluency in reading a passage of grade-level text aloud and their ability to understand what they had read. Two different versions of this subtask were implemented consecutively for this EGRA – always in the same order – with each student. Each version consisted of two parts:
 - a. **Oral reading fluency – 60 seconds with diacritics**: The ability to read passages fluently is considered a necessary component for reading comprehension. In the two versions of this subtask, children were given a different story of local relevance comprised of 57 words. Both stories and their questions were in formal Arabic. For the first version, the words were written *with diacritics* for all syllables (except the final syllable) and children were asked to read aloud the words correctly with their diacritics. For this first version of the subtask, they were given 60 seconds to read – the standard duration for this EGRA subtask. Before starting, each child was instructed to pay attention to the story as they read because he or she would be asked questions about the story after finishing. The oral reading fluency score was the number of correct words read per minute (cwpm).
 - b. **Reading comprehension – 60 seconds with diacritics**: After the children finished the passage, or the one minute ended, the story was removed. The assessor then asked questions that required children to recall basic facts from the passage or the part they read. The maximum number of questions asked was 6. But children were only asked those questions that

could be answered from the specific narrative that they had read.⁷ The reading comprehension score was the number of correct answers, with a maximum possible score of 6. This subtask was untimed but students who did not reply to a specific question within ten seconds were scored as “No Reply” on that question. Each question was asked only once with no repeat.

- c. ***Oral reading fluency – 90 seconds without diacritics:*** The second story directly followed the reading comprehension subtask of the first passage reading. The 2nd story was identical in length (57 words), comparable in difficulty and as locally relevant to Grade 2 students as the 1st story. The 2nd story differed in just 2 respects: i) children were given 90 seconds to read instead of 60 seconds, so as to give more time for children to read more of the story, and ii) the story did not include diacritics. Otherwise, this subtask was implemented exactly as the 1st oral reading fluency test. The aim of this 2nd story was to compare reading fluency and reading comprehension outcomes when these two variables were changed. Children were scored on this 2nd oral reading fluency subtask on whether they correctly pronounced the letters in the words, regardless of what short vowels they used for the syllables (without diacritics). To compare the results of the two stories, the same EGRA metric was used for both: the number of correct words read per minute (cwpm).
 - d. ***Reading comprehension – 90 seconds without diacritics:*** Immediately after completing the 2nd story, or at the end of the 90-second duration, students were asked direct questions (no inferential questions) about the narrative that they read. This 2nd reading comprehension subtask was implemented and scored exactly as the 1st reading comprehension test.
5. ***Listening comprehension*** is considered a critical skill for reading comprehension because it shows the ability to make sense of oral language. In this subtask, the examiner read clearly and at moderate pace (approximately 0.5 seconds per word) a short narrative story of 58 words to the children – comparable in length to the 57-word oral reading passages. Before starting, the assessor instructed each child to listen carefully as he or she would be asked several questions about the story. After hearing the passage, each child was asked all 6 questions, always in the same order and exactly as written in formal Arabic. The listening comprehension score was the total correct answers, with a maximum possible score of 6. This subtask was untimed but students who did not reply to a specific question within ten seconds were scored as “No Reply” on that question. Each question was asked only once. All students were given the listening comprehension subtask, even those nonreaders who had been “early stopped” on one or more previous subtasks.

⁷ Weak and slow readers were not asked questions for text they did not read in the one minute. Students who did not correctly read any word on the first line of the story (7-8 words) – zero scores – were not asked any reading comprehension questions. The iPads “activated” only those reading comprehension questions that could be answered from the text read by that student in the 60- and 90-second timed tests of oral reading fluency.

All written components of the EGRA were in Modern Standard Arabic, including the stories and all questions in the reading comprehension and listening comprehension subtasks. All assessors asked the subtask questions and read the listening comprehension passages exactly as written in formal dialect without variance. The oral instructions given to children for each subtask, however, were explained by assessors in the home language of Palestinian dialect. These instructions were written on the instrument in formal Arabic but presented orally by the assessor, as written, in simple, vernacular Arabic. Children were asked to confirm that they fully understood the instructions before starting each subtask. Once started, no subtask was interrupted. The only comment permitted for an assessor to make was to say “go on” after three seconds to a student stalled on a specific letter or word in one of the timed subtasks.

In administering the EGRA, assessors were very attentive to making each child feel comfortable and at ease. The child’s name was not recorded and assessors presented the test as a “game” that the child would enjoy and an “experimental activity” to test the instrument’s utility. Assessors were very explicit that the EGRA was not an exam and students were not being graded. Participating students were told they were lucky to have been



chosen for this experiment to test the instrument. Before beginning each assessment, children were pointedly asked for their assent to participate in the assessment. Any child who declined was thanked and invited to leave. Very few children refused to participate. At the end of the assessment and regardless of how well they read, the great majority of children responded, when asked if the assessment was difficult or easy, that it was “easy.”

Many children, however, were nonreaders or limited readers on some subtasks, especially the more difficult subtasks. For these students, all timed subtasks – letter sound knowledge, familiar word, nonword, and oral reading fluency subtasks – included an “early stop” rule that required assessors to discontinue the subtask if a child did not respond correctly to any of the items on the first line (i.e., the first 10 letters, the first 5 nonwords, or the first line of 7-8 words of the oral reading fluency story). If just one item on the first line was read correctly, the subtask continued through its full time period. This rule was established to avoid frustrating children who did not understand the subtask or lacked the reading skills to respond. If a subtask was halted by the “early stop” rule, the assessor went on the next subtask. If the oral reading fluency subtask, however, was halted by the “early stop” rule, the student was not asked any of the reading comprehension questions. All subtasks halted by the “early-stop” rule were marked clearly.

The Electronic and Paper EGRA Instruments

This EGRA was largely conducted using an electronic tool. After first developing and pre-testing the instrument in paper format, an identical electronic version was prepared for implementation using iPads. All subtasks were implemented in the same order with the same rules. The development and functionality of the iPad data collection program used for this West Bank EGRA closely followed the iPad program developed for the Egypt EGRA baseline in 2013.

Implementing this West Bank Baseline for Grade 2

Developing and Testing the Instrument

The EGRA instrument was developed with a team of Arabic language and reading experts in February 2014 following specific technical guidelines for key subtasks to enhance the standardizing and comparability of Arabic EGRA instrument design. The resulting EGRA instrument was then tested, using a paper format, with 22-25 students in each of four MOEHE primary schools in Ramallah districts. The MOEHE selected the schools and the school principals and Grade 2 teachers selected the students for this pilot implementation. The sole criterion for selection was that two schools be “weaker” schools of generally lower learning achievement and the other 2 schools be “strong” schools of superior student performance. Similarly, principals in each school were asked to provide an equal mix of weaker and more capable readers from Grade 2 to be tested. In most schools, all or most Grade 2 students were tested. The purpose of this pilot—to test that the instrument was neither too easy nor too difficult and appropriate for the range of reading abilities in Grade 2—was explained at each school and school officials were supportive and complied fully. The pilot confirmed the effectiveness and reliability of the instrument in differentiating a wide range of reading abilities.

Development of the electronic EGRA tool proceeded in parallel with the testing of the paper instrument. The final paper instrument is included in [Annex A](#).

Training the Assessors

The MOE identified thirty-five candidates from candidate MOEHE early grade teachers (assistant teachers awaiting appointment) across the West Bank to be trained as EGRA assessors. The large majority of candidates were younger women (both married and single) in their late twenties and thirties. All candidates were trained for 6.5 days in early March with the expectation that all assessors would be selected. The first three days familiarized assessors with the subtasks and trained assessors to implement the paper form. The final 3.5 days trained assessors to use the electronic (iPad) instrument and complete their familiarization with subtask content. It is very important that assessors learn thoroughly and by heart the sound of all subtask content so that they can concentrate fully on listening to students’ reading and not have to check each sound that they hear against the written form. This familiarity comes from repeated drilling of assessors on all subtasks using both the paper and

electronic tools. All assessors were continuously monitored for proficiency in applying both instruments. All training was conducted in Arabic. The 32 assessors who successfully completed the training and conducted all assessments are listed in *Annex A*.

The Sample Schools and Students

The population and random, stratified sample of 150 schools and Grade 2 students selected from all West Bank districts for this EGRA baseline are presented in *Annex B* along with the sample design. At MOEHE request, the random selection of the 150 schools applied stratification by school gender: 50 schools were girls' only primary schools with Grade 2, 50 schools were boys' only schools, and 50 schools were mixed (co-ed) schools with Grade 2. From single-gender schools, 20 children were randomly selected from enrollment lists for all Grade 2 classes in each school with approximately equal numbers drawn from each class. From mixed (co-ed) schools, 10 boys and 10 girls were selected from across all Grade 2 classes. Random number tables were used to select the sample students from numbered enrollment lists from each school. A reserve list of 4 students was also pre-selected for each gender in each school to replace any sample students absent on the day of the assessment visit to that school. A complete list of replacement schools was also provided with a specific replacement school indicated for each sample school if implementation in a given sample school was not possible. In fact, no replacement schools were needed. The sample used only the original list of sample schools drawn by RTI statisticians. The list of 20 Grade 2 sample students and additional reserve students for each school was drawn by the Field Director and four Field Coordinators responsible for field implementation of this baseline EGRA.

Field Implementation

The 32 assessors and 16 assessor team leaders were organized into 16 assessor teams, each with two assessors and one assessor team leader. Four teams were each directed by four Field Coordinators deployed across the West Bank. Each team assessed twenty students in a school each day. With each assessor testing 10 students, the school assessment typically required 2.5 – 3.0 hours. The assessor team leader was responsible for gathering and confirming the identity of the randomly pre-selected students for each school and delivering them, one by one, to the two assessors conducting the tests in the school library or classroom vacated for their use. Assessor team leaders ensured that students and assessors were not disturbed or interrupted.

The teams typically completed 16 schools per day. The assessment of all 150 schools was completed in twelve school days in the second half of March.

West Bank EGRA Baseline Findings for Grade 2

Summary Scores and Levels of Student Performance

This section presents summary statistics for all subtasks of this West Bank EGRA for Grade 2. First, we identify the percentage of nonreaders for each subtask, i.e. the percentage of students scoring zero on that subtask. Then we report the average scores,⁸ identify the proposed Grade 2 benchmarks for each subtask, and report the percentage of students now reading at or above these benchmark scores.

The benchmark scores presented in *Table 1* are those developed by participants of the benchmarking workshop held with key MOEHE and NGO participants in September 2014. No benchmarks were proposed at the workshop for the two optional subtasks of familiar words reading and reading comprehension for 90 seconds without diacritics. These are shown as “na” = “not applied” in Table 1. The workshop also did not propose a benchmark for the listening comprehension subtask. See the textbox here for explanation of the *illustrative* benchmark applied in this Report for listening comprehension.

Table 1 below reveals strong performance in the pre-reading skills learned in Grades 1 and 2 and in listening comprehension. Many Grade 2 students have mastered these pre-reading skills of letter sounds knowledge and nonwords reading and are reading at benchmark levels. Forty-one (41%) of sample students read letter sounds at benchmark level and 29% read unfamiliar (nonwords) words with strong proficiency. The percentages of students with zero scores on these first two subtasks were also low: 6.8% and 10% respectively on these two pre-reading skills. These are strong, positive findings of this assessment.

Listening comprehension scores were also high, indicating generally good comprehension by Grade 2 students of story text written in formal Arabic. These Grade 2 students are clearly accustomed to hearing classroom instructions and narrative text in formal Arabic.

But zero scores were higher and average reading scores lower in the oral reading fluency and reading comprehension subtasks than might be expected from the generally good student performance in pre-reading and listening comprehension skills. The average (mean) score of sample Grade 2 students on the standard EGRA

An Illustrative Benchmark for Listening Comprehension

Unlike the two optional subtasks of familiar words reading and reading comprehension for a text read for 90 seconds without diacritics, the results of the listening comprehension subtask are central to EGRA analysis and understanding student performance in reading comprehension. For these reasons, this Report applies the same benchmark of 50% of questions answered correctly that is proposed for the reading comprehension subtask. This benchmark for listening comprehension is, however, only *illustrative*, for the purposes of presenting results in this Report, and is not proposed for official adoption. It is important that this be clear.

⁸ All average scores are means that include zero scores in their calculation.

subtask of oral reading fluency for 60 seconds of a story of *familiar* words with diacritics was just 16.7 correct words per minute (cwpm). This result was lower than their performance in reading familiar words without context; and not much better than their success in reading unfamiliar (invented) words. Nearly one-quarter (22%) of Grade 2 students could not correctly read a single word on the first line of this story and only 16% of students read at the proposed benchmark level of at least 30 correct words per minute.

Their diminished performance in oral reading fluency is the direct cause of their low performance in reading comprehension. Too few Grade 2 students are able to read connected text in a narrative passage with sufficient fluency to enable them to comprehend that text. The average score on the standard reading comprehension

Table 1: Summary of EGRA Average Scores for Grade 2

Subtask	Percentage of Grade 2 students with zero scores	Grade 2 average score	Proposed provisional benchmark for Grade 2	Percentage of students performing at or above benchmark
Letter sound identification (clspm)	6.8%	36.0	45	41%
Familiar Word reading (cwpm)	10.0%	17.5	na ⁹	na
Nonword reading (cnonwpm)	8.7%	13.9	20	29%
Oral reading fluency (ORF) – with full diacritics (cwpm)	22.1%	16.7	30	16%
Oral reading fluency (ORF) – without diacritics (cwpm)	10.9%	24.9	35	27%
Reading comprehension – ORF with diacritics (max. 6)	35.7%	1.6	3.0 (50% correct)	26%
Reading comprehension – ORF without diacritics (max. 6)	25.9%	2.2	Na	na
Listening comprehension (max. 6)	7.1%	3.3	3.0 (50% correct)	68%

Note: clspm = correct letter sounds per minute; cnonwpm = correct nonwords per minute; cwpm = correct words per minute. The percentage zero scores for the two Reading Comprehension subtasks include “early stop” students on the respective Oral Reading Fluency exercises (22.1% and 10.9% respectively) who were subsequently not asked any comprehension question on the passage that they “early stopped”.

subtask after one minute of story reading (with diacritics) was just 1.6 questions answered correctly. Just one-quarter (26%) of sample students could answer at least 3 questions correctly – the proposed benchmark for Grade 2 reading comprehension of 50% of questions answered correctly. At the bottom end of reading comprehension proficiency, a much larger percentage – nearly 36% – of Grade 2 students could not correctly answer a single comprehension question.

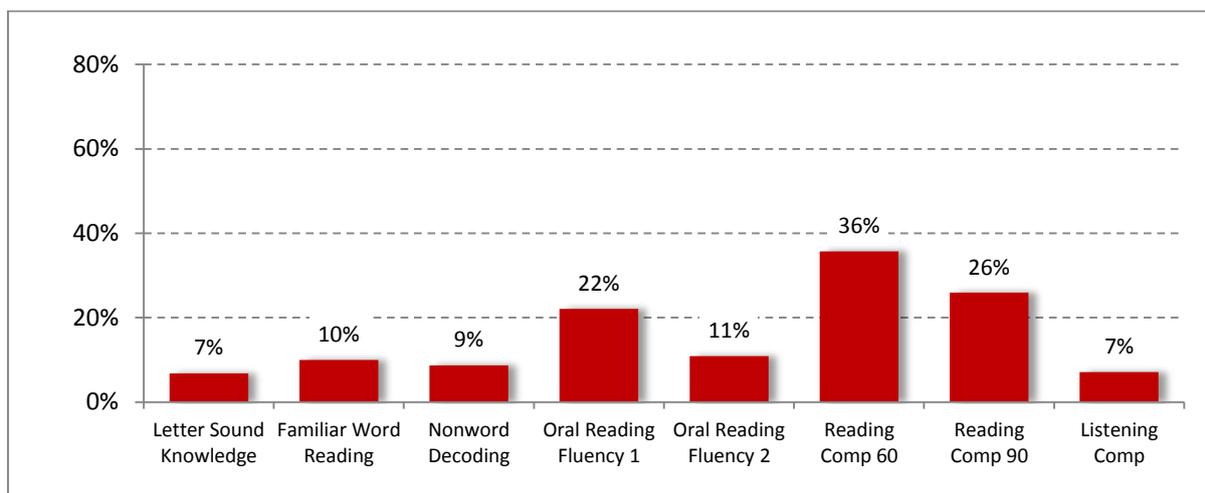
⁹ See footnote 1.

Nonreaders: In addition to considering the percentage of students who were unable to complete a single item on individual subtasks, there was a subgroup of students who were **nonreaders**. Nonreaders were students who scored zero on all three of the letter-sound identification, nonword reading, and oral reading fluency (with diacritics) subtasks. These students could not read correctly a single word nor correctly identify a single letter sound on the first line of each test. The results of the West Bank EGRA on this overall measure of reading performance were also very good. Less than 5% of sampled Grade 2 students were **nonreaders** or truly illiterate. There was, however, a sharp difference between genders on this measure. The percentage of boys who were nonreaders – 6.5% – was more than double the percentage of girls (3.1%). This pattern of girl outperformance in early grade reading has also been observed in many other countries and languages, and in all previous Arabic EGRAs. The gender gap is, however, greater in this sample of West Bank students from Grade 2 than observed in Egypt or Jordan.

The percentages of zero scores on the EGRA subtasks are shown in **Figure 1**. The low percentages in each the pre-reading skills of letter sounds knowledge and nonwords reading – in addition to familiar words reading – as well as listening comprehension are all positive. By comparison, the higher shares of zero scores in oral reading fluency (except oral reading fluency without diacritics, for which a lower standard of scoring is applied) and reading comprehension stand out.

The Grade 2 benchmarks proposed for the different reading skills include 2018 benchmarks for reductions in the percentage of zero scores on each subtask. These too are important benchmarks. These **zero score benchmarks** are identified for each subtask in the Subtask Analyses section below.

Figure 1: Percentages of Grade 2 Students with Zero Scores on the EGRA Subtasks

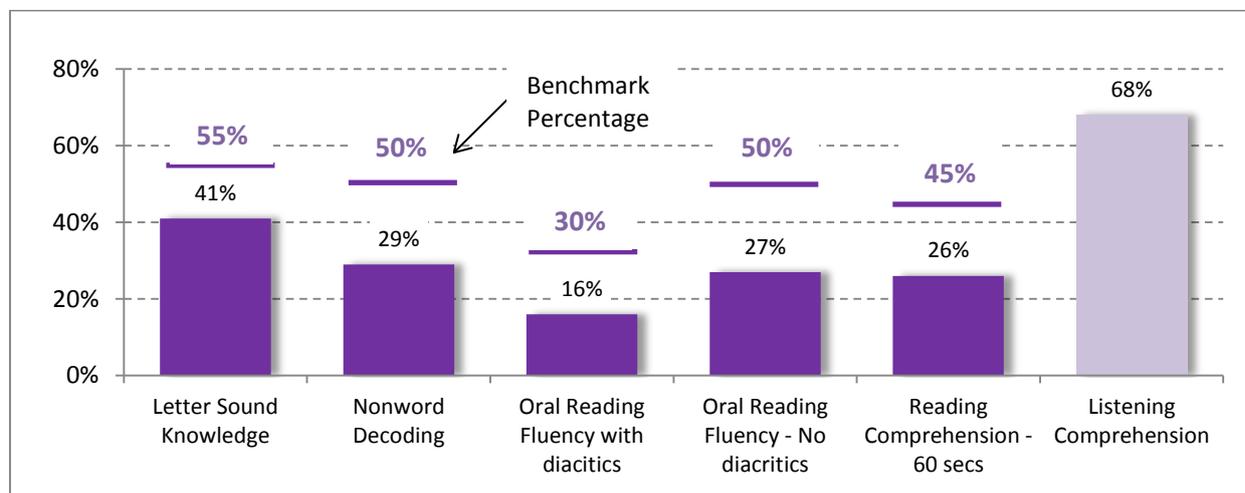


Students Reading at Benchmark Level: It is also important to consider how many students are now performing well. **Figure 2** shows sizable percentages of Grade 2 students already reading at or above the average score benchmarks proposed for the respective subtasks to be achieved in 2018. With the sole exception of oral reading

fluency with diacritics, the percentages of Grade 2 students reading at grade level on each of the EGRA subtasks is strong. The tall column for listening comprehension is colored separately to remind the reader that this result is only *illustrative* and no benchmark was developed at the benchmarking workshop for this subtask.

Shown above each column of current benchmark achievement in Figure 2 is a line and a percentage. This is the **benchmark percentage** of Grade 2 students that should read at benchmark level by 2018 on this EGRA subtask. For example, in this EGRA baseline, 41% of Grade 2 students read at or above the above score baseline of 45 correct letter sounds per minute. The 2018 benchmark *percentage* is that **55%** of Grade 2 students will read at least 45 correct letter sounds per minute. The benchmark percentages for each subtask are presented in the Subtask Analyses section below.

Figure 2: Percentages of Grade 2 Students reading at or above Average Score Benchmarks on the EGRA Subtasks



Accuracy of Grade 2 Reading: To this point, our analyses first looked at average scores on subtasks; second, at the percentage of Grade 2 students with zero scores on subtasks; third, at the percentage of students reading at or above the proposed benchmarks for each subtask. A fourth way to analyze reading proficiency is to compare the correct items or answers to the number of items *attempted* on each subtask. “Percentage correct of attempted” is an important metric of reading accuracy. Comparing scores to the number of items attempted on the subtask provided valuable insight into students’ mastery of early reading skills. Attempted scores are always higher than total scores.

Table 2 presents the average number of items attempted for each subtask and the average percentage of correct attempts. Children were most successful in letter identification and oral reading fluency without diacritics, an optional subtask with a lower standard for correct performance. Percentage correct of attempted scores on both subtasks exceeded 70%: Grade 2 children were correctly answering at least 70% if each item attempted in these subtasks. Student performance in nonword reading was also strong, with a percentage correct score (60.7%) only slightly below that for

familiar words (63.3%). Not surprisingly, percentage correct scores were lowest in the oral reading fluency with diacritics (53.1%) and both reading comprehension subtasks (less than 45%). Even in listening comprehension, the comprehension subtask on which Grade 2 students generally performed better, there is room for improvement. Student accuracy in answering the listening comprehension questions was not high: on average, just 55% of questions were answered correctly. Comprehension skills can be taught, learned and mastered.

Taken together, these results indicate that most Grade 2 students have a good foundation of pre-reading skills but are struggling to read familiar words with diacritics in passage text and to comprehend what they read.

Table 2: Summary of EGRA Scores for the Number of Items Attempted

Subtask	Average number attempted	Percent correct of attempted
Letter sound identification (clspm)	37	74.6%
Familiar word reading (cwpm)	23	63.3%
Nonword reading (cnonwpm)	20	60.7%
Oral reading fluency <u>with</u> diacritics (cwpm)	24	53.1%
Oral reading fluency with no diacritics (cwpm)	37	70.4%
Reading comprehension – 60 seconds (max. 6)	2.7	43.8%
Reading comprehension – 90 seconds (max. 6)	4.1	44.8%
Listening comprehension (max. 6)	6.0	54.6%

Note: clpm = correct letters sounds per minute; cnonwpm = correct nonwords per minute; cwpm = correct words per minute.

Subtask Analysis

In this section, we look at each subtask separately and analyze the range of student performance on each.

Letter Sound Knowledge

Letter sound knowledge is a basic reading skill taught in Grade 1. Letter sound knowledge, or the alphabetic principle, is considered a prerequisite skill for beginning reading and has been found to be a strong predictor of reading growth in consonant-based alphabets, such as Arabic. The test of letter sound knowledge was included to appraise the skill levels of these Grade 2 students on this basic reading skill. Each student was presented with a chart of 80 random letters in initial word or independent form with diacritical marks. They were asked to pronounce the sounds associated

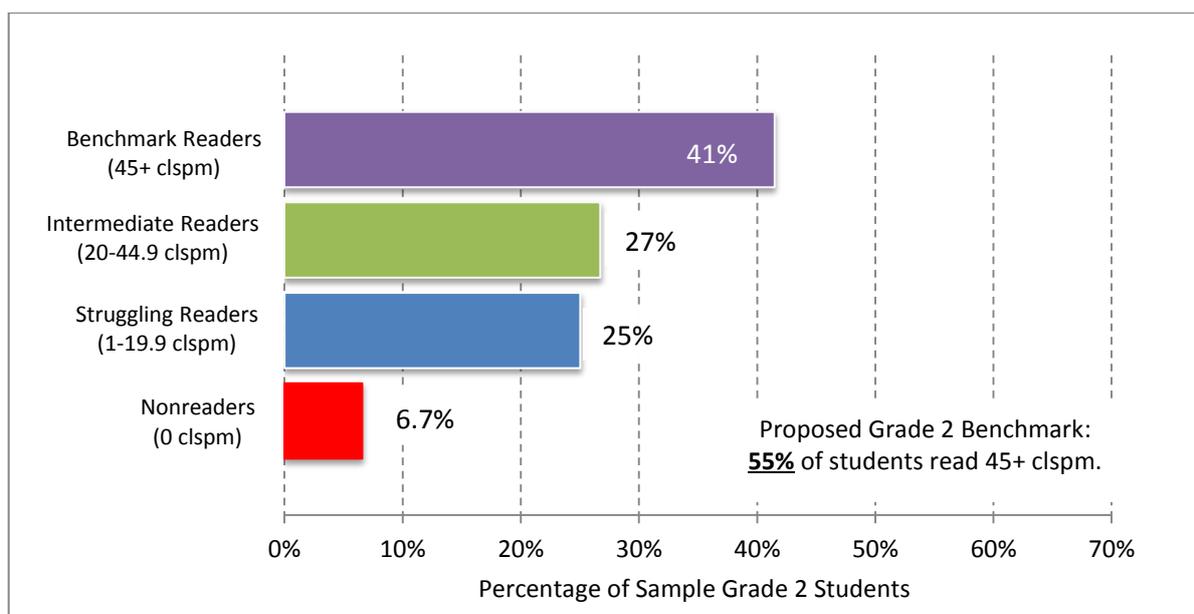
with as many of these letters as they could within one minute. Scores for this subtask were the number of letter sounds the student could correctly pronounce within one minute (correct letters sounds per minute [clpm]).

The 2018 benchmarks for letter sound knowledge proposed by participants of the September 2014 benchmarking workshop with the MOEHE are:

- a. Grade 2 children correctly read at least 45 letter sounds in one minute;
- b. 55% of Grade 2 children read letter sounds at benchmark level;
- c. Just 3% of Grade 2 students cannot correctly read a single letter sound from the first line of 10 letter sounds for this subtask (zero scores).

Figure 3 presents the results of Grade 2 student performance on letter sound knowledge. Overall, 93% of tested students could identify at least one letter correctly. Less than 7% could not correctly read any of the first ten letter sounds, which halted this subtask. This is an excellent result.

Figure 3: Distribution of Grade 2 Students on the Letter Sound Identification Subtask



One quarter of Grade 2 students (25%) are, however, still struggling with letter sounds. They could produce only 1 to 19.99 correct letter sounds in one minute. However, a much larger percentage of students (40%) are strongly proficient in this reading subtask, performing at or above the proposed benchmark level of 45 correct letter sounds per minute. This is both a sizable benchmark and a strong result. Between these high and low performers is another one-third (35%) of “intermediate” readers and those approaching benchmark level. These students can correctly read 20-44.99 letter sounds in one minute.

These are strong results: the lowest percentage of zero scores of any Arabic EGRA for Grades 2 or 3 and the highest percentage of students reading at the proposed

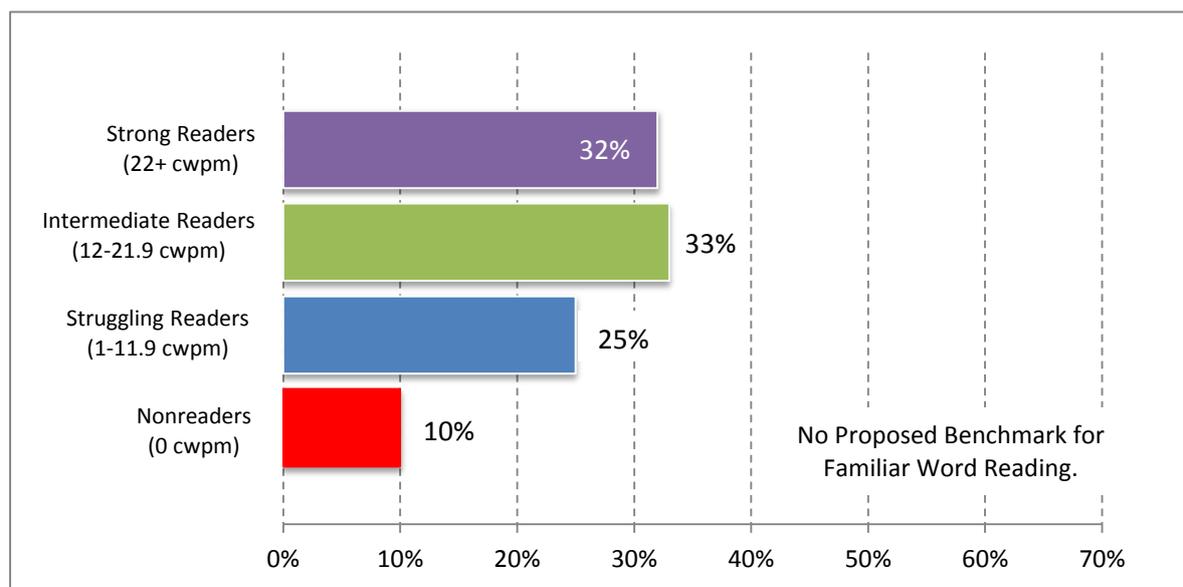
benchmark level of 45 correct letter sounds per minute. Reading letter sounds at a success rate of 75% correct of attempted (*see* Table 2), Grade 2 students are close to mastering this basic pre-reading skill and achieving all three benchmarks for letter sounds knowledge proposed in 2018. This is a solid foundation for positive achievement in strengthening higher-level reading skills.

Familiar Word Reading

At MOEHE request, the EGRA included a test of familiar word reading as the second subtask. Students read a chart of 50 words with diacritics on all but the last syllable. All words were taken from the Grade 2 textbooks for Arabic (both terms) so that, in principle, all students had previously seen the words.

Figure 4 presents the range of student performance in reading familiar words. One third of Grade 2 students were strong readers, reading at least 22 words correctly in one minute. Another one-third was intermediate readers, correctly reading 12 to 21.99 familiar words in one minute. The percentage of students – one-quarter – who

Figure 4: Distribution of Grade 2 Students on the Familiar Words Subtask



struggled with familiar words was identical to the percentage (25%) of students who struggled with letter sounds. They were the same students. Just 10% of Grade 2 students could not correctly read a single familiar word from the 5 words on the first line of this subtask. On average, these Grade 2 students had a success rate of 63% correctly read of attempted words. These too are positive results.

The familiar words subtask is, however, an uncertain measure of *reading* skill. Because these are familiar words that students were presumably taught in Grade 2, their performance on this subtask might reflect their memorization of specific vocabulary. In this case, the subtask might not be a true measure of student reading.

For this reason, no benchmark is proposed for familiar word reading. It is a better test of student knowledge of their Grade 2 textbook vocabulary than a test of student *skill* in decoding and correctly reading words. The familiar words subtask is not an essential subtask in the battery of EGRA skills test. The subtask of nonwords reading is a more certain test of student skill in word decoding.

Nonword / Invented Word Reading

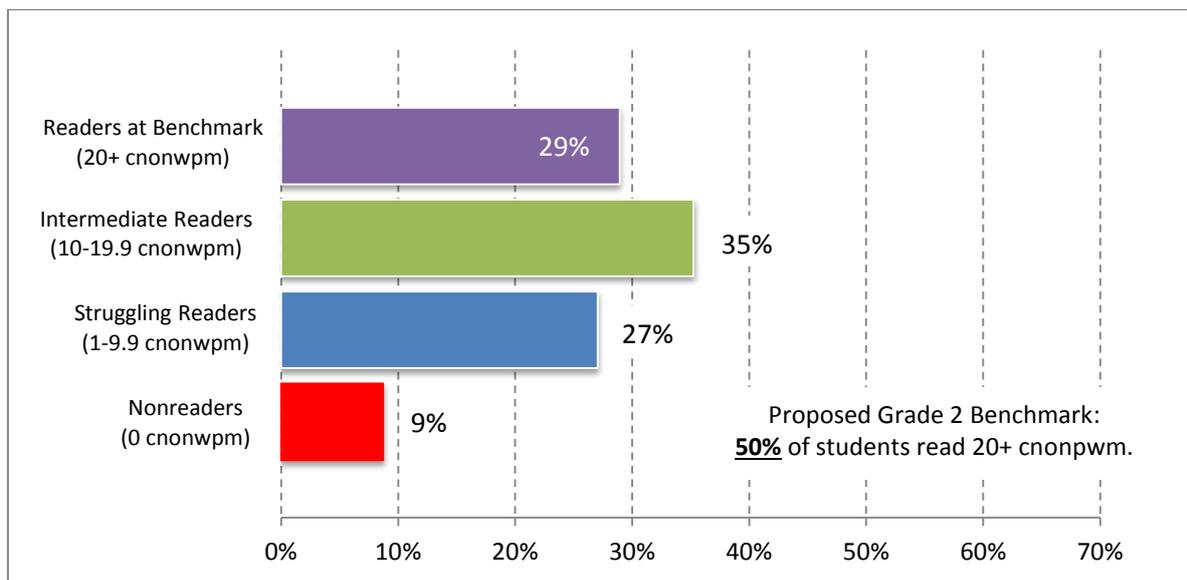
In the nonword subtask, students were presented with a chart containing 50 invented (or nonsense) words—most of them 3 or 4 letters—with diacritics on all but the final syllable and were asked to pronounce as many of the words as they possibly could within one minute. All 50 nonwords were constructed consistent with Arabic orthography for word construction. Skill in reading nonwords is a purer measure of decoding than using real words, because children cannot recognize the words by sight. Decoding is considered a self-teaching skill that enables children to read new and unfamiliar words independently. Nonword reading for one minute is a standard EGRA subtask.

Scores for this subtask were the number of words the student could correctly read within one minute (correct nonwords per minute or “cnonwpm”). The proposed 2018 benchmarks for this higher pre-reading skill in Grade 2 are:

- a. Grade 2 children correctly read at least 20 nonwords in one minute;
- b. 50% of Grade 2 children read nonwords at benchmark level;
- c. Just 5% of Grade 2 students cannot correctly read a single nonword from the first line of 5 nonwords in this subtask (zero scores).

The results presented in **Figure 5** show strong performance by Grade 2 students in decoding nonsense words. Indeed, the results are comparable to the previous familiar

Figure 5: Distribution of Grade 2 Students on the Nonword / Invented Word Reading Subtask



words subtask. Only 9% of Grade 2 students could not correctly read a single nonsense word on the first line; 29% are now reading at or above the benchmark level of 20 correct nonwords per minute. Another 35% of students are intermediate readers approaching benchmark proficiency, correctly reading 10 to 19.99 nonwords per minute.

These results are little different from the results of familiar word reading. The percentage correct of attempted (*see* Table 2) was also nearly the same: an average of 60.7% of nonwords attempted was read correctly by these Grade 2 students. This too is a strong achievement. Grade 2 students in MOEHE schools are well on their way to achieving the 2018 benchmark for this higher pre-reading skill that is best mastered in Grade 2. They are proficient in decoding words.

Oral Reading Fluency – with and without diacritics

Two oral reading subtasks were included in this EGRA. The first was the standard EGRA subtask of 60-seconds reading a story of local relevance with diacritics. The second subtask was an alternative design: 90-seconds reading a different story (also of local relevance and similar word difficulty) without diacritics. Despite these different durations, however, the same metric of oral reading fluency was applied to both subtasks: the number of correct words read *per minute*. Using the same metric is essential to compare results.

Oral reading fluency is a core index of reading competence, as it measures the skill and speed with which children translate letters into sounds, decode unfamiliar words, recognize known words, and simultaneously make sense of the text's meaning. Weakness in any one of these processes can slow or disrupt children's reading fluency.

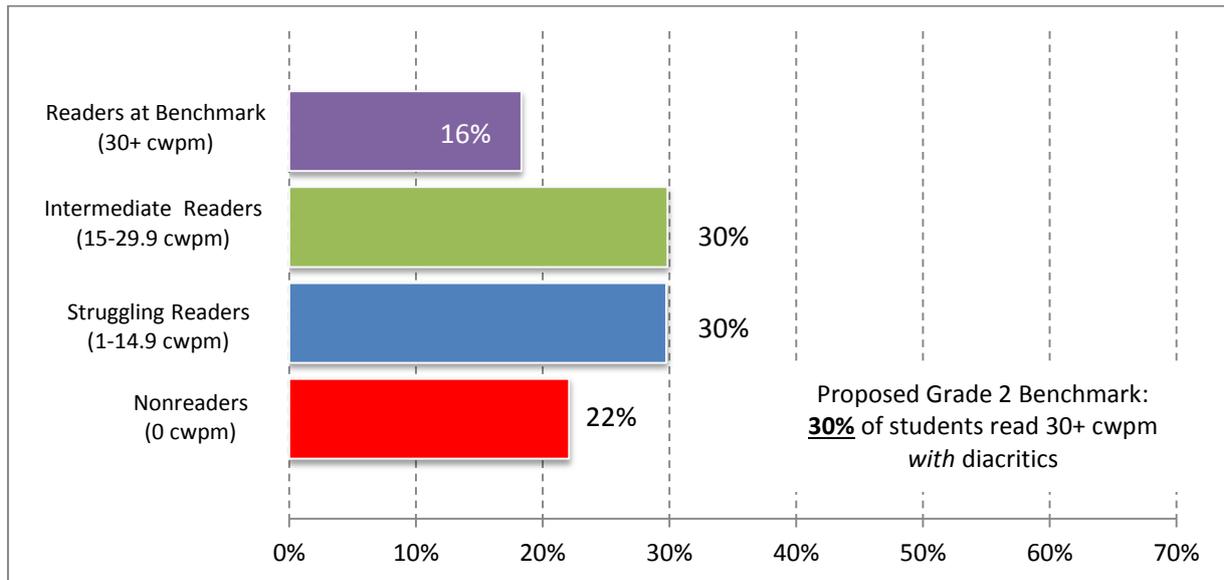
The proposed 2018 benchmarks for Grade 2 students in oral reading fluency *with* diacritics are:

- a. Grade 2 children correctly read at least 30 words per minute of a narrative passage;
- b. 30% of Grade 2 children read at benchmark level;
- c. Only 10% of Grade 2 students cannot read a single word correctly (with diacritics) on the first line of the reading passage.

Figure 6 shows that 22% of the students in Grade 2 could not correctly read a single word of the first eight words of the passage. This is a large percentage of nonreaders of known and familiar words. Just 16% of sampled students were reading at or above the benchmark of 30+ correct words per minute, with equal and larger shares of struggling and intermediate readers. Overall, these results indicate a low level of reading fluency. Grade 2 children are struggling to read connected text.

This result is also indicated by the EGRA metric of percentage correct of attempted (*see* Table 2). On average, these students correctly read just half of the words they attempted: 53.1%. This is a lower percentage with familiar than their comparable

Figure 6: Distribution of Grade 2 Students on the Oral Reading Fluency Subtask – with diacritics



average score when reading unfamiliar words (nonwords / invented words): 60.7% of attempted words read correctly. These Grade 2 children have more difficulty reading familiar words in a passage than unfamiliar words in isolation. This finding recommends greater reading practice of connected text (passage reading) in class.

Passage Reading without Diacritics: When the children read the 2nd reading passage without diacritics, there was a large increase in the total number of words read: on average, 37.9 words without diacritics compared to 16.7 words read correctly *with* diacritics. But it is uncertain whether the words without diacritics were read “correctly”. It is only certain that these children correctly read the component letters of these words. This is a lower standard for reading proficiency than reading with diacritics. It is, however, a useful measure of the progress of student reading proficiency in the earliest grades (Grades 1-2).

The benchmarking workshop proposed the following benchmarks for passage reading without diacritics to be achieved by Grade 2 students in 2018:

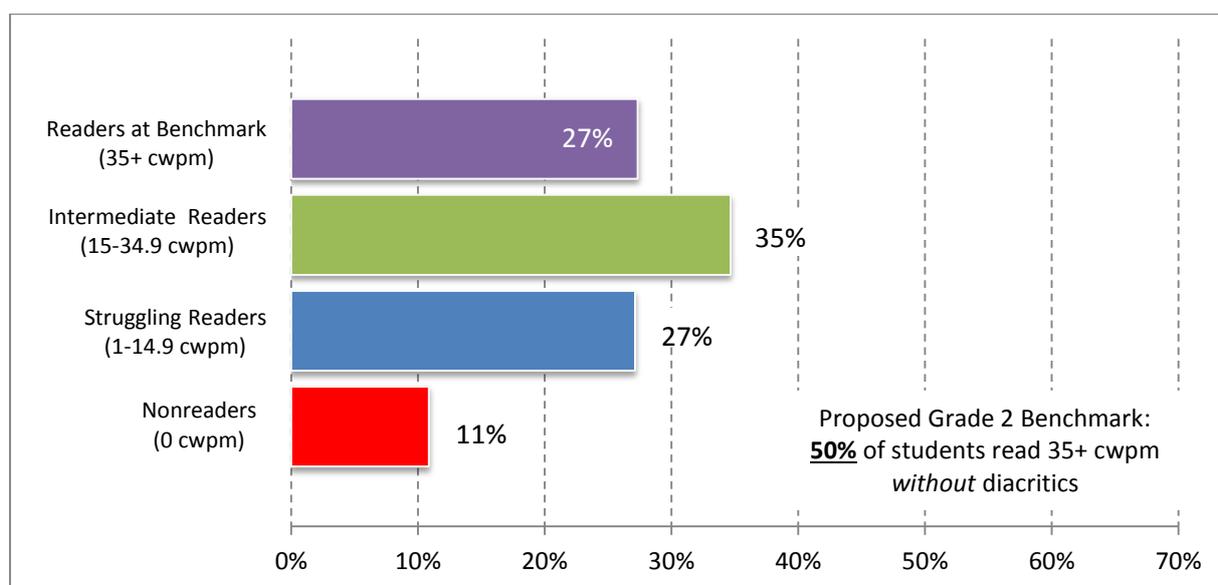
- a. Grade 2 children correctly read at least 35 words per minute without diacritics;
- b. 50% of Grade 2 children read at benchmark level;
- c. Just 5% of Grade 2 students cannot read a single word correctly (*without* diacritics) on the first line of the reading passage.

Figure 7 presents the range of student proficiency in reading words of connected text in a story passage without diacritics. Just 11% of students could not correctly read any of the 7 words in the first line of this passage.¹⁰ More than one-quarter (27%) of tested students read at the benchmark level of 35 or more correct words without

¹⁰ The passage included two very familiar words of just two letters on the first line: “أم” and “في”.

diacritics. And another one-third of students (35%) performed this subtask at an intermediate level of proficiency, correctly reading 15 to 34.99 words without diacritics. On average, the students were quite successful on this subtask, correctly reading over 70% of the words without diacritics that they attempted to read. This was the second highest average of reading accuracy across all subtasks (*see* Table 2).

Figure 7: Distribution of Grade 2 Students on the Oral Reading Fluency Subtask – without diacritics



Reading Comprehension – 60 seconds with diacritics

Immediately after children had read the short reading passage (57 words) with full diacritics for one minute, *all those who were able to read at least one word correctly on the first line* were asked questions about the story. Students who scored zero on oral reading frequency were not asked questions. The number of questions asked of each student depended on how far in the story each child had read. Children were asked only those questions that could be answered from the text that they read.¹¹ All six questions in this subtask were literal and could be answered directly from information provided in the story. There were no inferential questions that required students to combine information from the story with their background knowledge to derive a correct answer.

Children’s reading comprehension scores were recorded as the number of correct responses. The proposed 2018 benchmarks for Grade 2 reading comprehension after one minute reading of a story with diacritics are:

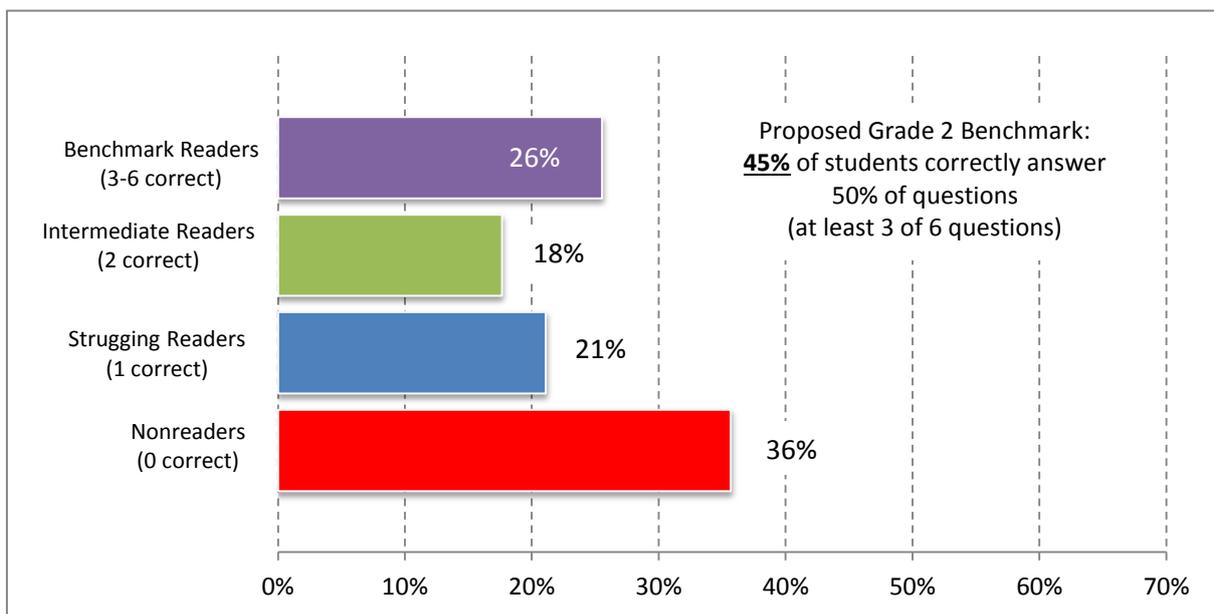
- a. Grade 2 children correctly answer 50% of the reading comprehension questions (3 of 6 questions);
- b. 45% of Grade 2 children are reading at benchmark level;

¹¹ The iPads would only “activate” those questions that could be answered from the text preceding the last word of the story read by the student in one minute, as recorded by the assessor.

- c. Only 15% of Grade 2 children cannot correctly answer any question (zero scores).

Overall, children had weak reading comprehension scores (*Figure 8*). More than one-third (36%) of students could not correctly answer a single question and another 21% answered just 1 question correctly. These are high percentages of children not comprehending what they read. Just one-quarter of children (26%) were performing at the proposed Grade 2 benchmark level of 3 or more correct answers. The 2018 benchmark is 45% of children correctly answering 3 or more reading comprehension questions after one minute reading.

Figure 8: Distribution of Grade 2 Students on the Reading Comprehension Subtask – 60 seconds with diacritics



Total shares exceed 100% due to rounding.

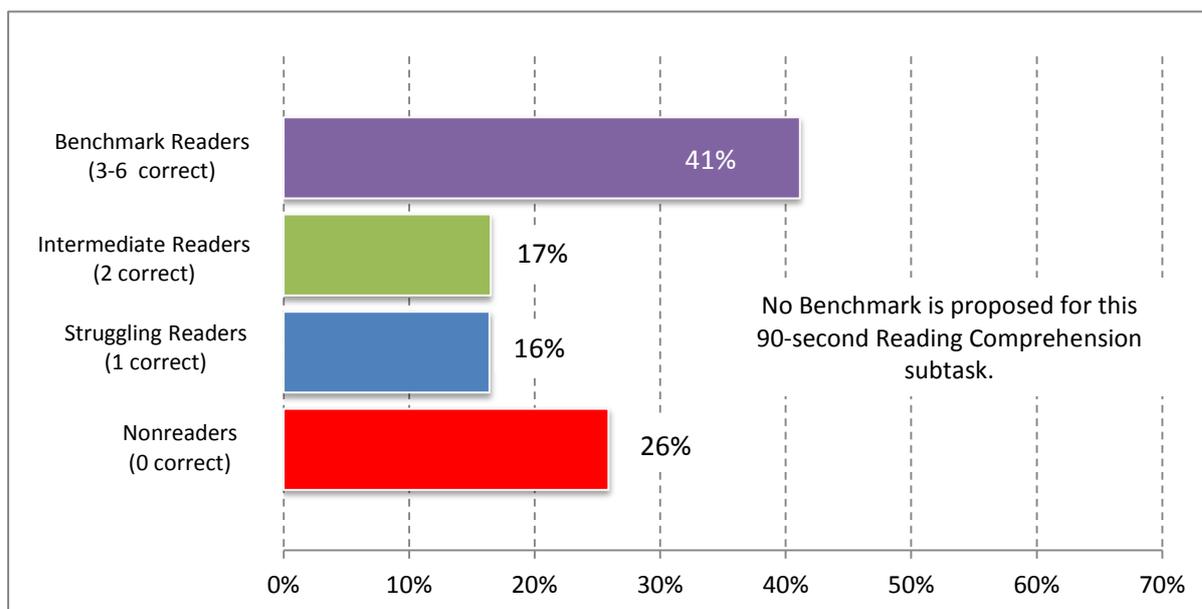
The weakness of reading comprehension skills is also indicated by the low percentage correct of attempted. The sample students correctly answered less than half (43.8%) of the questions that they attempted (*see* Table 2). Having to concentrate hard on the mechanics of reading even familiar words, the majority of Grade 2 students were unable to think about and recall what they are reading. The subtask results show this clearly.

Reading Comprehension – 90 seconds without diacritics

The reading comprehension subtask of 60 seconds with full diacritics is the standard for EGRAs in Modern Standard Arabic. The addition of a 2nd oral reading passage of 90 seconds without diacritics in this Grade 2 EGRA provided an excellent opportunity to compare reading comprehension results when two key variables are changed: i) the time allowed for reading, and ii) diacritics are omitted from the text.

The results of this 2nd reading comprehension exercise are shown in *Figure 9*. Comparing the results here with Figure 8 reveals shows *somewhat better* results when children are given more time to read. The percentage of children who could correctly answer at least 3 questions (41%) was larger; and the percentage of students who could not answer any question correctly – zero score – were fewer (26%) when given 90 seconds to read the story. The comparable results for students reading just 60 seconds were 26% at benchmark and 36% with zero scores.

Figure 9: Distribution of Grade 2 Students on the Second Reading Comprehension Subtask -- 90 seconds without diacritics



But the key observation here is that these results are only *somewhat better*. With 50% more time to read the story (90 instead of 60 seconds), the percentage of students reading at benchmark was only 15 percentage points higher and the percentage of students unable to answer *any* question correctly – zero scores – was still a high 26%.

The more revealing analysis, however, is the percentage correct of attempted (*see* Table 2). The result on this EGRA measure was no better than the result when children were given just 60 seconds to read. Given more time, children were reading farther in the story and thus attempting more questions. The average number of attempted questions on the 90-second reading passage was 4.1 questions – versus 2.7 questions attempted as the average for students reading just 60 seconds. *But students with more time were not more successful in answering the questions:* just 44.8% of attempted questions were answered correctly when children read 90 seconds, versus 43.8% when students read just 60 seconds. Greater time to read did not improve this measure of successful comprehension.

This last result also suggests that diacritics neither significantly impede nor improve comprehension. The 90-second reading passage did not include diacritics. Not having to concentrate on reading the diacritics correctly, children were able to read more words. But they were neither more nor less successful in answering questions

correctly about what they read than when diacritics were shown. Diacritics are essential to know the meaning of a word in isolation. But the context of a story will typically indicate the correct meaning of a word when no diacritics are provided.

Overall, the results of these two reading comprehension exercises are not markedly different. The results confirm that reading comprehension is not greatly improved by allowing children more time to read. Children must progress beyond mechanical reading and learn to think about the meaning of what they read. Reading comprehension is a parallel process of reading cognition that takes place at the same time as the mechanical reading of words. Only when children can free more of their attention from mechanical reading – that is, when their word recognition skills need less of their conscious control – can they begin to focus more on what they are reading. Reading fluency is an essential pre-condition for reading comprehension. But fluency alone is not sufficient. Comprehension is a specific reading skill that also must be mastered.

Numerous large-scale studies and meta-analyses have reported robust correlations between oral reading fluency and reading comprehension.¹² In other alphabetic languages, the relationship between decoding speed and reading comprehension is particularly strong among beginning readers because their word recognition skills still require conscious control.¹³ This was supported by the strong correlations ($r^2 = 0.83$ and $r^2 = 0.80$) between students' scores in oral reading fluency and reading comprehension on the 2 versions of this subtask: 60 seconds and 90 seconds respectively. These findings confirm that fluent oral reading is a critical component for reading comprehension, but that it is not sufficient for reading comprehension.

Listening Comprehension

In this final subtask of the reading assessment, the assessor read a short narrative story (58 words) to the child, followed by six questions about that story. Both the story and the questions were read in formal Arabic. This was purely a listening subtask: the child was not given a copy of the story to follow along or refer to when answering the questions. All students, including the nonreaders who did not correctly read a single item in one or more of the preceding subtasks, were tested for listening comprehension.

Although the listening comprehension subtask typically assesses a range of language and skills, such as attention, vocabulary knowledge, comprehension strategies, processing of oral language, and generation of appropriate replies, for Arab children, it also assessed their proficiency in the formal dialect of Arabic. Modern Standard

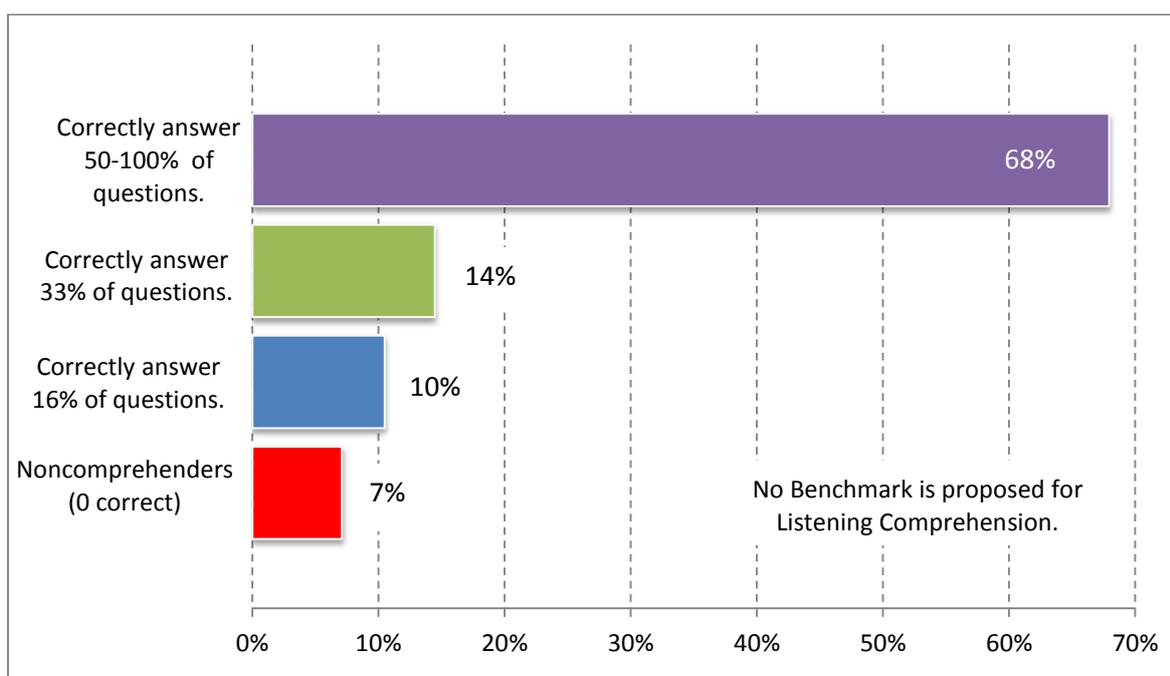
¹² See Abu-Rabia (2007); and also: M.C. Daane, J.R. Campbell, W.S. Grigg, M. J. Goodman, & A. Oranje. (2005). *Fourth-grade students reading aloud: NAEP 2002 special study of oral reading (NCES 2006-469)*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. Washington, DC: Government Printing Office. G.S. Pinnell, J.J. Pikulski, K.K. Wixson, J.R. Campbell, P.B. Gough, & A.S. Beatt. (1995). *Listening to children read aloud: Data from NAEP's Integrated Reading Performance Record (IRPR) at grade 4 (NCES 95-726)*. Washington, DC: U. S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

¹³ W.A. Hoover & P.B. Gough. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127–160.

Arabic differs substantially from the vernacular dialect used in children’s homes. Comparing children’s comprehension in these two modalities is important, because it allows determination of whether poor reading comprehension can be attributed to limited reading skills or to more general difficulties in comprehending the formal Arabic dialect used in schools.

In general, the students performed quite well on the listening comprehension subtask (**Figure 10**). Their scores on this subtask were much stronger than their reading comprehension scores. Yet the different reading passages were of similar length (57-58 words) and word difficulty. Just 7% of the children were unable to correctly answer any listening comprehension questions. Nearly the same percentage of children (7.6%) answered all 6 questions correctly. And more than two-thirds of all students (49.5%) could correctly answer 3 or more questions.

Figure 10: Distribution of Grade 2 Students on the Listening Comprehension Subtask



Total shares less than 100% due to rounding.

This result makes clear that the lower result in reading comprehension is chiefly due to children’s difficulty in *reading fluency* and not their difficulty in understanding formal Arabic. When listening to a story read to them in formal Arabic, most students understood the story quite well. The large majority of Grade 2 children in MOEHE schools in the West Bank do not have great difficulty comprehending a simple story in formal Arabic.

This conclusion is also supported by the low correlation between reading comprehension and listening comprehension for this sample of Grade 2 students. Whereas oral reading fluency was strongly correlated with reading comprehension

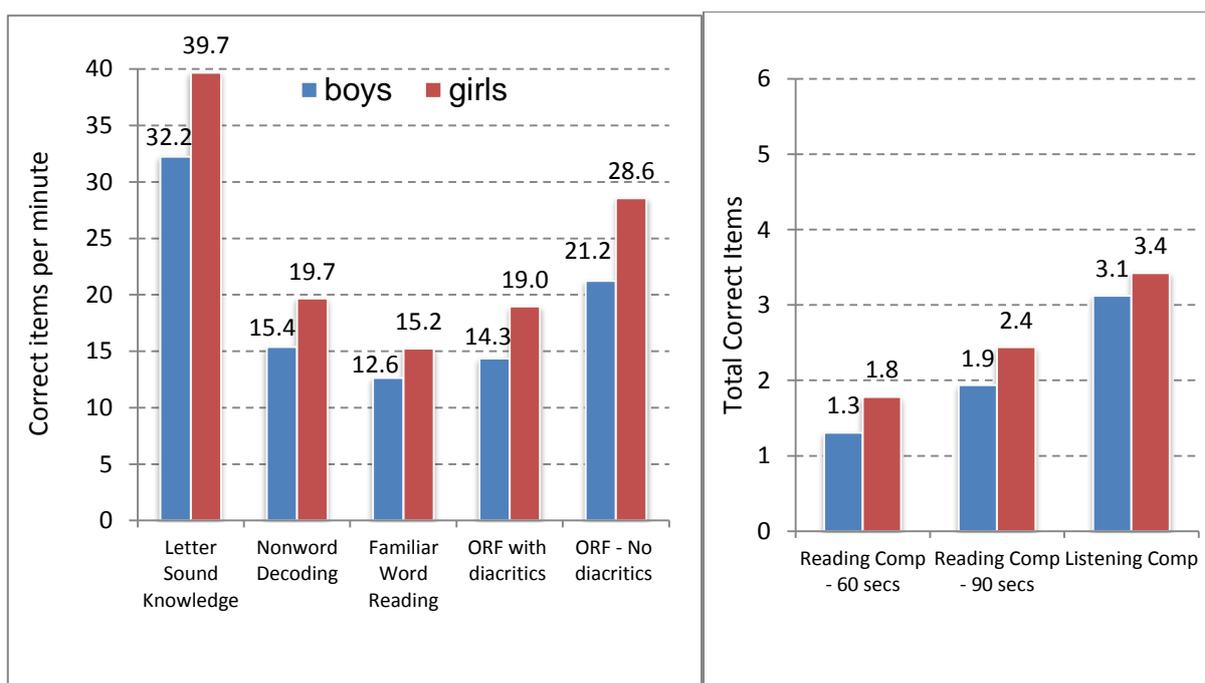
($r^2 = 0.83$), listening comprehension's correlation with reading comprehension was much weaker ($r^2 = 0.43$).

This good result on listening comprehension can – and should – be improved. On average, Grade 2 students were successful in answering only half (54.6%) of the listening comprehension questions (*see* Table 2). Listening comprehension is also an acquired skill. And while the sample Grade 2 students showed good results in listening comprehension, they have not yet mastered this skill.

Gender Differences in Reading Performance

The findings point to significant gender differences in Grade 2 reading proficiency in all subtasks, and marked differences in some (**Figure 11**). The gender gap between boys and girls in Grade 2 reading skills is especially pronounced in oral reading fluency. The differences are less, however, in the comprehension subtasks. Grade 2 girls markedly outperform boys in reading accuracy and speed. But they do not comprehend what they read or hear greatly better than boys. We have also previously

Figure 11: Reading Proficiency of Grade 2 Boys and Girls



Note: Separate scales were used for the two parts of Figure 11. The graph on the left shows student performance on the timed tasks and uses items per minute as the unit of measurement. The graph on the right shows student performance on the 3 untimed comprehension subtasks.

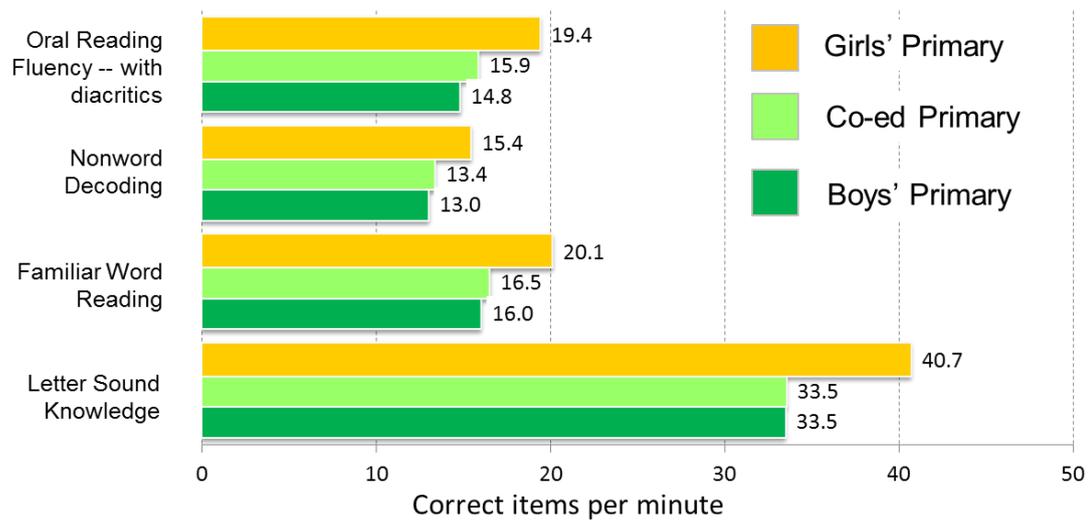
noted (page 11) the significantly lower percentage of Grade 2 girls who are nonreaders – less than half the percentage of boys. These West Bank results are consistent with the findings of previous Arabic EGRAs in Grades 2 and 3 in Jordan and Egypt. But the gender gap is more pronounced in the West Bank.

School Gender Differences in Reading Performance

At ministry request, the sample of 150 MOEHE primary schools was stratified by school gender with 50 boys' schools, 50 girls' schools and 50 mixed (co-ed) schools randomly selected for testing. The gender gap in reading proficiency, presented above, presages select results by school gender. Other results are, however, surprising.

Figure 12 presents the results for selected subtasks by school gender. Not surprisingly, Grade 2 students in girls' schools outperformed their peers in both boys' and mixed primary schools on all subtasks. Mixed schools, however, did insignificantly better than boys' schools on most subtasks.

Figure 12: Selected Average Subtask Scores by School Gender



The surprising – and yet unexplained – finding of analysis by school gender is that both boys and girls in *mixed schools* lag the reading performance of their gender peers respectively in girls' schools and boy's schools (**Table 3**). In mixed schools,

Table 3: Average Subtask Scores by Gender in Mixed and Single-Gender Schools

EGRA Subtask	Boys' School Grade 2	Boys in Mixed Schools, Grade 2	Girls' School Grade 2	Girls in Mixed Schools, Grade 2
Correct Letter Sounds per minute	33.4	29.3	40.6	37.5
Correct Familiar Words per minute	16.0	13.9	20.1	18.6
Correct Nonwords per minute	13.0	11.8	15.4	14.8
Oral Reading Fluency – 60 seconds with diacritics (cwpm)	14.8	13.3	19.4	17.9

both genders had slightly higher percentages of zero scores than in single-gender schools. Girls in mixed schools still outperformed their male classmates in Grade 2 reading. But both genders did less well in reading than their peers in single-gender schools.

The explanation may be sociological. Mixed primary schools are typically newer schools constructed on the edges of expanding towns and villages. The population served by these schools may be new families whose parents are less schooled. Their teachers may also be newer, less experienced and less trained. This finding, however, merits educational research to determine if specific interventions or supplemental attention might be recommended for mixed primary schools to enhance the early reading outcomes of their students.

Significance of Class Size for Reading Results

The MOEHE was keen to know if class size might be a significant variable in the reading outcomes of children. West Bank schools vary significantly in the average size of primary classes, with small classes in some small rural communities and large classes in some urban and large village schools. In this random stratified sample of 150 MOEHE primary schools with Grade 2, the smallest class had 11 Grade 2 students and the largest had 45 students. Five schools had classes of 40-45 students. Nine schools had classes of 20 students or less.

Table 4 presents the average scores for each reading subtask for three categories of class size: *small classes* of 11-20 students per class, *medium-size classes* of 21-30 students per class, and *large classes* of 31 or more students. The findings show

Table 4: Average Subtask Scores by Class Size

Subtask	Small Classes (11-20 students)	Medium Classes (21-30 students)	Large Classes (31+ students)
Letter sound identification (clspm)	34.91	36.73	35.35
Familiar word reading (cwpm)	16.96	17.80	17.35
Nonword reading (cnonwpm)	13.84	14.04	13.85
Oral reading fluency <u>with</u> diacritics (cwpm)	16.23	16.88	16.51
Oral reading fluency - no diacritics (cwpm)	23.68	25.09	24.93
Reading comprehension – 60 seconds (max. 6)	1.37	1.57	1.54
Reading comprehension – 90 seconds (max. 6)	1.92	2.22	2.20
Listening comprehension (max. 6)	3.07	3.31	3.28
Number of sample Grade 2 children	177	1443	1339

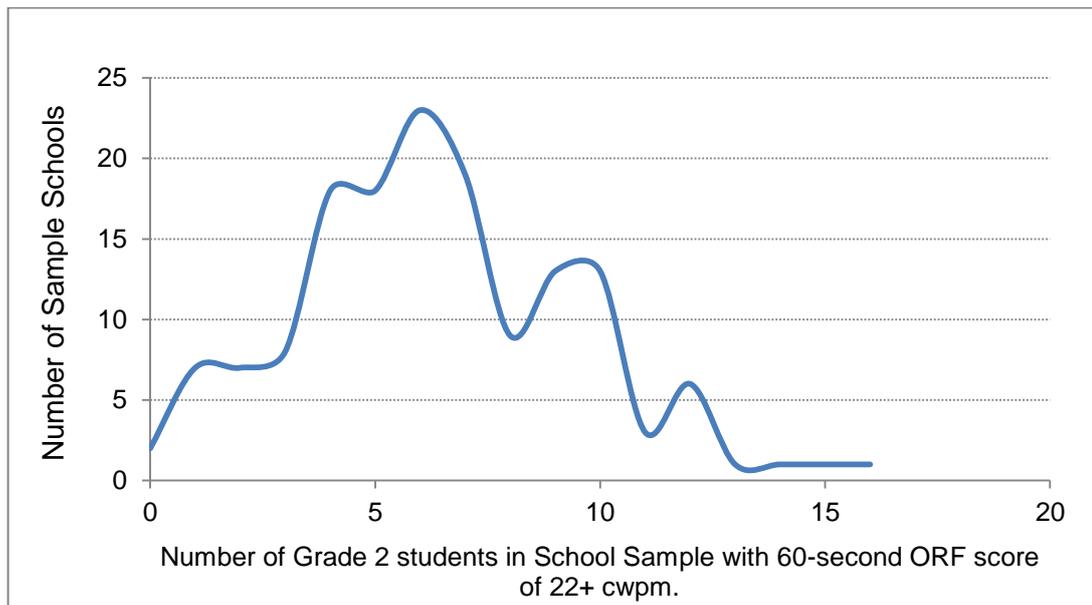
slightly lower results in the smallest classes – a difference that is statistically significant at the 0.05 level. The difference is not, however, meaningful. There was no significant difference in reading outcomes between medium and large Grade 2 classes. The finding is that there are no meaningful differences in reading results by class size.

Variability across MOEHE Schools in Grade 2 Reading Performance

All education ministries are interested to know the extent of school variability in reading proficiency: Are there marked differences in average reading performance between schools? Do the average scores truly reflect the performance of most schools? Or are there marked differences in average reading performance between schools that are obscured by the overall mean scores?

Figure 13 presents an analysis of school variability on the priority subtask of oral reading fluency (60 seconds with diacritics). The *x-axis* of this line graph is the number of sample students in each school who read at an intermediate level or higher of 22+ correct words per minute on the 60-second ORF subtask with diacritics. The *y-axis* of this graph is the number of MOEHE schools in the sample.

Figure 13: Divergence in Reading Performance among Sample Schools



The data plotted in Figure 13 clearly indicate low divergence between MOEHE schools on this core reading skill. The shape of the line graph approximates a bell curve with few schools at either end of the x-axis and the majority of schools at the graph center. The very large majority (75%) of sample schools had 6-12 of their 20 sample students reading 22+ correct words per minute on this standard oral reading fluency subtask. Only a few schools had fewer or greater numbers of capable readers in their EGRA samples. Similar plots of average school performance on other EGRA subtasks are comparable to Figure 13.

The conclusion of this analysis is that MOEHE schools show only limited variability (low divergence) in their Grade 2 reading outcomes. Most schools are performing within a central range of moderate reading proficiency. It is important that the significance of this finding be clear: the average mean scores for the sample of all Grade 2 students are “true” measures of central tendency that reflect the reality of Grade 2 reading proficiency in most MOEHE schools of the West Bank.

Reading Results by Governorate and Districts

Stakeholders can be expected to ask for the EGRA results of specific governorates or districts. The answer is that this EGRA sample of 150 schools was not large enough to provide results by governorates or districts. The stratification of West Bank schools by school gender required a minimum, statistically-valid sample of 120 schools: 40 schools of each of the 3 school genders (boys’ schools, girls’ schools and mixed schools). The statistical minimum for each level of stratification is 40 schools. To provide statistically valid results for each of the 11 governorates would have required a sample of 440 schools; 640 schools if results were desired for the 16 districts. The purpose of this EGRA baseline was to provide system-level results. Results at governorate and districts level would not have added significant insight to these findings. The findings of low divergence between sample MOEHE schools confirm this conclusion. Enlarging the sample to report results by governorate or district would not have justified the much greater assessment cost.

Summary Conclusions

The priority, summary conclusions of this EGRA baseline for Grade 2 in MOEHE schools of the West Bank are the following:

7. ***Good foundation of pre-reading skills:*** The sample Grade 2 students demonstrated good proficiency in the basic pre-reading skills of letter sounds knowledge and nonword / invented word decoding. Average scores on these two subtasks are, by far, the highest of any Arabic EGRA in Grades 2 or 3; and the percentage of zero scores are the lowest. These are the priority reading skills of Grades 1 and 2. **There is a strong foundation of reading proficiency in these pre-reading skills among MOEHE Grade 2 students. This is a strong and positive finding.**
8. ***Students struggle to read connected text in passages:*** Most Grade 2 students are struggling to read familiar words in connected text (passages, stories), indicative of low reading fluency. Average scores on the oral reading fluency subtask (with diacritics) are low and the percentage of students scoring zero is high. This finding suggests that early grade instruction in the Arabic language arts now chiefly teaches vocabulary as isolated words. **Students need greater practice reading new vocabulary in short and simple sentences (with diacritics) in class.**

9. **Low reading comprehension:** Most Grade 2 students do not comprehend what they are reading. Average scores on reading comprehension (60 seconds reading with diacritics) are low and the percentage of students with zero scores is high. This is both the direct outcome of low reading fluency and diminished attention to directed teaching and student practice of comprehension skills. The generally good results on the listening comprehension subtask suggest that most Grade 2 children capably understand formal Arabic. Hence **the low scores on reading comprehension in Grade 2 chiefly result from low reading fluency and insufficient practice of reading comprehension skills**. Comprehension skills must be learned, practiced and mastered – they do not come automatically to students who mechanically read well.
10. **Few illiterates in Grade 2:** **The percentage of illiterate Grade 2 students in MOEHE schools is low: less than 5%. This is an important and positive finding.** Few students are now being “left behind” in early grade instruction in Arabic. The number of nonreaders can and should, however, be reduced. For all reading subtasks, targets were set at the benchmarking workshop for lower percentage zero scores to be achieved by 2018. *It is important that these targets for fewer nonreaders be a top priority for benchmark achievement.*
11. **Low variability among MOEHE schools in Grade 2 reading proficiency:** **There is low variability (low divergence) among MOEHE primary schools in the reading proficiency of Grade 2 students. This finding has important social equity implications.** The large majority of MOEHE schools are performing comparably in Arabic reading instruction of their Grade 2 students. There are not significant numbers of “strong” schools and “weak” schools; nor a large divide between schools in the reading performance of their Grade 2 students. The overall EGRA results are “true” and general indicators of Grade 2 reading outcomes in most MOEHE schools. Most Grade 2 classes include a full range of student reading abilities, from weak to strong readers.
12. **Significant gender gap in Grade 2 reading proficiency:** Girls typically outperform boys in most early reading subtasks, especially the timed tests of pre-reading skills and oral reading fluency. This result is common to the large majority of languages and countries in which EGRAs have been implemented. It is true of all Arabic EGRAs. The gender gap in reading proficiency for this sample of West Bank Grade 2 students is, however, the largest of all Arabic EGRAs. The results cannot be extrapolated to later grades. The gender gap may diminish among older students. These Grade 2 results do recommend that **schools, teachers and parents be aware of this gender gap and ensure that boys have equal or greater opportunity for reading practice in class.**

Priority Recommendations and Next Steps

The May 2014 presentation of results of this EGRA baseline for Grade 2 in West Bank public schools proposed the following recommendations from the findings:

1. **Set and publicize reading standards (*benchmarks*) for the early grades:**
The MOEHE is advised to establish expected levels of reading proficiency for specific subtasks, for each grade. Early grade classroom and Arabic teachers, and their supervisors, should be trained in the expected levels of reading proficiency for each grade. Consider social marketing and video demonstrations of early grade students reading at these expected levels of reading proficiency.
2. **Significantly lower the percentages of zero scores – nonreaders – in oral reading fluency (with / without diacritics) and reading comprehension.**
3. **Increase average scores in oral reading fluency.**
4. **Apply teaching strategies and classroom exercises that directly strengthen comprehension – both listening and reading comprehension.**

The May 2014 presentation of EGRA results to the MOEHE pointedly included a separate presentation – *Thinking about Standards* – by Dr. Luis Crouch. Subsequent to these presentations and recommendations, the MOEHE invited technical expertise from RTI to lead a workshop for MOEHE and selected Palestinian NGO representatives in establishing early reading benchmarks. That workshop was led by Dr. Crouch in Ramallah on September 21-22, 2014. The MOEHE has effectively taken action on the first recommendation and proposed “next step” from this West Bank EGRA for improved reading outcomes by Palestinian children.

The 2nd and 3rd recommendations beg the question: *How much improvement in formal Arabic reading skills is possible in Grade 2?* The experience of the pilot USAID Girls’ Improved Learning Outcomes (GILO) Project in Egypt may be indicative. After implementing a single-year program of teach training, providing supplemental teacher resources and exercise routines for Grade 2 reading instruction, and technical support to Arabic supervisors for teacher coaching, impressive gains in reading outcomes were observed and measured in the 30 pilot-supported public schools.

Before the intervention, *nearly half* of randomly-selected Grade 2 students had zero scores on the letter sounds knowledge (48%) and oral reading fluency (44%) subtasks, as measured in a 2009 EGRA baseline in GILO-supported and control schools. At the immediate end of that Grade 2 intervention – in April 2011 of the same school year in which the reading intervention commenced (October 2010) – the post-intervention EGRA in April 2011 found that percentage of zero scores on these two subtasks had dropped sharply to 11% and 21% respectively in supported schools. The percentages of nonreaders had declined impressively. And average Grade 2 scores on these same subtasks had improved dramatically: up 192% in letter sound knowledge and 91% in oral reading fluency. These are impressive gains from a 6-month pilot intervention. The gains in Grade 2 reading performance were equivalent

to a full year of additional schooling. The Grade 2 students who benefitted from the intervention were reading better than Grade 3 students tested in the 2009 baseline (conducted in Grades 2-4).

A system-wide intervention in hundreds of public schools might not replicate these pilot results in a single year. But the potential for very significant improvements in reading proficiency is demonstrably possible in just 2-3 years of professional support for improved teacher training, improved teacher resources, strengthened supervision and constructive coaching of early grade teachers in reading instruction, and improved accountability from empirical assessment.

Annex A: The West Bank EGRA Instrument for Grade 2

List of Participants: MOEHE Tool Development Team for the EGRA Instrument

List of Assessors, EGRA

أداة تقييم مهارات القراءة في الصف الثاني – الضفة الغربية ، فلسطين
وزارة التربية والتعليم العالي

نموذج التعليمات للفاحص
أداة تقييم مهارات القراءة في الصف الثاني
الضفة الغربية ، فلسطين - مارس 2014

تعليمات عامة :

من المهم جداً إيجاد بيئة مريحة وقائمة على اللعب وخلق حوار من خلال مناقشة موضوعات تهتم التلميذ/التلميذة وذلك لتكوين ألفة مع الطلبة موضع الفحص. كما يجب أن يدرك التلميذ/التلميذة بأن تطبيق استمارة الفحص هي عبارة عن مواقف يسود فيها اللعب (على سبيل المثال سوف نقرأ بعض الكلمات التي ليس لها أي معنى). من المهم أيضاً قراءة الأجزاء الموجودة داخل الصناديق بصوت عالٍ وواضح. فإذا لم يفهم الطفل التعليمات، ففسرها له بلهجته الأصلية. بعد انتهاء المقابلة، اشكر التلميذ/التلميذة على الوقت والجهد الذي تم بذله ثم قدم له قلم رصاص هدية.

الموافقة اللفظية :

اسمي :

قل للطالب ما يلي :

- نحاول ان نفهم كيف يتعلم الأطفال القراءة . تم اختيارك بمحض الصدفة.
- نحتاج الى مساعدتك في إنجاز هذا العمل ، فإذا لم ترغب في ذلك فلك مطلق الحرية في عدم المشاركة.
- سوف نلعب معاً لعبة القراءة. سوف أطلب منك أن تقرأ بصوت عالٍ أحرفاً وكلمات وقصة قصيرة.
- باستخدام ساعة التوقيت هذه ، سوف أعرف المدة التي استغرقتها في القراءة.
- هذا ليس اختباراً ولن يؤثر على درجاتك المدرسية.
- لن أكتب اسمك وبالتالي لا أحد يمكنه أن يعرف أن هذه هي استجاباتك.
- مرة أخرى ، إذا لم ترغب في المشاركة أو لم ترغب في الإجابة عن سؤال ما فلك ذلك .
- هل يمكننا أن نبدأ ؟

تم الحصول على الموافقة : نعم لا (أوقف التقييم وأطلب فحص تلميذ آخر)

إسم الفاحص		إسم المدرسة	
رقم كود المدرسة	رقم كود المدرسة	نوع المدرسة	1 = بنين 2 = بنات 3 = مختلطة
يوم شهر سنة	تاريخ الميلاد	النوع	1 = ذكر 2 = أنثى
..... / / / /	عدد فقرات لدراسة بالمدرسة	1 = فترة واحدة 2 = فقرتين

ابدأ وحدد الوقت (ساعة ودقائق) : — إلى : —

إستخدام قائد الفريق

إسم القائد	تتبعيد من العينة الاصلية؟
..... لا	نعم

جزء 0 - معرفة الاتجاهات

اعرض على التلميذ/ التلميذة قطعة القراءة وقل له: رجاءً لا تقرأ فقط أشر لي من أين تبدأ القراءة . ثم حرك إصبعك في الاتجاه الصحيح للقراءة ، ثم ضع إصبعك على آخر كلمة.

01 هل وضع التلميذ/التلميذة إصبعه على أول كلمة	نعم	لا
02 هل حرك التلميذ/التلميذة أصبعه من اليمين الى اليسار	نعم	لا
03 هل وضع التلميذ/التلميذة إصبعه على آخر كلمة في النص	نعم	لا

جزء 1 - الحروف مع الحركات

قدم للتلميذ/ للتلميذة ورقة الأحرف الموجودة في الصفحة الثانية من كراس التلميذ/ التلميذة وقل له :

هذه الصفحة مليئة بالأحرف. رجاءً انطق بصوت كل حرف مع الحركة. فعلى سبيل المثال اسم الحرف مع الحركة "سـ" (أشـر الى الحرف "سـ" هو "سـ" ، الآن حاول انت . انطق لي أنت صوت هذا الحرف (أطلب من التلميذ/ التلميذة أن يقرأ حرف "سـ"). فإذا استطاع التلميذ/التلميذة أن يقرأ الحرف مع الحركة بشكل صحيح فقل له "أحسنـت"، صوت الحرف "كـ" . أما إذا أخطأ التلميذ/التلميذة فقل له صوت هذا الحرف "كـ" . الآن حاول مرة أخرى: انطق لي صوت هذا الحرف (اطلب من التلميذ/التلميذة أن يقرأ الحرف مع الحركة بشكل صحيح ، فقل له "أحسنـت"، صوت هذا الحرف "نـ" . أما إذا أخطأ التلميذ/التلميذة فقل له القراءة الصحيحة لهذا الحرف "نـ" . هل تفهم ما هو مطلوب منك ؟ عندما أقول لك " ابدأ " اقرأ الحروف بأفضل طريقة ممكنه، سوف ألتزم الصمت وأستمع لما تقول مستعداً؟ ابدأ

يجب على التلميذ قراءة الحروف بتشكيل لكي تحسب له صحيحة. رجاءً لاحظ أن التلميذ/التلميذة يلتزم بالقراءة من اليمين الى اليسار ابتداءً من السطر الأول. يجب على التلميذ قراءة الحروف بتشكيل لكي تحسب له صحيحة 

شغل ساعة التوقيت عندما يبدأ التلميذ/ التلميذة في القراءة. تابع التلميذ/ التلميذة باستخدامك قلم الرصاص وضع إشارة " / " لأي حرف يخطئ التلميذ/ التلميذة في نطقه. احسب التصحيح الذاتي الذي يقوم به التلميذ/التلميذة على أنه صحيح. التزم الهدوء إلا عندما تكون استجابات التلميذ/التلميذة على النحو التالي: إذا تردد التلميذ/ التلميذة لمدة 3 ثوان، قل له "كامل" ، واطلب منه أن يستمر في نطق الأحرف. احسب هذا الحرف الذي قراه التلميذ/ التلميذة على أنه غير صحيحة. بعد 60 ثانية ضع علامة [عند الحرف الأخير الذي قراه التلميذ/التلميذة ثم أطلب من التلميذ/التلميذة أن يتوقف .

10/ بِ س ع ت ن م ه ر ث غ
20/ أ ط ذ ش د ص ظ ك و ي
30/ ل م غ ت ا ع ض ف ق ز
40/ س د غ ط ك ن ق ش م ن ي
50/ ج ح ب ه غ غ و ق ث ص

الوقت الذي تشير إليه الساعة _____
تم إيقاف التريب نظراً لأن التلميذ/ التلميذة لم يستطع قراءة السطر الأول

جزء 2 - قراءة الكلمات المألوفة

قدم للتمييز/التمييز نموذج الكلمات الموجودة في الصفحة الثالثة من النموذج الخاص بالتمييز/التمييز، قل:

فيما يلي بعض الكلمات. أريد منك أن تقرأ لي جميع ما تستطيع قراءته من الكلمات المعروضة أمامك هذا مثال على ذلك: "باب"

الآن حاول أنت: اقرأ لي هذه الكلمة "أمي" إذا استجاب التمييز/التمييز بشكل صحيح، فقل أحسنت، هذه الكلمة هي "أمي". أما إذا أخطأ التمييز/التمييز، فقل هذه الكلمة هي "أمي"

الآن حاول مرة أخرى اقرأ لي هذه الكلمة "صف" إذا استجاب التمييز/التمييز بشكل صحيح، فقل أحسنت، هذه الكلمة هي "صف". أما إذا أخطأ التمييز/التمييز، فقل هذه الكلمة هي "صف"

. هل فهمت المطلوب؟ هل أنت مستعد؟ ابدأ

شغل ساعة التوقيت عندما يبدأ " التمييز/ التمييز في القراءة. ضع علامة / " لكل كلمة يقرأها التمييز/التمييز بشكل غير صحيح.

احسب الكلمات التي يقوم التمييز/ التمييز بتصحيح نفسه فيها على أنها صحيحة مع دائرة حول الكلمة. التزم الهدوء إلا عندما تكون استجابة التمييز/التمييز على النحو التالي: إذا تردد التمييز/التمييز لمدة 3 ثوان، اقرأ له الكلمة ثم أشر إلى الكلمة اللاحقة واطلب منه أن يستمر في قراءة الكلمات. احسب هذه الكلمة التي قرأها على أنها غير صحيحة. إذا أخطأ التمييز في خمس كلمات انتقل فوراً إلى السؤال التالي بعد 60 ثانية، ضع علامة [عند آخر كلمة قرأها التمييز/التمييز. ثم قل "توقف"

5/	حَدَّثَ	مَا	قِصَّتُهُ	اللُّصُّ	فَأَنَارَ
10/	تَرَعَى	ابْتَسَمَ	الْحَلْوَى	ذَهَبَ	بِطَبَقٍ
15/	فِي	قَالَ	صَخْرَةَ	يَلْتَهُمُ	فَكَرَّتْ
20/	عِنَايَةَ	مَصْنَعٍ	هَنْفٍ	أَشْعَبَ	مُرَّرَعْتَهُ
25/	تَرَضَعَ	بَاعَةَ	مِنْ	أَبِي	يَتَفَقَّدُ
30/	العَجُوزِ	بُئِي	فَادِي	أَمْسَكَ	سَجَنَ
35/	رَهْرَ	حِفَاظًا	فَارِسِ	رَدَّ	تَفْرَعِ
40/	يَا	عِيدِ	بُيُوتِ	الدَّهَابِ	مَغِيبِ
45/	ثِمَارِ	عُلْبَةٍ	رَأَتْ	الْحَجَلِ	تَرْتَفِعِ
50/	أَخْبَارِ	طَوْلُكْرَمِ	سَأَلَتْ	مُعْتِمَةَ	رَجُلٍ

الوقت الذي تشير إليه الساعة _____
 تم إيقاف التدریب نظراً لأن التمييز/التمييز لم يستطع قراءة المسطر الأول.

The EGRA Instrument for Grade 2 – 2014

جزء 3 - قراءة كلمات غير مألوقة

قدم للتمييز/ للتمييز صفحة الكلمات التي ليس لها معنى في الصفحة الرابعة من نموذج التمييز/التمييز وقل له :

انظر إلى هذه الكلمة ، إنها كلمة ليس لها معنى معروف لديك : هذا مثال على ذلك : " تَمَش " .
الآن حاول أنت : اقرأ لي هذه الكلمة " راك " إذا استجاب التمييز/التمييز بشكل صحيح، فقل أحسنت ، هذه الكلمة هي " راك" .
أما إذا أخطأ التمييز/التمييز ، فقل هذه الكلمة هي " راك" .
الآن حاول مرة أخرى : اقرأ لي هذه الكلمة " لَمك " إذا استجاب التمييز/التمييز بشكل صحيح، فقل أحسنت ، هذه الكلمة هي " لَمك " .
أما إذا أخطأ التمييز/التمييز ، فقل هذه الكلمة هي " لَمك " .
هل فهمت المطلوب ؟ هل أنت مستعد ؟ ابدأ .

شغل ساعة التوقيت عندما يبدأ " التمييز/التمييز في القراءة . من المهم وضع إشارة " / " لكل كلمة يقرأها التمييز/التمييز بشكل غير صحيح .

احسب الكلمات التي يقوم التمييز/التمييز بتصحيح نفسه فيها على أنها صحيحة وضع دائرة حول الكلمة. التزم الهدوء إلا عندما تكون استجابة التمييز/التمييز على النحو التالي: إذا تردد التمييز/التمييز لمدة 3 ثوان ، وأطلب منه أن يستمر في قراءة الكلمات. احسب الكلمة التي قرأها على أنها غير صحيحة . إذا أخطأ التمييز في السطر الأول انتقل فوراً إلى السؤال التالي بعد 60 ثانية، ضع علامة [عند آخر كلمة قرأها التمييز / التمييز ثم قل "توقف".

5	صَمَح	قَاذ	سَهْرَامِي	جَرَق	رَا
10	أَشِغ	ضَخْرَص	حَجِيث	بِي	لِهْس
15	حَرْجَل	تَاَف	ضَا	زَمْفَل	فَرْخ
20	وَلِح	جَمَطَح	ذَف	نَادِك	رَكْش
25	نَصْرَج	رِخِب	خَرِيض	سَا	صَعْفَط
30	زَطَع	ذِمْتَاخ	مُذَاكِب	تَمْرَخ	فِع
35	ظَصِيح	دَهْوَلَم	حَمْلَاك	هَنْكِر	عَمَش
40	لُكَذ	تِيث	دَاَف	قَنْط	رُبْح
45	شُدِخ	رَبِك	دِعْج	أَفِب	نَرْدَخَة
50	شَرِيك	عَوْدَر	طَاهِس	لَنَا	أَفِي

الوقت الذي تشير إليه الساعة _____
تم إيقاف التدريب نظراً لأن التمييز/التمييز لم يستطع قراءة السطر الأول.

جزء 4 - أ - قراءة قطعة

قدم للتمييز/ للتمييز قطعة القراءة الموجودة في آخر صفحة من نموذج التمييز/ للتمييز وقل له :

أريد منك أن تقرأ هذه القطعة بصوت عالٍ ، ركز على الحركات، وعندما تنتهي من قراءتها سوف أسألك بعض الأسئلة عنها.
هل فهمت المطلوب ؟ عندما أقول "ابدا" اقرأ القصة
هل أنت مستعد ؟ ابدأ

شغل ساعة التوقيت عندما تقول للتمييز/ للتمييز " ابدأ" . ضع علامة " / " لكل كلمة يقرأها التلميذ/ التلميذة بشكل غير صحيح.
احسب الكلمات التي يقوم التلميذ/ التلميذة بتصحيح نفسه فيها على أنها صحيحة. التزم الهدوء إلا عندما تكون استجابة التلميذ/ التلميذة على النحو التالي: إذا تردد لمدة 3 ثوان ، وشجعه على أن يستمر في قراءة الكلمات، استمر. احسب الكلمة التي تقرأها أنت للتمييز/ التلميذة على أنها غير صحيحة. بعد 60 ثانية، ضع علامة [عند آخر كلمة قرأها التلميذ/ التلميذة ثم قل " توقف"

8 يَزُورُ أَطْفَالَ مَدِينَةِ طُولُكْرَمِ الْحَدِيقَةِ كُلَّ يَوْمِ جُمُعَةٍ،
18 وَتَلْعَبُ حَنِينٌ عَلَى أَرْجُوْحَتِهَا الْمُعَلَّقَةِ عَلَى الشَّجَرَةِ، وَيُطَيِّرُ سَامِي طَيَّارَتَهُ
29 الْوَرَقِيَّةَ فِي الْهَوَاءِ عَالِيًا. ذَاتَ يَوْمٍ هَبَّتْ عاصِفَةٌ قَوِيَّةٌ، فَأَمْتَلَأَتْ أَرْضَ
41 الْحَدِيقَةِ بِالْأَوْرَاقِ وَالْغُبَارِ، وَلَمْ تَعُدْ طَيَّارَةُ سَامِي تَعْلُو فِي الْهَوَاءِ كَمَا كَانَتْ.
51 وَوَجَدَتْ حَنِينٌ أَرْجُوْحَتَهَا مَكْسُورَةً ، فَطَلَبَتْ حَنِينٌ مِنْ سَامِي النَّفْكَيرَ فِي
57 طَرِيقَةٍ لِإِعَادَةِ تَنْظِيفِ الْحَدِيقَةِ وَإِصْلَاحِ أَرْجُوْحَتِهَا.

الوقت الذي تشير إليه الساعة _____
تم إيقاف التدريب نظراً لأن التلميذ/ التلميذة لم يستطع قراءة السطر الأول.

لا اسمح للتمييز/ للتمييز بالاحتفاظ بقطعة القراءة عقب الانتهاء من قراءتها. بعد قراءة كل سؤال أعط التلميذ/ التلميذة 10 ثانية على الأكثر للإجابة عن كل سؤال. ضع إشارة ✓ في العمود الأول إذا كانت الإجابة صحيحة وفي العمود الثاني إذا كانت الإجابة غير صحيحة. وفي العمود الثالث في حلة عدم الإجابة.

الآن سوف أسألك عدة أسئلة عن القطعة التي قرأتها ، حاول الإجابة عن الأسئلة.			
صحيح	غير صحيح	لا إجابة	1- متى يزور الأطفال الحديقة ؟ (يوم الجمعة)
صحيح	غير صحيح	لا إجابة	2- أين تلعب حنين ؟ (في الحديقة / على أرجوحيتها)
صحيح	غير صحيح	لا إجابة	3- ماذا يفعل سامي؟ (يطير طيارة)
صحيح	غير صحيح	لا إجابة	4- لماذا امتلأت أرض الحديقة بالغبار؟ (هبت عاصفة)
صحيح	غير صحيح	لا إجابة	5- ما الذي حدث لأرجوحة حنين؟ (انكسرت)
صحيح	غير صحيح	لا إجابة	6- ماذا طلبت حنين من سامي؟ (تنظيف الحديقة وإصلاح أرجوحيتها)

جزء 4 - ب - قراءة قطعة

قدم للتمييز/ للتمييز قطعة القراءة الموجودة في آخر صفحة من نموذج التمييز/ للتمييز وقل له :

أريد منك أن تقرأ هذه القطعة بصوت عالٍ ، ركز على الحركات ، وعندما تنتهي من قراءتها سوف أسألك بعض الأسئلة عنها.
هل فهمت المطلوب ؟ عندما أقول "ابدا" اقرأ القصة
هل أنت مستعد ؟ ابدأ

شغل ساعة التوقيت عندما تقول للتمييز/ للتمييز " ابدأ" . ضع علامة " / " لكل كلمة يقرأها التمييز/ للتمييز بشكل غير صحيح.
احسب الكلمات التي يقوم التمييز/ للتمييز بتصحيح نفسه فيها على أنها صحيحة. التزم الهدوء إلا عندما تكون استجابة التمييز/ للتمييز على النحو التالي: إذا تردد لمدة 3 ثوان ، وشجعه على أن يستمر في قراءة الكلمات، استمر. احسب الكلمة التي يقرأها أنت للتمييز/ للتمييز على أنها غير صحيحة. بعد 90 ثانية، ضع علامة [عند آخر كلمة قرأها التمييز/ للتمييز ثم قل " توقف"

- 7 تعيش أم حمدان وأبناؤها في قرية صغيرة ،
16 ولها خمس دجاجات ، وفي كل صباح ، تفتح باب الخم ؛
26 ليأكل الدجاج الحب والعشب ، وفي أحد الأيام اختفت دجاجة سوداء ،
37 فبحثت عنها أم حمدان ولم تجدها ، وبعد ثلاثة أسابيع إذ بالدجاجة
48 السوداء تخرج من كومة حطب ، ومعها عشرة فراخ صغيرة، قال حمدان
57 بصوت عال: أمي! الدجاجة عادت! طارت أم حمدان فرحاً.

الوقت الذي تشير إليه الساعة _____
تم إيقاف التريب نظراً لأن التمييز/ التمييز لم يستطع قراءة السطر الأول.

لا اسمح للتمييز/ للتمييز بالاحتفاظ بقطعة القراءة عقب الانتهاء من قراءتها. بعد قراءة كل سؤال أعط التمييز/ التمييز 10 ثانية على الأكثر للإجابة عن كل سؤال. ضع إشارة ✓ في العمود الأول إذا كانت الإجابة صحيحة وفي العمود الثاني إذا كانت الإجابة غير صحيحة. وفي العمود الثالث في حلة عدم الإجابة.

الآن سوف أسألك عدة أسئلة عن القطعة التي قرأتها ، حاول الإجابة عن الأسئلة.			
صحيح ___	غير صحيح ___	لا إجابة ___	1- أَيْنَ تَعِيشُ أُمُّ حَمْدَانَ ؟ (قَرْيَةً)
صحيح ___	غير صحيح ___	لا إجابة ___	2- كَمْ دَجَاجَةً لِأُمِّ حَمْدَانَ ؟ (خَمْسُ دَجَاجَاتٍ)
صحيح ___	غير صحيح ___	لا إجابة ___	3- لِمَاذَا تَفْتَحُ أُمُّ حَمْدَانَ بَابَ خَمِّ الدَّجَاجِ ؟ (لِيَأْكُلَ)
صحيح ___	غير صحيح ___	لا إجابة ___	4- مَا الَّذِي بَحَثَتْ عَنْهُ أُمُّ حَمْدَانَ ؟ (دَجَاجَةً)
صحيح ___	غير صحيح ___	لا إجابة ___	5- مَاذَا كَانَ مَعَ الدَّجَاجَةِ جِئِمَا خَرَجَتْ مِنْ كَوْمَةِ الحَطَبِ؟ (فِرَاحٌ)
صحيح ___	غير صحيح ___	لا إجابة ___	6- لِمَاذَا فَرِحَتْ أُمُّ حَمْدَانَ ؟ (عَوْدَةُ الدَّجَاجَةِ)

جزء 5 - فهم المسموع

هذا التمرين لا يعتمد على التوقيت . لا يوجد نموذج أو ورقة للتمديد/للتلميذة . في هذا التمرين يقوم الفاحص بقراءة النص التالي مرة واحدة وبصوت عالٍ ، تتم القراءة ببطء أي بمعدل كلمة في الثانية الواحدة .

عقب الانتهاء من قراءة النص، اعط التلميذ/التلميذة 10 ثانية على الأكثر للإجابة عن كل سؤال قل للتمديد/للتلميذة :

سوف أقرأ لك النص مرة واحدة ثم أوجه لك بعض الأسئلة . رجاء الاستماع بحرص والإجابة عن الأسئلة على أفضل وجه ممكن . هل فهمت المطلوب ؟

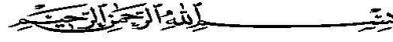
النص:

غَابَتِ الْمُعَلِّمَةُ عَنِ الْمَدْرَسَةِ، فَعَلِمَ الطُّلَابُ بِمَرَضِهَا.
قَالَتْ سُهَيْرٌ: سَنَرُورُ مُعَلِّمَتُنَا ، وَلَكِنْ لَا نَسْتَطِيعُ أَنْ نَذْهَبَ جَمِيعًا،
فَعَدَدْنَا كَبِيرًا ، عَلَيْنَا أَنْ نَخْتَارَ ثَلَاثَةً مِّنَّا. اشْتَرَى الطُّلَابُ بَاقَةَ مِنَ الزُّهُورِ وَذَهَبَ الثَّلَاثَةُ
إِلَى بَيْتِ الْمُعَلِّمَةِ ، وَقَالُوا: يَسُرُّنَا أَنْ نُقَدِّمَ لَكَ هَذِهِ الْهَدِيَّةَ ، وَنَرْجُو لَكَ الصِّحَّةَ وَالْعَافِيَةَ.
قَدِّمَتِ الْمُعَلِّمَةُ الْحَلْوَى لَهُمْ، وَقَالَتْ: أَنَا سَعِيدَةٌ بِزِيَارَتِكُمْ، أَرْجُو أَنْ تَثَابِرُوا عَلَى اجْتِهَادِكُمْ.

اسأل التلميذ/التلميذة :

1	لماذا غابت المعلمة عن المدرسة ؟ (مريضة)	إجابة صحيحة ___	إجابة غير صحيحة ___	لا إجابة ___
2	لماذا لا يستطيع جميع الطلاب الذهاب إلى المعلمة ؟ (عددٌهم كبيرٌ)	إجابة صحيحة ___	إجابة غير صحيحة ___	لا إجابة ___
3	ما الهدية التي اشترىها الطلاب ؟ (باقةٌ زهور)	إجابة صحيحة ___	إجابة غير صحيحة ___	لا إجابة ___
4	ماذا تمنى الطلاب للمعلمة ؟ (الصحة - العافية)	إجابة صحيحة ___	إجابة غير صحيحة ___	لا إجابة ___
5	ماذا قدمت المعلمة للطلاب ؟ (الحلوى)	إجابة صحيحة ___	إجابة غير صحيحة ___	لا إجابة ___
6	بماذا نصحت المعلمة الطلاب ؟ (المثابرة - الاجتهاد)	إجابة صحيحة ___	إجابة غير صحيحة ___	لا إجابة ___

تذكر: تسجيل وقت نهاية الفحص بالصفحة الأولى



State of Palestine

Ministry of Education & Higher Education
D. G of Assessment, Evaluation & Examinations



دولة فلسطين
وزارة التربية والتعليم العالي
الإدارة العامة للقياس والتقويم والامتحانات

Participants in the Tool Adaptation Workshop for the West Bank EGRA
February 23-24, 2014

الرقم	الاسم	الوظيفة	السكن
١	أ. ذياب عقل نايف تركمان	مشرف مرحلة / الوزارة	أريحا
٢	أ. ماجد عودة محمد أبو داود	مشرف مرحلة ادبي / الوزارة	بيت لحم
٣	د. فيصل محمود سباعنة	مشرف مرحلة ادبي / الوزارة	قباطية
٤	أ. عاهد حسين عياش	مشرف تربوي / الوزارة	قرى نابلس
٥	أ. خالد حسن محمود ذياب	دائرة القياس والتقويم / الوزارة	جنين
٦	د. سهير محمد يوسف قاسم	مدير دائرة التدريب / الوزارة	رام الله
٧	أ. نهاية خالد شاهين	جمعية المصادر للطفولة المبكرة	رام الله
٨	أ. وحيد حسين عبد قننة	دائرة القياس والتقويم / الوزارة	قرى رام الله
٩	أ. أحمد محمد خطيب	مركز المناهج الفلسطينية / الوزارة	قرى رام الله
١٠	أ. محمود عيد	مدير النشاط الثقافي / الوزارة	قرى قلقيلية
١١	أ. صادق عيسى الخضور	نائب مدير عام المعهد الوطني	الخليل
١٢	أ. وليد احشيش	مركز تطوير المعلم	رام الله
١٣	د. محمد مطر	دائرة القياس والتقويم / الوزارة	عزون / قلقيلية

Ramallah Tel (02-2969392) Fax (02-2969349) Gaza Telfax (08-2820411)
e-mail : ac-moe@palnet.com

The West Bank EGRA, Grade 2 - List of Assessors

The Assessors trained and deployed for the West Bank EGRA

#	الاسم الكامل كما في بطاقة الهوية	رقم بطاقة الهوية	تاريخ الميلاد	مكان السكن	مكان العمل/المديرية	جوال	email
1	سناء قزاع علي دراغمة	850750795	1/30/1987	طوباس	طوباس	597515832	ala-sawafra2010@hotmail.com
2	ريما بسام محمد ميثاني	941164063	4/5/1984	سيلة الحارثية	جنين	598753751	reem.mt.16@facebook.com
3	هبة يونس محمد همشري	919093757	7/21/1982	طوكرم	طوكرم	569555934	khalid.khalid95@yahoo.com
4	سماح ابراهيم عايش فرعيش	852954619	24/8/1990	يطا	جنوب الخليل	569240890	
5	عبير يونس مصطفى نواس	907674337	1/10/1981	رافات	سلفيت	598758346	helal.alex@hotmail.com
6	رانياة برهان ابراهيم داود	900060815	9/9/1976	قلقية	قلقية	597776935	ashtayeh@hbif.com.io
7	نهوند وديع حامد الشنية	90025499	17\2\1973	سام	نابلس	599149928	AAlaa_kh199@hotmail.com
8	زاهرة محمود علان خطاطبة	902064849	8\2\1977	بيت فوريك	نابلس	599771328	em_el_jood@hotmail.com
9	ساهرة وجيهه عبد الله عازم	905167532	3\3\1979	سبسطية	نابلس	598101785	Hanan-love-f@hotmail.com
10	حنان عقل رجب حمدان	921728036	13/11/1983	بيت سيرا	رام الله	598135251	inashindi@gmail.com
11	ايناس حبيب محمد هندي	900072158	21/9/1974	بيت سيرا	رام الله	597602947	Az14185@gmail.com
12	عبير بدوي رجب ابو ماضي	921670428	1/14/1984	الخليل	الخليل	597493572	annnnnnnr@hotmail.com
13	رانياة رشيد عبد الغني الكركي	981889207	8/30/1973	الخليل	الخليل	599879758	rawiajawabreh31@gmail.com
14	علا وحيد كمال جوايزة	902757202	8/17/1977	بيت فجر	بيت لحم	569401408	
15	نجوى ابراهيم عبدالله حماد	81081770	4/9/1979	القدس	القدس	569456842	suha-20001@hotmail.com
16	سهى عبد الكريم اسماعيل احمدان	907106314	9/2/1981	بدو	ض القدس	598741968	Lama-zaid@hotmail.com
17	نجلاء فتحي احمد جماعات	941603086	3/2/1981	العوجا	أريحا	595052982	Mote3hamarsha@hotmail.com
18	لما زيد محمد دراغمة	911504983	4/27/1982	طوباس	طوباس	599119769	
19	بسمة محمود احمد حمارشه	917292831	22/5/1971	يعبد	جنين	599129237	Laval-86@hotmail.com
20	فاطمة احمد صلاح هريشه	850733015	6/3/1987	قفين	طوكرم	599591476	
21	رولا جودت فائق جاعوب	914499249	13/5/1986	بيتا	جنوب نابلس	599220481	iawabrehnabila@gmail.com
22	سهير بسام حسن عبد الله	901161208	1/28/1977	سلفيت	سلفيت	598591399	Facebook:Maen Ahmad
23	نبيلة عمر محمد منصور	911534360	16/1/1983	قلقية	قلقية	599321442	ah.16mad5hamad1985@yahoo.com
24	معن محمد احمد الشيش	850244278	2/7/1989	الدوحة	بيت لحم	592373102	
25	احمد حماد احمد دبابسة	851178384	16/5/1985	نوبا	شمال الخليل	598150993	
26	عبير علي محمد جمل	858589328	2/21/1977	بيت سريك	ضواحي القدس	599534151	

The West Bank EGRA, Grade 2 - List of Assessors (continued)

Facebook:Hakema Mustafa	599293882	بيت لحم	بيت ساحور	11/1/1983	943332445	حاكمة عبدالرزاق عبدالرحمن نوفل	27
	592670606	جنوب الجليل	يطا	1/5/1985	920106887	ابسان خالد أحمد الحريزات	28
ansamhasan-1991@hotmail.com	599152721	قباطية	عراية	28/11/1979	904738564	منال عبد الله محمد نصار	29
salah1987.sh@hotmail.com	598203503	نابلس	عورتا	12/10/1987	920317872	صلاح عز الدين محمد شراب	30
safasoso92@yahoo.com	592670606	الخليل	طر تومية	4/4/1983	920317872	صفاء محمد الفطاطة	31
www-10sana@hotmail.com	598573108	قباطية	سائور	25/11/1986	850433913	سناه لطفي أحمد عيبه	32

Annex B: Sample Design and Weighting

This annex presents additional details about the sample design for this West Bank EGRA study.

Stage 1: The Selection and Weighting of Sample Schools

The MOEHE Planning Department provided a list of all MOEHE primary schools with Grade 2 students in the 11 governorates of the West Bank. These 976 MOEHE primary schools included a total enrollment of 44,336 Grade 2 students (22,171 girls – 22,165 boys) in the 2013/14 school year. These 976 schools were the sample frame.

Before drawing the random sample of schools to be included in the study, the 976 MOEHE primary schools with Grade 2 students in the West Bank were stratified by *school gender* at ministry request. MOEHE primary schools in the West Bank included comparable numbers of all-boy primary schools (303), all-girl primary schools (332) and co-ed primary schools (341) with Grade 2 students in 2013/14. From each of the three school gender types, 50 schools were randomly selected, to allow for maximum statistical power for each school gender type. The sample is comprised of 150 schools.

To achieve a representative sample of schools across the 16 districts and to ensure a mix of genders, the school sample had explicit stratification by school gender and implicit stratification by districts. An additional research question is whether children in small schools read better than children in larger schools. To help answer this question, and to ensure that all schools and all students have a non-zero probability of selection, even schools with a smaller number of Grade 2 students were included in the sample frame with a sampling probability proportional to grade 2 student enrollment. The minimum Grade 2 enrollment for a school to be selected for the West Bank sample was 10 students.

For each selected school, one replacement school was selected, to be used if the sampled school could not be visited. Replacement schools were held privately and were very similar to the selected school. Wherever possible, the replacement and sampled school were in the same districts and had other similar school characteristics (like size and school gender type). No replacement schools were required for this EGRA.

To make the sample representative of the total population of all MOEHE primary schools in the West Bank, school weights were calculated as the inverse of the selection probability of the school (Weight1, Stage 1 selection) and then scaled to the total number of schools of each school gender type. **Table B1** shows that the weighted counts and percentages of the sampled schools in each school gender are, in fact, representative of the population.

$$Weight_School_{(s,i)} = \frac{[Total\ Number\ of\ Grade2\ Students]in\ SchoolType(r)}{[Number\ of\ Grade\ 2\ Students\ in\ Selected\ School(i)]} * \frac{1}{50}$$

$r = 1\ to\ 3\ SchoolGenderTypes$

Table B1. Distribution of Schools in the Total Population and Sample, by School Gender

School Gender	Population		Sample		
	Total number of MOEHE primary schools with Grade 2 students	Percentage of schools (%)	Sampled number of schools	Weighted number of schools	Weighted percentage of sampled schools (%)
Boys' Primary	303	31.04%	50	293.78	32.08%
Girls' Primary	332	34.02%	50	324.59	35.45%
Mixed Primary	341	34.94%	50	297.31	32.47%
Total	976	100%	150	915.68	100%

Stage 2: The Selection and Weighting of Sample Students

The second stage of sample selection was the random stratified selection of students to be tested in each sample school. Grade 2 students were stratified by gender prior to selection and were selected with equal probability. For each sample school, the Assessment Team obtained complete lists of all enrolled Grade 2 students prior to the field visits. The names of 20 boys were randomly selected from each boys' school, 20 girls from each girls' school, and 10 boys and 10 girls each co-ed school. Random number tables were used to select the sample students from class lists in each school. For each single-gender school, a "reserve" list of up to 8 students was also randomly prepared to replace the pre-selected sample students absent on the day of the assessment. In mixed schools, the "reserve" list included 4 students of each gender. Both sample and "reserve" students were selected randomly and equitably across all Grade 2 classes in each sample school.

The students' weights were calculated by multiplying the school weight by the probability of selecting the student in the given school. This was then multiplied by the student scaled weights to guarantee that the sampled students were representative of the Grade 2 population in the West Bank.

$$Student\ Weight(g, s) = School_Weight(s, i) * Weight2(j)$$

Where: Weight2 (g,s) represents the weight of the second stage of selection: student by gender (g) within the selected school (s)

$$Weight2(g, s) = \frac{Total\ Number\ of\ Students\ by\ Gender\ in\ School\ s(g, s)}{Sampled\ Number\ of\ Students\ by\ Gender\ in\ School\ s(g, s)}$$

The gender distribution of the population and sample are shown in **Table B2**. In the total population of enrolled Grade 2 students in the 2013/14 school year, boys and girls are roughly evenly distributed (49.993% to 50.007%). Girls were slightly more numerous in our sample (49.68% to 50.33%).

Grade 2 male representation by school type can be seen in **Table B3** and Grade 2 female representation by school type can be seen in **Table B4**.

Table B2. Gender Distribution of Grade 2 Students in the Population and Sample

Student Gender	Population		Sample		
	Total Grade 2 Students	Percentage of Grade 2 Students (%)	Sampled number of Grade 2 Students	Weighted number of Grade 2 students	Weighted Grade 2 students (%)
Males	22,165	49.99	1,461	22,024	49.68
Females	22,171	50.01	1,492	22,312	50.33
Total	44,336	100.00	2,953	44,336	100.00

Table B3. Distribution of Grade 2 Boys in the School Gender Population and Samples

School Gender	Population		Sample		
	Total Grade 2 Males	Percentage of Grade 2 Males (%)	Sampled number of Grade 2 Males	Weighted number of Grade 2 Males	Weighted Grade 2 Males (%)
Boys Primary	15,750	71.1	994	15,750	71.5
Girls Primary	0	0.0			
Co-ed Primary	6,415	28.9	467	6,274	28.5
Total	22,165	100.0	1,461	22,024	100.0

Table B4. Distribution of Grade 2 Girls in the School Gender Population and Samples

School Gender	Population		Sample		
	Total Grade 2 Females	Percentage of Grade 2 Females (%)	Sampled number of Grade 2 Females	Weighted number of Grade 2 Females	Weighted Grade 2 Females (%)
Boys' Primary	0	0.0			
Girls' Primary	15,986	72.1	991	15,986	71.7
Mixed Primary	6,185	27.9	501	6,326	28.4
Total	22,171	100.0	1,492	22,312	100.0

Note about Precision Estimates:

The sample’s overall proportion of the population is not relevant in a study with a large population. In this study, our population is 44,336 Grade 2 students and our sample size is 2,953 (6.66%) of the population. The sample size, compared to the population, is not relevant because regardless of how large our population is, a sample size of 2,953 students provides us with extremely high statistical precision. For example, a 95% confidence band width of ± 3.5 is considered an acceptable precision for oral reading fluency (ORF). So with a mean ORF score of 16.67, we would say that a 95% confidence interval of (13.17, 20.17) is acceptable.

Figure B1 compares the accepted 95% confidence interval with the actual 95% confidence interval (15.78, 17.55) for the mean ORF score of 16.67 in this West Bank EGRA. As we can see in the figure, the actual 95% confidence interval is smaller (or ‘tighter’) than the acceptable 95% confidence interval, thus a sample size of 2,953 students provides more precise estimates than the acceptable precision level.

Figure B1: Accepted and Actual 95% Confidence Interval, Oral Reading Fluency Subtask

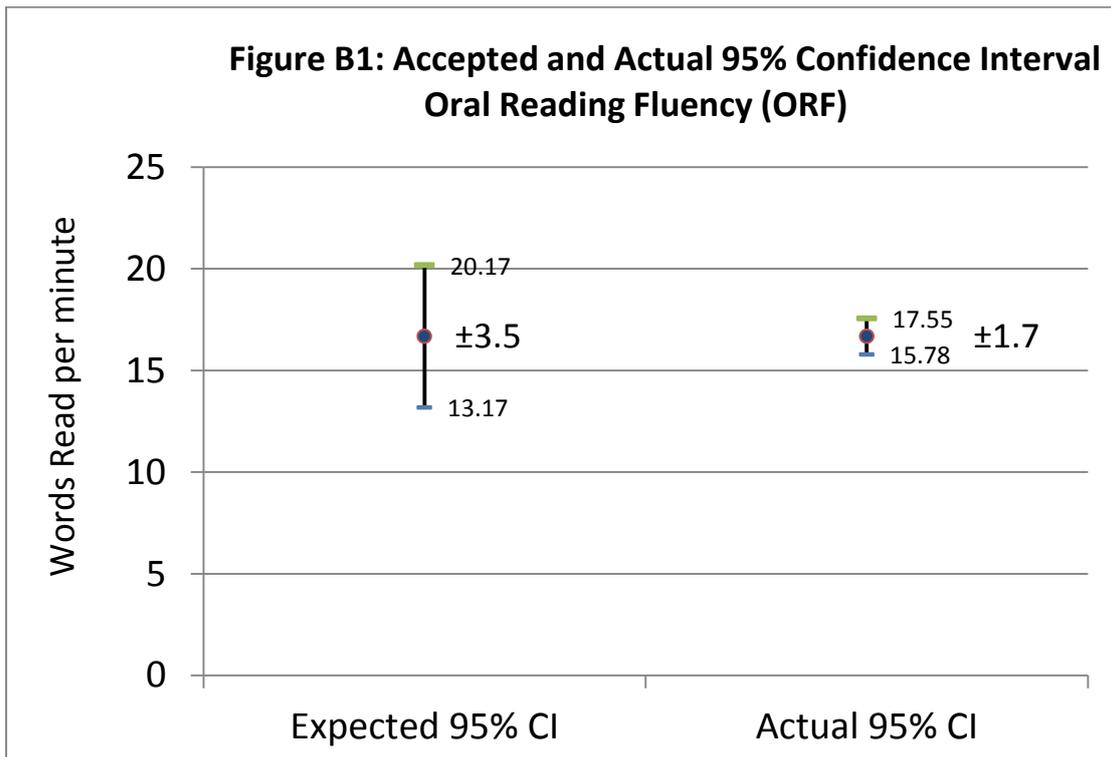


Table B5. EGRA means and 95% confidence intervals

EGRA Subtask	Grade 2 Average Score	95% CI Low End of Range	95% CI High End of Range	Range
Correct Letters Sounds Per Minute	35.98	34.39	37.57	3.18
Correct Familiar Words Per Minute	17.54	16.77	18.31	1.54
Correct Non-Words Per Minute	13.94	13.43	14.45	1.02
Oral Reading Fluency – 60 seconds with diacritics (cwpm).	16.67	15.78	17.55	1.77
Oral Reading Fluency – 90 seconds without diacritics.	24.93	23.75	26.10	1.17
Reading Comprehension – 60 seconds with diacritics	1.55	1.46	1.63	0.17
Reading Comprehension – 90 seconds without diacritics	2.19	2.10	2.29	0.10
Listening Comprehension # of correct answers (max 6)	3.28	3.19	3.36	0.17

Annex C: The Benchmarking Workshop – Agenda, Key Presentation, Proposed Benchmarks, and List of Participants

Proposed Agenda for the *Benchmarking for Early Grade Reading* workshop: Sunday-Monday, September 21-22, 2014 -- Ramallah, West Bank

Objectives:

1. Familiarize participants with *how* to use results of the Grade 2 EGRA to establish benchmarks (standards) for Grade 2 reading in MOEHE schools;
2. Familiarize participants with *other* information and objective inputs to inform the setting of benchmarks / standards for early grade reading in the West Bank;
3. Assist participants to reach consensus on *proposed*, measurable benchmarks for specific reading skills for early grades in MOEHE schools. Proposed, consensus benchmarks would be presented to the MOEHE at the conclusion of the workshop for internal MOEHE review and possible adoption.
4. Ensure that participants understand the process and uses of benchmarking and are technically able to reproduce results / reading benchmarks for early grades using EGRA or other objective sources.

Agenda – Sunday: September 21, 2014		
Time	Activity	Proposed Session Leaders / Participants
9:00 – 9:30	Arrival coffee and greeting of participants	
9:30 – 9:50	Welcoming remarks and Introduction of Participants Presentation of Workshop Objectives, Workshop Agenda and detailed Day 1 Agenda. Introduction of Workshop Presenters	MOEHE EGRA Steering Committee: Dr. Basri Saleh, Dr. Tharwat Zeid, Dr. Mohammad Matar Dr. Luis Crouch Dr. Robert LaTowsky
9:50 – 10:45	Session 1: Review of Key Tasks (Sub-Tests) of the Grade 2 WB EGRA. Review of WB EGRA Results on these Sub-tests.	Dr. Robert LaTowsky
10:45 – 11:15	Session 2: Review of benchmarking and its sources of information – <i>Part 1</i>	Dr. Luis Crouch
11:15 – 12:30	Session 3: Small group assignment – benchmarking and its uses. Short plenary discussion of small group outputs and key points. Q&A.	Participants work in small groups facilitated by Dr. Luis / Dr. Robert. Plenary review of small group work led by Dr. Luis.

12:30 – 13:00	Coffee Break with light snacks	
13:00 – 14:00	Session 4: Review of benchmarking and its sources of information – <i>Part 2</i> <i>Explain Session 5 tasks for work groups.</i>	Dr. Luis Crouch
14:00 – 15:00	Session 5: Small Working Groups implement 1 st exercise on benchmarking Grade 2 reading skills.	Participants work in small groups facilitated by Dr. Luis / Dr. Robert.
15:00 – 15:30	Plenary debrief and reflection on <u>1st</u> exercise results	Dr. Luis Crouch Dr. Robert LaTowsky
15:30 – 16:00	Summary Wrap-up of key Day 1 achievements	Dr. Luis Crouch Dr. Robert LaTowsky
16:00	Lunch and Departure	

Agenda – Monday: September 22, 2014		
Time	Activity	Proposed Session Leaders / Participants
9:00 – 9:30	Arrival coffee and greeting of participants	
9:30 – 9:45	Brief Review of Day 1 outputs and Day 2 Agenda	Dr. Luis Crouch Dr. Robert LaTowsky
9:45 – 10:30	Session 6: Small Working Groups implement <u>final</u> exercise on benchmarking Grade 2 reading skills.	Participants work in small groups facilitated by Dr. Luis / Dr. Robert
10:30 – 11:00	Plenary discussion of results of final exercise on benchmarking Grade 2 reading.	Dr. Luis Crouch Dr. Robert LaTowsky
11:00 – 11:45	Session 7: Presentation of additional information to benchmark reading skills in Grades 1 and 3.	Dr. Luis Crouch
11:45 – 12:30	Session 8: Small Working Groups implement <u>preliminary</u> exercise in benchmarking reading skills for Grades 1 & 3.	Participants work in small groups facilitated by Dr. Luis / Dr. Robert.
12:30 – 13:00	Coffee Break with light snacks	
13:00 – 13:30	Plenary discussion of <u>preliminary</u> results in benchmarking for Grades 1 and 3.	Dr. Luis Crouch Dr. Robert LaTowsky
13:30 – 14:30	Session 9: <u>Final</u> exercise in benchmarking for Grades 1 and 3.	Participants work in small groups facilitated by Dr. Luis / Dr. Robert.
14:30 – 15:30	Final review / discussion of results of benchmarking reading skills in Grades 1, 2 and 3. <u>How to achieve benchmarks?</u>	Dr. Luis Crouch Dr. Robert LaTowsky
15:30 – 16:00	Closing remarks Presentation of certificates	MOEHE EGRA Steering Committee: Dr. Basri Saleh, Dr. Tharwat Zeid, Dr. Mohammad Matar Dr. Luis Crouch Dr. Robert LaTowsky

Towards Possible Early Grade Reading Benchmarks in the West Bank

USAID | WEST BANK/GAZA

Education Data for Decision Making (EdData II):

Towards Possible Early Grade Reading Benchmarks for the West Bank

September 2014

Overview of presentation

- Definitions
- Why standards or benchmarks
- Sources of benchmarks
- Possible ideas for benchmarks



2

Recall this Recommendation from the Results Presentation:

- **Set and publicize reading standards (benchmarks) for the early grades.**
 - Establish expected levels of reading proficiency for specific subtasks, for each grade.
 - Train classroom teachers and primary grades in the expected level of proficiency for each grade.
 - Make frequent use of audio or video recordings of girls reading at the expected level of proficiency, so all stakeholders – teachers, supervisors, principals, parents, and students – clearly know the level of reading expected for each grade, and whether their students are reading at this level.

This presentation and your group work will show how you yourselves will do it -- at least in draft form.

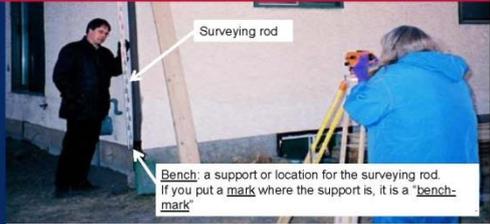
3

What is "bench-mark?"

- Relative standard of performance
- Origin (probably):
 - Bench
 - Mark
 - Mark for a "bench" on which to put a surveying rod so it'd be the same height each time
- Used now to mean any standard-setting
 - Benchmark errors in a factory: not more than 1/1000 defects
 - Benchmark performance of athletes: 11 seconds to run 100 meters
 - Benchmark skills (e.g., reading)

4

Bench-mark



Surveying rod

Bench: a support or location for the surveying rod. If you put a mark where the support is, it is a "bench-mark".

Photo by Gerd Mickelina



Towards Possible Early Grade Reading Benchmarks in the West Bank

A very old one. You could put the "bench" into those holes, and then hold the rod on the bench.



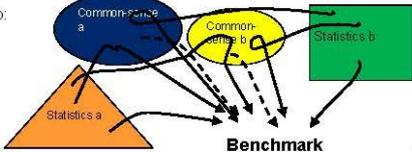
Photo: Statista

Why benchmarking?

- Most curricular statements are not specific enough.
- Curriculum: "child should read fluently" or "child should understand simple paragraph"
- But
 - Using what metric? Correct words per minute? Percentage correct answers to simple questions?
 - And, what are the benchmarks? 80% of simple questions correct? 60%? 20 correct words per minute? 120?
- Benchmarks give specificity to the curriculum.
- Create clear expectations.
- Provide assistance to teachers/coaches.

Bench-marking: science or art ?

- Both
- Statistics + common sense
- Science + wisdom
- And: not just one type of statistics, not just one group's wisdom
- So:



International trends

- As world moves towards meeting goal of 100% access to education, the United Nations, governments, and other bodies ask:
- "The children are in school, but are they learning?"
- As a result, a movement has arisen to set benchmarks of learning. Benchmarks of access to school are easy: are children there? Benchmarks of learning are very hard. This is partly why so much emphasis on access: it is easy to measure.
- A few countries are leaders. In the Arab world West Bank and Egypt could take a lead.
- In this discussion at international level there is a specialized vocabulary. We will introduce that vocabulary in the next slide.
- Not everyone uses these words in exactly the same way (but fairly similarly), but within the context of one presentation or one paper it is important to make the meaning totally clear.

Important distinctions - 1

- **Goal**
- **Metric**
- **Benchmark**
- **Indicator**

- Goal is an long-term aspiration, maybe without numerical value
- Metric is a valid, reliable unit of measurement
- Benchmark is a numerical step towards the goal, using the metric
- Indicator is a variable using the benchmark

- **Goal**: All our children should read
- **Metric**: "correct words per minute in passage reading"
- **Benchmark**: 45 correct words per minute, understand 80% of what they read
- **Indicator**: % of children at or above benchmark, or average achieved by the children, using the metric.

Important distinctions - 2

- Also, it is good to specify the metric as well as you can: "correct words per minute in connected text, of level of difficulty x" – and in that case the benchmark needs to "inherit" that specificity.
- In this case it is important to specify well the level of difficulty of the text, this can be done by judging the length of the words, the length of the sentences, how difficult the words are. Some of this can be done by formula (there are formulas for this), some by expert judgment.

Towards Possible Early Grade Reading Benchmarks in the West Bank

Uses of benchmarking in early grade reading

- Teachers themselves
 - Inspectors and coaches
 - Select / decide who to help more
 - Reporting to communities, parents
 - Report cards but with an objective bench-mark
 - Reporting to national level
- The kinds of metrics and benchmarks in this presentation are not recommended for high-stakes situations (e.g., holding children back, bureaucratically and strongly determining teacher rewards, etc.).

13

Group Exercise – Purposes and Uses of Benchmarks

- ✓ What are your ideas for how to use the benchmarks?
- ✓ For teachers to use in classroom?
- ✓ To report to parents whether their children's achievement is coming closer to benchmark?
- ✓ For teachers and schools to report to the Ministry?
- ✓ For the Ministry to report to society?
- ✓ Would you use them for rewarding and pressuring?
- ✓ What are some of the potentials and dangers?

14

Where can benchmarks come from?

1. Evidence from brain science
2. Evidence from other countries' standards
3. Evidence from West Bank itself
 1. Distribution of where you are now
 2. Note the correlations between skills
4. What kind of increases have been seen elsewhere?

15

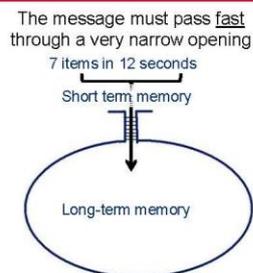
Insights from brain science*

- There is a paradox about the brain
- Long-term memory is nearly infinite
 - But short-term memory is very limited
- To understand a sentence we must read it within the deadline of our working (short-term) memory
- The sentence must be understood before it is processed by short-term memory
- If child reads too slowly, forgets the beginning before gets to the end
- Then, does not understand, cannot process it, and cannot penetrate into long-term memory

*This and next 3 slides benefit from insights of H.Abadzi

16

Insights from brain science



Insights from brain science: Minimum reading speed benchmark

- A sentence of about 7 words read in about 12 seconds gives roughly:
- One word per 1 – 1.5 second
- 45-60 words per minute
- And must be done automatically, without effort
- The mechanics of reading must become fluid and automatic so the brain can concentrate, while reading, on the meaning of what is being read.

Towards Possible Early Grade Reading Benchmarks in the West Bank

In setting benchmarks: consider where you are now: averages and % at 0

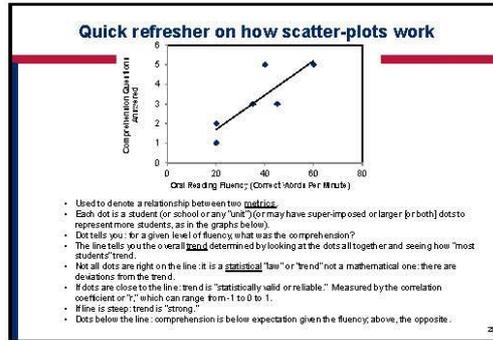
Grade 2		
Subtask	Percentage of students with zero scores	Grade 2 average score
Letter Sound Identification (lpm)	6.8%	36.0
Familiar Word Reading (fwp)	10.0%	17.5
Nonword Reading (nwp)	8.7%	13.9
Oral Reading Fluency – 60 seconds with diacritics (owp)	22.1%	18.7
Oral Reading Fluency – 90 seconds without diacritics (owps)	10.9%	24.9
Reading Comprehension – after 60 seconds with diacritics (max. 6)	35.7%	1.6 (out of 6 or 27%)
Reading Comprehension – after 90 seconds without diacritics (max. 6)	25.9%	2.2 (out of 6 or 37%)

In setting benchmarks: consider where you are now: look at the full distribution

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Letter Sound Identification (lpm)	35.9	4	11	18	26	37	46	51	58	67	100
Familiar Word Reading (fwp)	17.6	0	4	8	12	15	19	24	29	38	79
Nonword Reading (nwp)	13.9	1	5	8	11	13	16	19	22	27	66
Oral Reading Fluency – 60 seconds with diacritics (owp)	16.7	0	0	5	10	14	18	23	28	39	90
Oral Reading Fluency – 90 seconds without diacritics (owps)	24.9	0	6	11	16	21	27	33	41	54	127
Reading Comprehension – after 60 seconds with diacritics (max. 6)	1.55	0	0	0	1	1	2	2	3	4	6
Reading Comprehension – after 90 seconds without diacritics (max. 6)	2.2	0	0	1	1	2	3	3	4	5	6

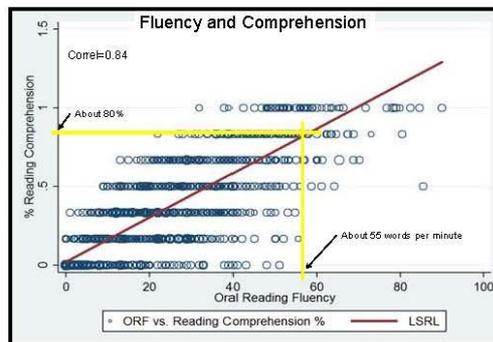
Look at evidence from West Bank itself

- Now some correlations

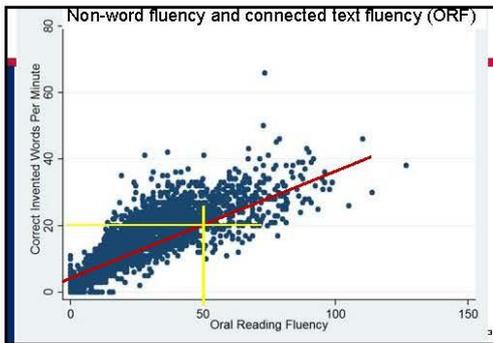
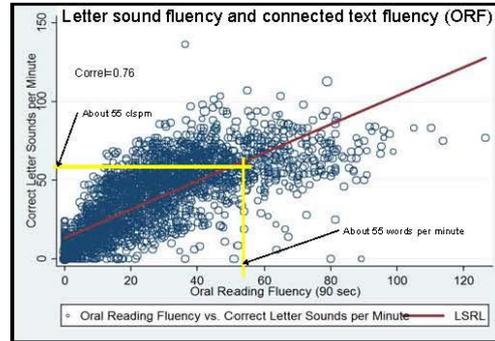
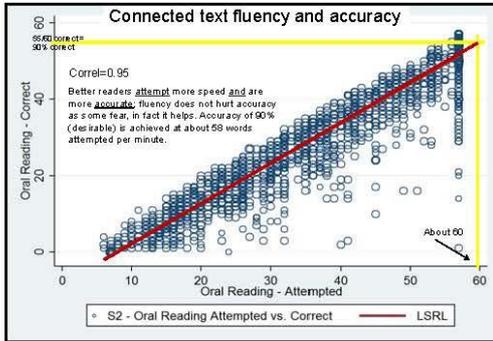


Let's do our own scatter graph

	CLSPM	ORF
Child 1	24	15
Child 2	56	50
Child 3	42	16
Child 4	0	0
Child 5	57	19
Child 6	12	23
Child 7	14	8
Child 8	36	33
Child 9	0	1
Child 10	39	12
Child 11	57	20
Child 12	11	7
Child 13	39	16
Child 14	55	22
Child 15	70	41
Child 16	2	3
Child 17	62	25
Child 18	50	31
Child 19	7	7
Child 20	50	40
Child 21	55	13
Child 22	59	34
Child 23	0	0



Towards Possible Early Grade Reading Benchmarks in the West Bank



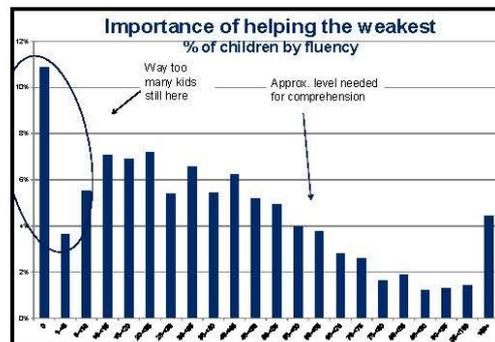
What can happen with a good pilot?

	ORF	Letter Sounds
2009	11	9
2011	21	22
	Nearly doubles	Grows by about 140%

*Egypt GIL0 Project

What is the message so far?

- About 80% comprehension is a "decent" conventional benchmark to shoot for in grade 2 or 3.
- 80% comprehension is correlated with ORF (connected text) of about 55-60 correct words per minute (cwpm).
- ORF at around 55 words per minute is associated with about 90% accuracy and that's the minimum needed to understand.
- ORF at around 55 correlates with about 55 correct letter sounds per minute (cisp), but note that developed country standards are lower for cisp.
- This is consistent with brain science.
- Some of these are lower than rich-country norms but those countries have more resources, longer experience, and no diglossia.



Towards Possible Early Grade Reading Benchmarks in the West Bank

Suggested place to begin

- Set desired bench-mark for:
 - Comprehension
 - Oral fluency in connected text (correct words per minute)
 - Letter sounds
 - Non-words
- Set desired % of children at or above the benchmark within 5 or 8 (? your decision) years
- Set maximum tolerable level of children scoring zero

37

Suggested place to begin – Grade 2

- Something like this. First choose a date in the future. 2020?
- The based on all the above information, choose benchmarks.

Grade 2	Comprehension	ORF	Letter sounds fluency	Non-word fluency
Benchmark:	80%?	45? 50? 55?	30? 40?	20? 30?
% at or above the benchmark	30%?	40%?	60%?	30%?
% scoring zero	15%?	0%?	5%??	10%??

- These are goals: goals are long-term, aspirational, and should be "full ambition"
- These are time-bound benchmarks to reach within X years (5-8 years? 8 years?)
- Are these the right skills?

38

Next steps? Questions?

- These will not be perfect; but it is a beginning.

- Any other questions?
- Now for group work



39

Worksheet for Group Work – Grade 2

- Suggested procedure: start with comprehension, reason back.
- Look at the correlation analysis, at the international evidence, and the current averages, % at zero, and distribution by percentile.
- It is easier to do the comprehension column first, then ORF, then letter sounds, then non-word fluency.

Grade 2	Comprehension	ORF	Letter sounds fluency	Non-word fluency
Benchmark:				
% at or above the benchmark				
% scoring zero				

40

Worksheet for Group Work – Grade 3

- Suggested procedure: start with comprehension, reason back.
- But this is more speculative until you have your own data.
- You can use some of the international evidence, but adapt.
- You can also look at distributional data from Grade 2:
 - Can you expect the "norm" for Grade 3 to be the top 10% in Grade 2? Think about it.

Grade 3	Comprehension	ORF	Letter sounds fluency	Non-word fluency
Benchmark:				
% at or above the benchmark				
% scoring zero				

41

Worksheet for Group Work – Grade 1

- Suggested procedure: start with comprehension, reason back.
- But this is more speculative until you have your own data.
- You can use some of the international evidence, but adapt.
- You can also look at distributional data from Grade 2:
 - Can you expect the "norm" for Grade 1 to be the bottom 20%? in Grade 2? Think about it.

Grade 1	Comprehension	ORF	Letter sounds fluency	Non-word fluency
Benchmark:				
% at or above the benchmark				
% scoring zero				

42

Towards Possible Early Grade Reading Benchmarks in the West Bank

Recommendations – How to achieve benchmarks

- How to **achieve** reading benchmarks?
 - Provide *pre-service* teacher training (at university) in reading instruction.
 - Teach new vocabulary in phrases and short sentences – not as isolated words – to promote reading connected text.
 - Promote reading practice in class. Encourage opportunities for students to practice reading paragraphs and short stories in school.
 - Add teaching strategies and classroom exercises that *directly strengthen comprehension*.
 - Use supplemental teaching and student practice materials for oral reading fluency and reading comprehension in addition to the Grade 2 textbooks.
 - Identify nonreaders early and provide reading coaching in class and after school.

Recommendations – How to achieve benchmarks

- How to **achieve** reading benchmarks? - *continued*
 - Identify early those who cannot read at all and give them additional instruction in school and out of school.
 - Raise awareness among lesser-educated parents of the importance of practice reading at home. Make practice reading materials (short stories and exercises) appropriate to Grades 2 and 3 readily available to all school children at low cost.
 - Have capable readers and “readers approaching benchmarks” practice silent and oral reading and reading comprehension exercises *with each other* in small groups in class while the teacher directs reading instruction chiefly to struggling readers.
 - Design and disseminate to all primary schools model programs and materials for one-week “reading camps” at mid-year and summer breaks for Grades 2 and 3 students. Be sure to include separate materials for:
 - i) each grade, and ii) struggling readers and readers approaching benchmarks in each grade.

Next steps? Questions?

- Any other questions?
- Thank You !



Outputs of Group Work – Benchmarking Workshop – Towards Proposed Benchmarks for Grade 2

10/12/2014

1st draft of small group benchmarks, Grade 2. ورقة عمل للعمل في المجموعة – الصف الثاني

الصف الثاني	الفهم	طلاقة القراءة الشفهية	التعرف على أصوات المقاطع	قراءة الكلمات غير المألوفة
العلامة المرجعية	- %٥٠ - %٦٧ %٤٠	٢٥ ٢٨ ٣٥	٤٥ ٤٠ ٤٥	١٧ - ١٨ - ٢٠
% عند أو فوق العلامة المرجعية	- %٤٠ - %٣٠ %٦٧	- %٢٤ - %٣٠ %٧٠	- %٥٥ - %٥٥ %٨٠	- %٥٥ - %٣٠ %٦٥
% نسبة يحصلون على الصفر	- %١٥ - %١٥ %١٥	- %١٠ - %١٠ %١٠	- %٣ - %١ %٥	- %٥ - %٥ %١٥
المهلة	٢٠١٨ - ٢٠١٨ - ٢٠١٨			

Revised small group benchmarks, Grade 2. ورقة عمل للعمل في المجموعة – الصف الثاني

الصف الثاني	الفهم	طلاقة القراءة الشفهية	التعرف على أصوات المقاطع	قراءة الكلمات غير المألوفة
العلامة المرجعية	- %٥٠ - %٥٠ %٤٠	٢٨ ٣٥ ٢٥	٤٠ ٤٥ ٤٥	١٨ ١٩ ١٧
% عند أو فوق العلامة المرجعية	- %٤٠ - %٤٠ %٥٥	- %٢٤ - %٣٠ %٥٥	- %٥٥ - %٥٥ %٦٠	- %٥٥ - %٤٠ %٥٥
% نسبة يحصلون على الصفر	- %١٥ - %١٥ %١٥	- %١٠ - %١٠ %١٠	- %٣ - %١ %٥	- %٥ - %٥ %٧
المهلة	٢٠١٨			

1

Proposed Grade 2 Benchmarks for 2018

Dr. Crouch reflection on revised small group benchmarks, Grade 2.

Grade 2	Reading Comprehension	Oral Reading Fluency	Letter Sound Knowledge	Nonword Reading
Benchmark score	50%, 50%, 40% 50% of questions answered correctly is a nice round number for reading comprehension benchmark (Grade 2: 60 second story <u>with</u> diacritics).	28, 35, 25 Average of above is ~30 correct words per minute. But without diacritics 30% already at 30 cwpm . With diacritics, 17% now read at 30 cwpm. Reasonable to set benchmark at 35 cwpm with 50% meeting (no diacritics), and 30 cwpm with 30% meeting (with diacritics).	45, 45, 40 45 correct letter sounds per minute (clspm) is nice round number and most common value (recommended by 2 of 3 working groups).	19, 18, 17 20 correct nonwords per minute (cnonwpm) is nice round number for a benchmark.
Percentage of students reading at or above <u>benchmark</u> score	40%, 40%, 55% Without diacritics, 40% of G2 students are already at 50% of questions correct. With diacritics, only 20% of students are at 50% or above. So perhaps 45% is a good target?	35%, 24%, 55% See above: 50% of Grade 2 students reading at or above benchmark of 30 cwpm <u>without</u> diacritics. Perhaps 30% <u>with</u> diacritics.	55%, 55%, 60% 40% of Grade 2 students are already reading 45 clspm. Target of 55% is reasonable.	40%, 55%, 55% About 30% of Grade 2 students now meet 20 cnonwpm benchmark. 50% or 55% reading at or above benchmark of 20 cnonwpm is reasonable target.
Percentage nonreaders (score zero)	15%, 15%, 15% Target of 15% of Grade 2 students with zero score on reading comprehension (60 second story) seems reasonable. Now 36%.	10%, 10%, 10% Perhaps 10% of Grade 2 students scoring zero <u>with</u> diacritics; 5% <u>without</u> diacritics. Currently 22% score zero with diacritics.	1%, 3%, 5% 1% seems ambitious. Now 7% of Grade 2 students score zero on letter sounds. Perhaps 3% is good; 5% if you prefer a round number.	5%, 5%, 7% Currently 9% of Grade 2 students score zero on nonword reading. 7% is not ambitious. 5% is a nice round number.

List of Participants, West Bank EGRA – Benchmarking Workshop

Participants in the Benchmarking and EGRA Analysis Workshops West Bank EGRA, Grade 2 - September 21-25, 2014

EGRA Analysis Workshop - Day 2		EGRA Analysis Workshop - Day 1	Benchmarking Workshop - Day 2	Benchmarking Workshop - Day 1	Internet	Cellphone	Place of Work مكان العمل	Participant Name الاسم	#
X	X	X	X	X	ahmedsaleh@moec.pna.ps	0599393848	وزارة التربية والتعليم العالي	أحمد صباح موسى حسان	1
X	X	X	X	X	wahed2040@yahoo.com	0598200062	دايرة القياس والتقويم - وزارة التربية والتعليم العالي	أ. وحيد قنينة	2
X	X	X	X	X	momatar66@yahoo.com	0599323895	دايرة القياس والتقويم - وزارة التربية والتعليم العالي	د. محمد مطر	3
X	X	X	X	X	khabsharat@yahoo.com	0599787895	دايرة القياس والتقويم - وزارة التربية والتعليم العالي	أ. خالد بشارات	4
X	X	X	X	X	k.thiab@yahoo.com	0599358502	دايرة القياس والتقويم - وزارة التربية والتعليم العالي	أ. خالد ذياب	5
X	X	X	X	X	ahed_aywashi@hotmail.com	0598302194	وزارة التربية والتعليم العالي	أ. عاهد عياش	6
X	X	X	X	X	msbleh@pctbs.gov.ps	0599765150	دايرة إحصاءات - وزارة التربية والتعليم العالي	أ. ماهر صبيح	7
X	X	X	X	X	muniid_2000@yahoo.com	0599944690	دايرة إحصاءات - وزارة التربية والتعليم العالي	أ. منجد فضل سليمان	8
X	X	X	X	X	aheneiti@tamerinst.org	0597911015	مؤسسة تامر	أ. أحمد حنيطي	9
X	X	X	X	X	ikhalli@unrwa.org	0595392789	unrwa - الأونروا	أ. عصام أبو خليل	10
X	X	X	X	X	qutohsamar@gmail.com	0595778138	مؤسسة تامر	أ. سمر القطب	11
X	X	X	X	X	umansour2@unrwa.org	0568730769	unrwa - الأونروا	أ. تينيل منصور	12
X	X	X	X	X	umansour@unrwa.org	0569059355	unrwa - الأونروا	أ. نظير منصور	13
X	X	X	X	X	sassaf@amideast.org		AMIDEAST	د. سهام عساف	14
X	X	X	X	X	aanbankaat@hotmail.com	0597950811	مشارف مركزية - وزارة التربية والتعليم العالي	أ. سهام سلمة	15
X	X	X	X	X	munashalohub@gmail.com	0599225798	مشارف مركزية - المعهد الوطني - NSET	أ. منى محمد إبراهيم تلهوب	16
X	X	X	X	X	omararwan@yahoo.com	0599293877	دايرة الأبحاث - وزارة التربية والتعليم العالي	د. عمر عطوان	17
X	X	X	X	X	gh.hanar66@hotmail.com	0598205520	مشارف مركزية - المعهد الوطني - NSET	د. حنان أحمد خليل الجمل	18
X	X	X	X	X	sarazahran80@hotmail.com	0599035329	مؤسسة تامر	أ. سارة زهران	19
X	X	X	X	X	waleedhsheish@yahoo.com	0599838361	مركز تطوير - وزارة التربية والتعليم العالي	أ. وليد احتشيش	20
X	X	X	X	X	ayvoub@naiah.edu	0599701982	جامعة النجاح	د. عبد الكريم أيوب	21
X	X	X	X	X	akhateebzoly@gmail.com	0598040652	دايرة الأبحاث / المناهج - وزارة التربية والتعليم العالي	د. أحمد الخطيب	22
X	X	X	X	X	tamer@palnet.com	0599255231	مؤسسة تامر	أ. رناد الفتيح	23
X	X	X	X	X	aaabdelrazeq@birzeit.edu	0597339142	مؤسسة تامر	أ. جوانا ريفيدي	24
X	X	X	X	X			جامعة بيرزيت	د. أنور عبد الغرزوقي	25

Annex D: Presentation of West Bank EGRA Results - May 25, 2014

West Bank Grade 2 Early Grade Reading Assessment Baseline



Education Data for Decision Making (EdData II):

**West Bank Grade 2
Early Grade Reading Assessment Baseline**

May 2014

Overview of presentation

- Background Brief
- Major Findings
- Other Findings
- Recommendations



Background Brief – Purposes / Uses of West Bank EGRA

- Purposes and Uses of the West Bank EGRA Grade 2 baseline:
 - Inform policy decisions and planning by the Palestinian Authority (PA) for improved reading instruction and student learning outcomes.
 - Infuse EGRA findings into the design and development of MOE curricula and teaching resources for enhanced reading instruction in Grades 1-3.
 - Establish a national baseline of Grade 2 reading skills to measure future progress in enhanced reading performance in MOE schools.
 - Strengthen the capacities and deepen the knowledge base of MOE staff to implement EGRAs.
- This is the first baseline assessment of early grade reading skills representative of all Grade 2 students in MOE schools of the West Bank.

Background Brief – The Sample and Technical Details

- A nationally-representative sample and scientifically rigorous implementation to assess Grade 2 reading skills:
 - 150 MOE primary schools with 10+ Grade 2 students randomly selected from all 16 West Bank *muderiya*s, in proportion to each *muderiya*'s share of all WB schools.
 - 2953 students: typically 20 students in each school randomly pre-selected from school enrollment lists before arrival at the school.
 - Candidate MOE teachers (assistant teachers) for primary grades thoroughly trained as assessors and deployed in 16 teams of 2 assessors each. Each assessor team led by an MOE Arabic supervisor.
 - The 150 sample schools were stratified into equal numbers of co-ed, boys' and girls' primary schools: 50 schools each.
 - Assessment conducted in the second half of March 2014.

Background Brief – The Assessment Subtasks

- The choice and design of specific subtasks (subtests) for this Grade 2 assessment incorporated the standard subtasks of previous Arabic EGRAs *plus* two subtasks requested by the MOE.
- The subtasks included in the West Bank EGRA Baseline for Grade 2:
 1. Timed Test of Letter Sound Knowledge – 1 minute
 2. Timed Test of Familiar Word Reading – 1 minute
 3. Timed Test of Nonword / Inverted Word Reading – 1 minute
 4. Timed Test of Oral Reading Fluency – 60 seconds with diacritics (standard)
 5. Test of Reading Comprehension: 60 seconds with diacritics – 6 questions
 6. 2nd Timed Test of Oral Reading Fluency – 90 seconds without diacritics
 7. 2nd Test of Reading Comprehension: 90 seconds without diacritics – 6 questions
 8. Test of Listening Comprehension – 8 questions

Background Brief – The Assessment Subtasks

- The Letter Sound Knowledge subtask

ل	ك	س
ب	من	ع
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز	هـ
ث	ت	د
هـ	ز	ش
ط	ظ	ف
ق	ي	أ
ب	ن	ط
س	د	ع
ك	ل	م
ن	ي	ق
ف	ظ	ط
ش	ز</	

West Bank Grade 2 Early Grade Reading Assessment Baseline

Background Brief – The Assessment Subtasks

- The Familiar Word Reading subtask

صَف	أَبِي	بَاب		
قَانَار	لَلَّس	قِصَّتُهُ	مَا	حَدَّثَ
يَطْبِقُ	ذَهَبَ	الْحَلْوَى	اِبْتَسَمَ	تَرَضَى
فَكَّرَتْ	يُنْتَهِمُ	صَحْرَةٌ	قَالَ	فِي
مَرْزَعَتُهُ	الشَّعْبَ	خَفَّ	مَصْنَعٌ	عِنَايَةَ
يَنْقُدُ	أَبِي	مِنَ	بَاغَةَ	تَرَضَعَ

Background Brief – The Assessment Subtasks

- The Nonword Reading subtask

لَمَشَ	رَاكَ	لَمَكَ		
رَا	جَزَقَ	سَهْرَامِي	قَانَذَ	صَمَحَ
لِهَسَ	بِي	حَجَبَتْ	صَخْرَصَ	أَشْبَعَ
فَرَحَ	زَمَلَّ	ضَا	نَافَ	خَزَجَلَ
زَكَنَ	نَاكَ	ذَفَ	خَمَطَحَ	وَلَحَ
صَمَطَطَ	سَا	خَرِيضَ	رَجِبَ	لَمَزَجَ

Background Brief – The Assessment Subtasks

- The Oral Reading Fluency (passage reading) subtask – 60 seconds with diacritics

يَزرُورُ أطفَالِ مَدِينَةِ طُولُكْرَمِ الحَدِيثَةِ كُلَّ يَوْمٍ جُمُعَةً،
وَتَلْعَبُ خَنِينٌ عَلَى أَرْجُوخَتِهَا المَعْلُوقَةِ عَلَى الشَّجَرَةِ، وَيَطِيرُ سَامِي طَيَارَتَهُ
الْوَرَقِيَّةَ فِي السَّمَاءِ عَالِيًا. ذَاتَ يَوْمٍ هَبَّتْ عاصِفَةٌ قَوِيَّةٌ، فَامْتَلَأَتْ أَرْضُ
الحَدِيثَةِ بِالأُتْرَاقِ وَالْعُبَارِ، وَلَمْ تَعُدْ طَيَارَةُ سَامِي تَعَلُّو فِي السَّمَاءِ كَمَا كَانَتْ.
وَوَجَدَتْ خَنِينٌ أَرْجُوخَتَهَا مَحْشُورَةً، فَطَلَبَتْ خَنِينٌ مِنْ سَامِي التَّفَكُّيرَ فِي
طَرِيقَةٍ لِإِعَادَةِ تَنْظِيفِ الحَدِيثَةِ وَإِصْلَاحِ أَرْجُوخَتِهَا.

Background Brief – The Assessment Subtasks

- The Oral Reading Fluency (passage reading) subtask – 90 seconds without diacritics

تعيش أم حمدان وأبناؤها في قرية صغيرة ،
ولها خمس دجاجات ، وفي كل صباح ، تفتح باب الخم ؛
ليأكل الدجاج الحب والعشب ، وفي أحد الأيام اخفت دجاجة سوداء ،
فبحثت عنها أم حمدان ولم تجدها، وبعد ثلاثة أسابيع إذ بالدجاجة
السوداء تخرج من كومة حطب ، ومعها عشرة فراخ صغيرة، قال حمدان
بصوت عال: أمي! الدجاجة عادت! طارت أم حمدان فرحاً.

Background Brief – The Assessment Subtasks

- Letter Sound and Non-word Reading subtasks test foundational reading skills taught in Grades 1 and 2 that are essential for reading fluency and comprehension.
- The Oral Reading Fluency (Passage Reading) subtasks test reader fluency and precision in reading familiar words – with and without diacritics – in story context.
- Three (3) comprehension tests – 2 Reading and 1 Listening Comprehension – measure student understanding of what they read. Learning to read for comprehension is an emergent objective of Grade 2 reading instruction.

Major Findings



West Bank Grade 2 Early Grade Reading Assessment Baseline

Major Findings – Introduction

- The different measures of EGRA results:
 - Average scores for each subtask.
 - Percentage zero scores.
 - Percentage of students reading at or above benchmarks.
 - Range of student performance.

Major Findings – Applying Provisional Benchmarks

- Proposed, *provisional* benchmarks applied to EGRA results
 - To make the EGRA results more meaningful, to initiate MOE reflection on reading benchmarks, and anticipating the question “*What percentage of Grade 2 students now read well?*” this presentation applied proposed, *provisional* benchmarks for Grade 2 to each EGRA subtask.
 - Provisional benchmarks for letter, word and connected-text reading subtasks were set at the performance level of the **top 33%** of students.
 - Provisional benchmarks for the Comprehension subtasks were set at 2/3 of questions (67%) answered correctly. 4 of 6 questions for each of the reading and listening comprehension subtasks. The comprehension benchmark for higher grades is typically 80% of questions answered correctly.
 - Setting Palestinian benchmarks for EGR is an MOE responsibility.

Major Findings – Key Conclusions

➤ **Strong pre-reading and listening comprehension skills.**
Average scores of Grade 2 students on the pre-reading subtasks and listening comprehension are strong. And the percentage of students with zero scores are low. Best results of all Arabic EGRAs (Grades 2-3) to date on these subtasks.

Table 1: Summary scores for pre-reading and listening comprehension subtasks

Subtask	Percentage of students with zero scores	Grade 2 average score	Proposed benchmark	Percentage of students performing at or above benchmark
Letter Sound Identification (c/pra)	8.8%	36.0	50	33%
Familiar Word Reading (wprt)	10.0%	17.5	23	32%
Nonword Reading (ncwprt)	8.7%	13.9	19	32%
Listening Comprehension (max. 6)	7.1%	3.3	4.0	50%

Major Findings – Key Conclusions

➤ **Weaker performance in oral reading fluency and reading comprehension.** Average Grade 2 scores on oral reading fluency and reading comprehension subtasks are weak. High percentages of zero scores; low percentages of students reading at the comprehension benchmark. Additional reading time does not greatly improve *comprehension*.

Table 2: Summary scores for oral reading fluency and reading comprehension subtasks

Subtask	Percentage of students with zero scores	Grade 2 average score	Proposed benchmark	Percentage of students performing at or above benchmark
Oral Reading Fluency – 60 seconds with diacritics (swrt)	22.1%	16.7	22	32%
Oral Reading Fluency – 90 seconds without diacritics (owrt)	10.9%	24.9	31	33%
Reading Comprehension – after 60 seconds with diacritics (max. 6)	35.7%	1.8	4.0	13%
Reading Comprehension – after 90 seconds without diacritics (max. 6)	25.9%	2.2	4.0	27%

Major Findings – Key Conclusions

➤ **Nonreaders few in pre-reading skills and listening comprehension; but many in oral reading fluency and reading comprehension.** Low percentage of functionally illiterate students in Grade 2 – but many students do not comprehend what they read.

Figure 1. Percentage Zero Scores for each subtask

Subtask	Percentage Zero Scores
Letter Sound Knowledge	7%
Familiar Word Reading	10%
Nonword Reading	9%
Oral Reading Fluency with Diacritics	22%
Oral Reading Fluency without Diacritics	11%
Reading Comp - 90 Sec	36%
Reading Comp - 60 Sec	26%
Listening Comprehension	7%

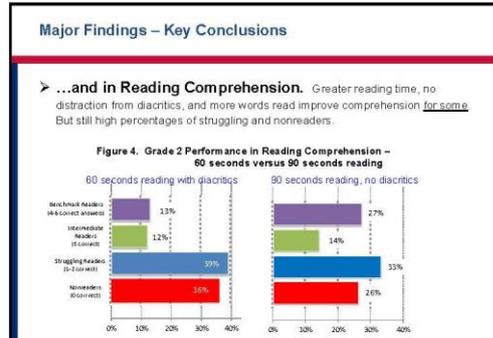
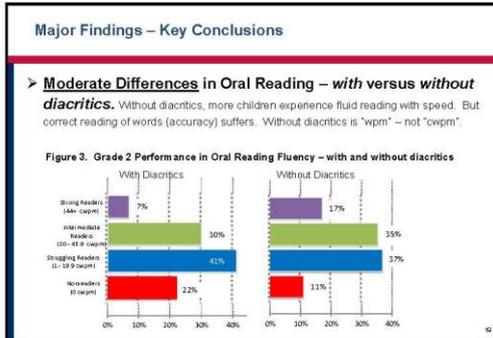
Major Findings – Key Conclusions

➤ **Many students mastering Grade 2 reading skills.** One-third or more of students now performing at or above meaningful benchmarks for most Grade 2 reading skills. But lower shares achieve proposed Grade 2 benchmarks for reading comprehension.

Figure 2. Percentage Grade 2 students reading at or above proposed, provisional benchmark levels

Subtask	Percentage at or above benchmark
Letter Sound Knowledge	33%
Familiar Word Reading	32%
Nonword Reading	32%
Oral Reading Fluency with Diacritics	32%
Oral Reading Fluency without Diacritics	33%
Reading Comp - 60 seconds	13%
Reading Comp - 90 seconds	27%
Listening Comprehension	50%

West Bank Grade 2 Early Grade Reading Assessment Baseline



Comparing Results with other Arabic EGRAs

➤ West Bank Grade 2 average scores in pre-reading skills and listening comprehension at the top of Arabic EGRA results.

- Highest average scores in letter sound knowledge and non-word reading.
- Results markedly higher than Jordan Grades 2-3 and Egypt Grade 3.

EGRA Subtask	West Bank 2014 Grade 2	Jordan 2012 Grade 2	Jordan 2012 Grade 3	Egypt 2013 Grade 3
Correct Letters Sounds Per Minute	36.0	26.5	26.3	18.8
Correct Familiar Words Per Minute	17.5	Not included	Not included	Not included
Correct Inventive Words (Non Words) Per Minute	13.9	4.4	7.0	5.9
Total correct Listening Comprehension questions	3.3	2.2	2.9	3.2

Comparing Results with other Arabic EGRAs

➤ **Lowest zero scores in pre-reading skills and listening comprehension of all Arabic EGRAs:**

EGRA Subtask	West Bank 2014 Grade 2	Jordan 2012 Grade 2	Jordan 2012 Grade 3	Egypt 2013 Grade 3
% Zero Scores on Letters Sounds	6.8%	21%	28%	18.3%
% Zero Scores on Familiar Words	10.0%	Not included	Not included	Not included
% Zero Scores on Non-Words	8.7%	4.9%	45%	27.4%
% Zero Scores on Listening Comprehension questions	7.1%	15%	8%	13.3%

Comparing Results with other Arabic EGRAs

➤ **But NOT superior in Oral Reading Fluency or Reading Comprehension:**

EGRA Subtask	West Bank 2014 Grade 2	Jordan 2012 Grade 2	Jordan 2012 Grade 3	Egypt 2013 Grade 3
% Zero Scores on Oral Reading Fluency – 60 seconds, full diacritics	22.1%	21%	20%	21.6%
% Zero Scores on Reading Comprehension questions	35.7%	27%	22%	35.4%
Correct words per minute Oral Reading Fluency – 60 seconds, full diacritics	16.7	16.2	23.7	21.9
Average # of correct questions Reading Comprehension 60 seconds, full diacritics	1.6	2.0	2.9	1.9

What are the implications of this?

- ### Recap of Major Findings
- Major Results of the Grade 2 EGRA Baseline:**
- ✓ Low zero scores: lowest of all Arabic EGRAs in pre-reading skills and listening comprehension.
 - ✓ Strong pre-reading and listening comprehension average scores.
 - ✓ Weaker performance in oral reading fluency (passage reading) and reading comprehension.
 - ✓ Reading without diacritics speeds reading and comprehension for some students. But nonreaders and struggling readers are still many.

West Bank Grade 2 Early Grade Reading Assessment Baseline



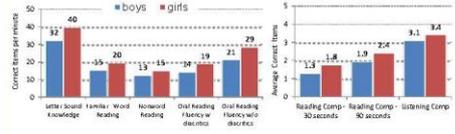

Other Findings



Other General Findings – Gender Differences in Reading

➤ **Gender differences in reading performance:** In the early grades, average scores for girls on reading subtasks are typically higher than boys. West Bank Grade 2 results confirm similar findings from Arabic and African EGRAs. But the gender gap is greater – except in listening comprehension.

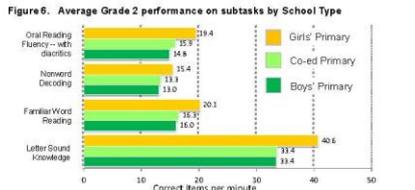
Figure 5. Gender performance on subtasks



Other General Findings – Gender Differences in Reading

➤ **Girls' Primary schools significantly outperform other School Types in Grade 2 reading.**

Figure 6. Average Grade 2 performance on subtasks by School Type



Other General Findings – Gender Differences in Reading

- Lower reading performance *for both boys and girls* in **co-ed schools**.

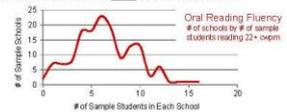
EGRA Subtask	Boys' Schools Grade 2	Boys in Co-Ed Schools Grade 2	Girls' Schools Grade 2	Girls in Co-Ed Schools Grade 2
Correct Letters Sounds Per Minute	33.4	29.3	40.6	37.4
Correct Familiar Words Per Minute	16.0	13.9	20.1	18.6
Correct Non Words Per Minute	13.0	11.6	15.4	14.8
Oral Reading Fluency – 60 seconds with diacritics (owgm)	14.8	13.3	19.4	17.9

- Percentage **zero scores** are also **higher** in co-ed schools.

Other General Findings – Differences in School Performance

➤ **Low divergence** among West Bank schools in the reading performance of their students.

- Average scores and % zero scores do not differ greatly among the 150 sample schools.
- West Bank MOE schools not divided into "weak" schools and "strong" schools.
- The average scores are "true" measures of central tendency. They reflect well the actual performance in most West Bank schools.



Other General Findings – 95% Confidence Intervals (CIs)

Finally... How reliable (replicable) are these results?

EGRA Subtask	Grade 2 Average Score	95% CI Low End of Range	95% CI High End of Range	Range
Correct Letters Sounds Per Minute	35.98	34.39	37.57	3.18
Correct Familiar Words Per Minute	17.54	16.77	18.31	1.54
Correct Non Words Per Minute	13.94	13.43	14.45	1.02
Oral Reading Fluency – 60 seconds with diacritics (owgm)	16.67	15.78	17.55	1.77
Reading Comprehension – 60 seconds with diacritics	1.55	1.46	1.63	0.17
Listening Comprehension # of correct answers	3.28	3.19	3.36	0.17

We are 95% certain that the true value is within this range for each subtask.

And this range is tight.

This is what is meant by the "95% Confidence Interval".

West Bank Grade 2 Early Grade Reading Assessment Baseline






Recommendations

Recommendations

- **Set and publicize reading standards (*benchmarks*) for the early grades.**
 - Establish expected levels of reading proficiency for specific subtasks, for each grade.
 - Train classroom teachers and Arabic supervisors for early primary grades in the expected levels of reading proficiency for each grade.
 - Make frequent use of audio or video recordings of boys and girls reading at the expected level of proficiency, so **all stakeholders** – *teachers, supervisors, principals, parents, and students* – clearly know the level of reading expected for each grade, and whether *their* students are reading at this level.

Recommendations

➤ **Priorities for enhanced Grade 2 reading as indicated by the EGRA results:**

- ✓ Significantly lower the percentage of **zero scores (nonreaders)** in oral reading fluency (with / without diacritics) and reading comprehension.
- ✓ **Increase average scores** in reading comprehension.
 - Apply teaching strategies and classroom exercises that **directly** strengthen *comprehension* – **both reading and listening** comprehension.

Recommendations – How to achieve benchmarks

- **How to achieve reading benchmarks?**
 - Provide *pre-service* teacher training (at university) **in reading instruction**.
 - Teach new vocabulary in phrases and short sentences – not as isolated words – **to promote reading connected text**.
 - Promote reading practice **in class**. Encourage opportunities for students to practice reading paragraphs and short stories in school.
 - Add teaching strategies and classroom exercises that **directly strengthen comprehension**.
 - Use supplemental teaching and student practice materials **for oral reading fluency and reading comprehension** *in addition* to the Grade 2 textbooks.
 - **Identify nonreaders early** and provide reading coaching in class and after school.

Recommendations – How to achieve benchmarks

➤ **How to achieve reading benchmarks? - continued**

- Have capable readers and “readers approaching benchmarks” practice silent and oral reading and reading comprehension exercises *with each other* in small groups in class while the teacher directs reading instruction chiefly to **struggling** readers.
- Design and disseminate to all primary schools model programs and materials for one-week “reading camps” at mid-year and summer breaks for **Grades 2 and 3** students. Be sure to include **separate materials** for: i) each grade, and ii) struggling readers and readers approaching benchmarks in each grade.
- Raise awareness among lesser-educated parents of the importance of **practice reading** at home. Make practice reading materials (short stories and exercises) appropriate to **Grades 2 and 3** readily available to all school children at low cost.



Thank you !

Robert J. LaTowsky, Ph.D. robert.latosky@gmail.com
 Luis A. Crouch, Ph.D. lcrouch@rti.org
 Statistical analysis by Susan Edwards, RTI