EdData II

Task Order 15: Data for Education Programming in Asia and the Middle East (DEP/AME)

Going to Scale: The Early Grade Reading Program In Egypt: 2008-2012

CASE STUDY REPORT April 4, 2013

EdData II Technical and Managerial Assistance, Task Number 15
Contract Number AID-EHC-E-00-04-00004
Task Number AID-OAA-BC-11-00001
April 4, 2014

This document was produced for consideration by the United States Agency for International Development. It was prepared by RTI International and Dean Nielsen, Consultant.

Task Order 15: Data for Education Programming in Asia and the Middle East (DEP/AME)

Going to Scale: The Early Grade Reading Program In Egypt: 2008-2012

CASE STUDY REPORT April 4, 2013

Prepared for
Christine Capacci-Carneal, Senior Education Advisor, Middle East Bureau
Mitch Kirby, Senior Education Advisor, Asia Bureau
Contracting Officer's Technical Representative
Data for Education Programming/Asia and Middle East
USAID / Washington
1300 Pennsylvania Avenue NW
Washington, DC 20523

Prepared by
RTI International
3040 Cornwallis Road
Post Office Box 12194
Research Triangle Park, NC 27709-2194

Table of Contents

Abbreviations

COTR	Contracting Officer's Technical Representative
DEP-ASIA/ME	Data for Education Programming in Asia and the Middle East
EGRA	Early Grade Reading Assessment
EGRP	Early Grade Reading Program
GILO	Girls' Improved Learning Outcomes Project, Egypt
GOE	Government of Egypt
MOE	Ministry of Education
NGO	Nongovernmental Organization
PAT	Professional Academy of Teachers
RTI	Research Triangle Institute International
SD	Standard Deviation
USAID	United States Agency for International Development

Preface

This Case Study has been commissioned by USAID Washington's Middle East Bureau through the Data for Education Programming in Asia and the Middle East (DEP-ASIA/ME) contract to help decision-makers and program managers understand how the Egypt Early Grade Reading Program (EGRP) unfolded between 2008 and 2012 and the context within which it operated. The study is not an evaluation – that has been done elsewhere. Instead it is designed to help partners better understand what took place and why, and, with this, be better able to design and implement similarly successful programs elsewhere.

The study relied on document review and field work in Egypt over a two-week period and is divided into six sections as follows: a) Piloting Early Grade Reading Assessment and Interventions; b) EGRP Scale-up and Institutionalization; c) Evidence of Scale-Up Success; d) Factors Contributing to and Constraining Success; e) Lessons Learned; and f) Possible Future Directions.

Field work for the study was conducted during November 2012 and the report draft in December 2012 and January 2013. Since then new steps have been taken by both the Ministry of Education (MOE) and USAID towards EGRP sustainability, including the conduct of a national Early Grade Reading Assessment (March 2013), the establishment of a permanent EGRP unit in the MOE to sustain the momentum of improved reading instruction, and the announcement of a new USAID project for continued support of early grade reading called *Core Skills and Organizational Reform in Education.* These initiatives, all mentioned in the Case Study as desirable future directions, are described in report footnotes.

The Case Study investigator is sincerely grateful for the cooperation and support he received during all phases of this review, from preparation to final drafting. Especially appreciated is the help he received in gathering key documents from the RTI home office technical manager in Washington, Michelle Ward-Brent, and in Cairo from the USAID Mission's Contracting Officer's Technical Manager (COTR) Hala El Serafy and the RTI/GILO project staff, headed by Barbara Toye-Walsh. Receptivity of key informants in the USAID mission, GILO, and the MOE to my requests for interviews made these information gathering events not only informative but pleasurable. Most of all, I am grateful for the gracious and enthusiastic welcome I received from office and school managers, teachers, teacher syndicates, community members, and students in the 13 schools I visited over four governorates. Never have I found such group discussions and classroom visits to be so stimulating and pleasurable. Key to the magic were the efforts of my translator/interpreter, Noha Haddad, whose tireless and lucid translations kept our conversations moving. She deserves much of the credit for this report.

Maplewood, New Jersey
March 31, 2013

Executive Summary

In 2008 the Girls' Improved Learning Outcomes (GILO) Project was launched in Egypt, funded by the U.S. Agency for International Development (USAID) and managed by an implementing contractor, the Research Triangle International (RTI International). The goals of the project were simple: to a) Expand girls' access to quality education in remote and deprived areas of Upper Egypt; b) Strengthen school management and governance; and c) Improve the quality of teaching in and learning in targeted schools and *idaras* (districts). Efforts to improve learning outcomes included many key determinants of student achievement, including enhanced early grade reading skills, a component which at first consisted mainly of piloting an Early Grade Reading Assessment (EGRA), but which eventually was expanded into an early grade reading program that reached virtually every primary school in the country.

EGRA 2009 and EGRP. The first EGRA in Egypt conducted in a set of 60 project and control schools in four Upper Egypt governorates in 2009, yielded results so jarring to authorities that the government, USAID, and GILO decided to develop a remedial program to help pilot schools and their teachers improve reading results. This intervention became known as the Early Grade Reading Program (EGRP). The EGRP was built cooperatively by RTI staff and consultants with Egyptian Arabic language specialists. It was based on RTI experiences in other countries, Egyptian specialists' contributions, and the results of both the EGRA and a 2010 assessment of the grade 1/2 language textbooks (which revealed, among other problems, an overly complex vocabulary set). EGRP was introduced in pilot school 1st and grade 2 classrooms in September 2010 through intensive teacher training and coaching methods.

EGRA 2011. In the spring of 2011 EGRA was administered in same pilot schools that participated in EGRA 2009 and the EGRP intervention, plus the 2009 control schools. This time it covered only second graders, whose results could be compared with those in the 2009 assessment. The results were dramatic. Before the EGRP intervention, students in the GILO pilot schools performed at roughly the same level as their peers in the control schools. After less than a full academic year of the EGRP, students in the EGRP pilot schools pulled away from their control group counterparts on all EGRA dimensions. For example, EGRP pilot school students identified 19 more letter sounds per minute at the end of the intervention period, an increase of 194% over baseline; whereas students in the control group gained just 2 letter sounds per minute, an increase of only 21% over baseline. In reading fluency, EGRP pilot school students could read, on average, a total of 10 more words per minute – an increase of 82% over baseline; whereas control school students only read 3 more, a 38% increase.

Scaling-up. With such results, and the positive feedback from teachers and managers, the Project partners moved quickly towards scaling up the EGRP intervention. GILO originally presented a design to scale-up gradually, with the first stage covering the remaining districts in its four pilot governorates. The national Ministry of Education (MOE) was determined to move ahead with a rapid nation-wide scale up. In the end, it was decided that Project would fund and technically support scale-up in the pilot governorates, and assist the MOE to scale up in the remaining 23 governorates (largely by training governorate planning teams and EGRP trainers/supervisors). Much of the training at the district level and all at the school level would be funded and implemented by the MOE through a cascade model of dissemination.

Institutionalization. At the same time that EGRP was being scaled-up, the Ministry also worked to safeguard sustainability through institutionalizing various features of the intervention, including learning materials (textbooks aligned with EGRP); EGRP teacher training and training materials (accredited by the Professional Academy of Teachers (PAT); PAT in-service training courses on EGRP (a requirement for new

teachers and accredited EGRP teacher trainers); various policies and ministerial decrees, e.g., those requiring 25 minutes a day for phonics and allowing teachers to coach in neighboring schools; and systems and tools for supervising EGRP instruction. The Ministry was also well on its way to institutionalizing the EGRA, in the sense that domestic testing specialists had acquired the capacity to administer it without external assistance.

Evidence of Scale-up Effectiveness. There is a widespread perception within and outside of Egypt that the scaling up of EGRP has been a success of unprecedented proportions. Given the fact that there has not yet been a national EGRA application, and little other systematic evidence has been collected, it is difficult to confirm these perceptions empirically. However, two recent GILO studies (2012) do shed some light on the subject; a telephone survey of scale-up implementation, and GILO's use of a rating scale -- along 23 dimensions of EGRP implementation -- to compare how effectively the 27 governorates have been in establishing the system.

The telephone survey was to show whether a school had become *operational* in EGRP by its receipt of EGRP materials. Of the 129 schools selected at random for the survey from all 23 non-GILO-supported governorates, about 90% had received at least 3 of the five EGRP resources materials. Given that distribution was done by cascade this shows impressive organization effectiveness in getting the system "out there."

The results of program implementation ratings were more nuanced. The dimensions that were rated (23 of them) were grouped into organizational levels: *muderiya (*governorate*)* planning team*, muderiya* cadre team*, idara* (district) cadre team, follow-up and supervision, and teachers. The results show relatively high ratings for program implementation at the higher levels (*muderiya*) and relatively low ratings at the lower levels (follow-up and supervision, teachers). These differences seem to illustrate the limitations of the cascade model, where quality becomes diluted as implementation moves from top to bottom, at least during the first year of the national scale-up: it is conceivable that in the second year (2012-2013) the cascade will work better. An analysis of the difference between high and low performing governorates, revealed big gaps in the area of training at all levels, particularly district, and school-level supervision. Overall, three of the five best ratings were found in governorates that *were not among those intensively supported by GILO* (North Sinai, Port Said, and Aswan), even at the district and school levels, showing how aggressively the program was taken up in some relatively remote areas.

Interviews. This Case Study's two week window for in-country data collection made possible the conduct of 20 key informant interviews (see Annex A – EGRP-Egypt Persons Interviewed), plus group interviews at all of 13 visited schools. The interviews revealed high levels of enthusiasm for the program at all levels, and virtually no criticism (see Textbox 1 for testimonials). All expressed amazement at how quickly the intervention spread nationally ("like wildfire" according to one informant). This occurred despite strong initial skepticism on the part of RTI, which expressed concern about the adequacy of existing resources. Many remarked on how determined the MOE's First Deputy Minister, Dr. Reda Abu Serie, was to making the needed resources available. In the end he was able to do so, partly by drawing upon volunteerism (dedicated people donating time) and local resources from governorates, districts, and schools (and their boards). *Finally, in virtually all group discussions, managers and teachers spoke of improved student learning outcomes: higher reading rates, lower dropout, improved school attendance. Unfortunately, however, up to now, there is no hard evidence for this.*

Direct observations. All of the 32 reading lessons observed were using EGRP routines and materials (in grade 1, as scripted) for teaching phonics, often with creative adaptations by the teachers themselves.

The use of student active learning processes was also observed in a majority of classrooms, but this was still difficult for many teachers. Student engagement in these lessons, even where active learning was not consistently applied, was at the rare level of 80-90%. Only one of the schools visited was a Project pilot school. The lessons observed there were clearly more polished, the teachers were more confident, and the components were better connected. This was undoubtedly a reflection EGRP being implemented there for two years rather than one as in the others. A quick test of reading skills among 2nd graders in a few schools (volunteers and those chosen at random) showed most students to be reading already -- some fluently, but most haltingly – an impressionistic result, but a sign of better than EGRA 2009 functioning . The best teaching observed during the Case Study was not at a Project pilot school, but in a low income area in Cairo and in a rural school in coastal Damietta. *Although this was not a representative sample (and the lowest performing governorates were not visited), it was varied enough in region and school quality to confirm earlier survey results that EGRP methods have been effectively implanted in the education system, despite the use of the cascade model*.

Factors contributing to scale-up success. The above successes were seen to be attributed, at least in part, to the following: 1) bipartisan endorsement of the learning model; 2) scientific underpinnings of EGRA/EGRP; 3) a well-led and skillful GILO team and effective RTI backstopping; 4) early MOE engagement and ownership; 5) a Ministry champion and a stable central leadership team ; 6) young energetic mentor teachers; 7) quick results; 8) teacher testimonials; 9) volunteerism and local resources; and 10) a skillful and politically adept USAID country team.

Factors constraining scale-up success. These were found to be: a) lack of national EGRA data,[1] b) lack of good empirical data on teacher performance, c) over-crowded classrooms, and d) the teachers' sense that they have not yet received enough training to reach mastery level in the system.

Lessons Learned: For the innovation managers and implementers. This Case Study has identified numerous lessons that have either already been articulated by key actors or have been constructed from the materials uncovered. An initial list includes the following:

1. *Science matters*: the cognitive scientific framework of EGRA gave it credibility in Egypt, and the rigor of the assessment, pre-intervention classroom observations, and textbook analysis furthered this credibility and demanded attention; however, there is potential long-term damage to the credibility when crucial pieces of the scientific foundation are left out (e.g., a national EGRA).[2]
2. *Bipartisan endorsement of the model:* multiple endorsement of the kind that was reached with EGRP in Egypt was a powerful motivator.
3. *True partnerships*: between the Project staff/subcontractors and their Egyptian counterparts generated energy.
4. *Innovations with simple steps*: the sequence of learning to read through phonics in EGRP caught on quickly.
5. *Early and easily demonstrable gains*: almost immediate improvement in reading skills; improved student engagement and attendance garnered attention and generated enthusiasm.
6. *Shifting from a teacher- to a student-centered pedagogy takes time: most observed teachers were able to adopt the phonics teaching approach to reading, but many did it in the traditional teacher-*

[1] Shortly after the Case Study mission a decision was made to fill this gap, by conducting a national EGRA at the third grade level in the spring of 2013. This will be touched upon in the main report.

[2] The recent conduct of a national EGRA helps to fill this gap (see previous footnote), but other measures of student and teacher performance are still needed.

centered manner, suggesting that changing teacher styles takes time and requires continuous professional support and coaching.

7. *Successful pilots can be brought to scale*: in Egypt this resulted from strong consensus on the need and way to change early grade reading; a moral-political mandate to improve reading for all; and the right kind of government ownership (facilitated by a wise contractor) and commitment, led by group of Program "champions," which remained throughout the pilot and dissemination period.
8. *Radical positive changes can occur "beneath the radar*": even during times of instability and conflict on the main stage major reforms can still be brought to fruition.
9. *Widespread scaling-up is possible*: even when funds are tight, government agencies can enhance system efficiency and thus scalability through increased volunteerism and local buy-ins -- dividends of a popular innovation; as the system moves towards institutionalization, nationally and locally, such contributions will be less necessary and difficult to sustain in any case.
10. *The cascade model of dissemination has both strengths and limitations*: without the model it clearly would not have been possible to disseminate EGRP nationally in such a short time; one consequence of its use, however, has been the appearance of regional disparities, as some localities are better able to manage the scaling-up challenges than others – a limitation that can be dealt with as long as it is recognized and proactively compensated for.

Lessons Learned: For Other Countries/USAID Support Programs. This detailed and systematic study of the Egyptian EGRP case has generated insights that might be helpful for other countries or the U.S. Agency for International Development as they consider expanding work in this field. Here are some concepts:

1. The scientific rigor of EGRA/EGRP is an asset that needs to be safeguarded as advocates, managers and implementers think about/work on cross- or within-country dissemination.
2. A viable model to consider for introducing and/or scaling up early grade reading assessments and interventions is the one followed by Egypt (not necessarily by design but more by default) and that is to implant a modestly funded program at the beginning and then, relying on skillful implementers/ contractors and local USAID managers, nurture its growth within and in response to the ever changing political/educational environment, moving briskly at times to seize opportunities, but also gradually growing local ownership and commitment, so when a chance does arise for scaling-up substantial political will and domestic funding will have materialized.
3. No single country, even one as successful as Egypt, can be used the prototype of how to scale-up an effective early grade reading program, given the unique combination of factors and forces in each country. However, countries should be able to learn lessons from one another for use within their own contexts.
4. An exciting and immediately successful learning system may be more effective in increasing school attendance than more conventional methods to increase access. Such methods, for example, using social promotions or family incentives may be able to bring reluctant children to school *but cannot hold them* if instruction is weak and little is learned. Experience with EGRP has shown (at least anecdotally) that where students enjoy learning and feel a sense of accomplishment attendance soars.
5. Macro-level instability and fragility is not always detrimental to the success of reform effects -- as long as there is a stable corps of visionary managers at the operational level, supportive and sensitive contractors and Agency managers, and some space to innovate under the radar. Once the government of former President Hosni Mubarak fell, even though chaos ensued in the wide political realm, the educational bureaucracy felt liberated and charged with a mandate to move into new realms of nation-building.

6. Effective training and real professional growth of teachers are the keys to creating improved student learning outcomes in low-performing educational systems. From the beginning of EGRP it was clear that in Egypt, as elsewhere, the educational system's main asset and best change agent was the classroom teacher.

Future Directions. Evidence was strong that the government is committed to covering the operational costs of continued EGRP functioning; however, development costs for continued EGRP model development may require continued external financial support. Needed further developments would include: a national EGRA, stronger data on teacher performance, a revised grade 2 textbook, EGRP routines/materials for grade 3, and eventual vertical and horizontal integration with other parts of the curriculum. Also, desirable would be more R&D on system efficiency (e.g. e-learning), continued use of monitoring and evaluation (M&E) for system improvement, and better linkages with higher education for pre-service and further EGR research and experimentation.

The Early Grade Reading Program (EGRP) in Egypt: 2008-2012

A. Piloting Early Grade Reading Assessment and Interventions

a.1 ***Labor force.*** The last decade of the Mubarak regime saw Egypt actively engaging in the global marketplace through, among other things, increased foreign investments and the establishment of knowledge-based enterprises. The manpower demands of these enterprises, including high literacy levels, has often been beyond what the Egyptian labor force could supply, frequently requiring firms to recruit workers from abroad. This has caused policy makers to take a hard look at the quality of the education system at all levels, from basic to tertiary.

a.2 ***USAID programs.*** During the last Mubarak era the government provided little funding in support of the reform strategies promoted by the Ministry of Education, including those to improve educational quality. For this reason the Ministry often turned to development agencies to help cover new development costs. In the education sector, one of the most important outside contributors has been USAID, especially in the area of basic education where it became the country's main external donor. USAID's support to Egyptian basic education early in the 21st century was focused on the country's quality improvement agenda. This occurred initially through improved inputs, including school buildings and equipment, training and professional development for teachers and administrators, community involvement in planning and management (through the New School Program and the Alexandria Reform Pilot), and library books (the USAID-supported National Book Program provided over 24 million books to public schools all across Egypt). By the middle of the Century's first decade, USAID was introducing tools and approaching for measuring improvements in teaching (SCOPE) and for assessing student learning outcomes (CAP). These efforts laid the ground work for the most recent round of USAID education projects, including the focal point of this Case Study; the *Girls' Improved Learning Outcomes (GILO)* project (funded 2008-2011 and extended to 2013). With improving learning outcomes as their main concern, GILO managers and overseers agreed to the urgings of its contractor, RTI International, to introduce a modest $40,000 oral reading assessment component into the Project, which, if accepted by Egyptian counterparts, could be adapted for use with the Arabic language. This was done, and already in 2008 the instrument was piloted in Cairo schools as a way of assessing early reading skills. This assessment instrument became known in Egypt and other locations around the world where it has been adapted and supported by USAID, the World Bank and other agencies, as the *Early Grade Reading Assessment* (EGRA).

a.3. ***EGRA in Egypt.*** The Arabic language EGRA instrument in Egypt was adapted from the standard template developed by RTI International with the support of the U.S. Agency for International Development (USAID) and the World Bank for use throughout the developing world.[3] The local adaptation of the EGRA instrument into Arabic for piloting in Egypt was based

[3] A detailed history of the instrument, its components, and the conceptual framework underpinning the assessment can be found in RTI, *Early Grade Reading Assessment Toolkit*. Prepared by RTI for The World Bank, Office of Human Development, 2009.

largely upon inputs from Egypt Arabic language specialists.[4] The skills assessed by the instrument in Egypt, as in most other countries using the instrument, were: a) letter naming fluency; b) syllable reading fluency; c) word reading fluency; d) unfamiliar word reading fluency; e) oral reading fluency, and f) reading comprehension.

a.4. Following completion of the initial instrument adaptation in 2008, agreement was reached to integrate EGRA more fully into the Project. The first step in this integration was to pilot the instrument in the GILO project areas in 2009 – an application that is now referred to as EGRA 2009. This piloting took place in 29 Project-supported schools from 11 districts of Fayoum, Minya, and Qena governorates, and in a sample of 30 control schools from non-project districts in the same three governorates having characteristics similar in many ways to the pilot ones.[5] Students from grades 2, 3 and 4 in the pilot and control schools were selected at random from class lists with girls and boys being represented equally. Data collection was conducted by Project-trained enumerators, recruited from NGOs and local universities, who tested students one at a time by showing them stimulus materials and recording their responses (a process taking about 30 minutes per student). The cost of the assessment, including fees, travel, accommodations, and breaks for enumerators, and testing materials, came to about $9 per student. It was conducted over the three-week period in January and February 2009 at the rate of approximately 5 schools per day.

a.5. ***Wake-up Call***. The results of the 2009 EGRA were sobering to the Egyptian educational community. Not only were most tested second graders in both pilot and control schools slow in recognizing syllables and reading words, but about half in both settings *could not recognize any syllables or read any words at all* (see Tables 2 and 3 in Annex E). These results, when rolled out to a gathering of education policymakers, created a crisis atmosphere. This galvanized a strong commitment to what was first thought of as "remediation," but what eventually became system/policy change in the teaching of reading. In the words of one informant, "EGRA was our wake-up call."

a.6. Not all parties were initially convinced of the need for change. For example, many curriculum specialists were confident that the existing Arabic reading program was well designed and appropriate. They viewed the EGRA 2009 results more as a reflection of weaknesses in Upper Egypt, where EGRA was piloted, than a portrayal of weak early grade reading skills throughout the nation. Also, there was a strong commitment to the official textbooks, which in Egypt were the drivers of instruction. Project managers, confident in the wider applicability of EGRA 2009 findings, eventually convinced curriculum specialists to join

[4] At first there was resistance to a US agency leading out in the creation of an Arabic language reading test, with voices asking, "Why are Americans here to teach us about Arabic?" But eventually it became clear that GILO was there not to teach Arabic but to introduce a globally-validated assessment instrument used in dozens of countries with a wide range of languages, now including Arabic.

[5] It should be pointed out that this was not an experimental design; since the pilot districts were not chosen at random, but purposely, based on relatively low girls' participation and attainment rates (improving these was one of the initial goals of the Project). By selecting control districts/schools of similar characteristics, and then selecting students at random in both pilot and control schools, the Project created a semi-experimental design, which allowed for valid comparisons of outcomes.

them in systematically observing reading lessons in a diverse sample of schools and the results revealed even to the skeptics how flawed and neglected reading instruction had become. In addition, the early reading textbooks came under scrutiny, at first informally and then, in 2010, through a systematic content analysis.[6] The findings for grade 1 revealed significant stumbling blocks for early grade readers. For example, of the 925 words introduced, 435 were only used once; the vocabulary included words up to six syllables long; many words were unfamiliar or of little relevance to children; and the books used small type font. This was an important turning point for the Program in Egypt: two new sets of scientific evidence showing the need for change contributed to a growing consensus on not only the need for reform but the best direction for it. GILO had already initiated work on improving learning outcomes: the shock waves from EGRA 2009 and the growing consensus about the need and direction for change prompted the Project to shift work on an early grade reading intervention into high gear.

a.7. ***The Early Grade Reading Intervention/Program (EGRP).*** Encouraged by a new level of consensus with Egyptian reading specialists, GILO early reading consultants and sub-contractors, led by world-class early reading specialist, Dr. Sylvia Linan-Thompson, went to work on drafting new instructional strategies, using a consultative and iterative approach. After discussions with local experts, the Project would draft segments of the intervention. These were then submitted to MOE and Arabic language specialists for review with the question, "will this work?" Reaching agreement would sometimes involve several iterations or week-long workshops. Small-scale trials in schools were also used, sometimes observed by Cairo's senior MOE leadership, including First Deputy Minister, Dr. Reda Abu Serie. The intervention/Program (EGRP) that took shape by the second half of 2010 included the following sections: *phonics, phonemic awareness, word lists, decodable sentences and stories, and comprehension routines.*

a.8. Teacher and student materials also emerged in stages: at first most materials were trainer/teacher made; but gradually a teachers' manual, flip charts, CD containing resource materials, and some student worksheets were produced by the Project (although teachers were still encouraged to create some of their own). EGRP was piloted in Project-supported schools in districts in four Upper Egypt governorates (the three that participated in EGRA 2009 plus Beni Suef). Intensive training of teachers took place during the school break in 2010; program implementation in both grades 1 and 2 started in October of that year, and continued through most of the school year, although the Arab Spring social/political instability forced the closing of schools in some locations during some weeks in early 2011.

a.9. ***Follow-up Assessment: EGRA 2011.*** EGRA 2011 was administered in the same schools where EGRA 2009 was conducted, but only among second graders (first graders were considered too young for rigorous assessment). Originally scheduled for January-February (as in the case of EGRA 2009), the assessment took place as soon as the calming of the Arab Spring crisis allowed (April-May, 2011). The test and processes were the same as for EGRA 2009, except that this time

[6] A study group was formed to carry out the textbook analysis. The group consisted of Egypt MOE and Project staff along with an Arabic language textbook expert from Ain Shams University in Cairo.

half of the enumerators were recruited from local MOE offices. This reduced the cost of the assessment by almost half (see Annex C. Main Features of EGRA 2009 and 2011 for details).[7]

a.10. ***Stunning Results.*** The results of EGRA 2011 were stunning. Even though the intervention had only been running for a few months, pilot school children scored vastly better that their control group counterparts (for details, see Tables 1-4 in Annex E). For example, on Syllable Reading, average scores in intervention schools increased from 9.8 in 2009 to 28.5 in 2011, a change of 192%; whereas those in control schools only increased from 8.6 to 10.1, a change of 18%. Concerning Oral Reading Fluency, intervention school scores almost doubled from the first assessment to the second (91%), whereas in the control schools the change was only 23%. These scores were celebrated throughout the Ministry. They were accompanied by direct senior MOE officials' experiences in intervention classrooms (e.g., according to one Case Study informant, during a training session for Governorate-level EGRP planning teams, team members visited a pilot classroom in a Fayoum governorate village and found the pupils there to be "cleverer" readers than they were). The conclusion that many took away from such evidence of effectiveness was that it would be unethical to deny such a dynamic intervention to the rest of the country's school children. Thus, the decision to scale-up was made.

B. EGRP Scale-up and Institutionalization

b.1. ***Scope of scale-up.*** The decision to scale-up was not without controversy. The Project team and RTI headquarters' staff advocated a gradual scale-up approach to be preceded by a national EGRA and a careful counting of costs and the marshaling of resources. An analysis was prepared by a senior RTI researcher which demonstrated resource requirements and presented alternative scenarios.[8] This was largely set aside by the MOE's First Deputy to the Minister Dr. Reda, as overly-cautious. Once he and his team were convinced of the scientific rigor and practical effectiveness of the EGRP model, they were determined to move as quickly as possible to a full-blown national scale up. Even though they realized that funding would be tight, they were determined to find the needed resources, even if it meant heightened levels of local and voluntary contributions, which, given the popularity of EGRP, they anticipated would be forthcoming. Their sense of urgency was reinforced by Egypt's Arab Spring turmoil (there had been three Ministers of Education appointed since President Mubarak departed): they wanted to scale-up and institutionalize this reform before the end of their tenures given uncertainties as to the future course of the Egyptian "revolution."

b.2. USAID Cairo supported this decision and cooperated with the MOE team in moving quickly, "striking while the iron was hot," in the words of one informant. A compromise was

[7] While the use of local enumerators was a formula for reducing the costs of an assessment like EGRA (and in this case, it did so by over 40%), it did come with a risk, since local enumerators could have been tempted to inflate student test scores in order to make their own territory look better. No specific cases of score inflation were found in EGRA 2011, but the bias could still be there undiscovered.

[8] See J. DeStefano, "GILO Support for Scaling-Up Improved Early Grade Reading Instruction," reproduced in the GILO Project *14th Quarterly Report - July 1 to September 30, 2011.*

reached with the more gradualist camp: GILO would still work on intense scaling-up in the four pilot governorates, and, at the same time assist the MOE with scaling-up in the remaining 23 governorates.[9] To get things moving, Dr. Reda asked the Project EGRP training consultant for an overall scale-up design. The plan that she developed employed a cascade model of dissemination.

b.3. ***Scale-up steps.*** The first step in the cascade was to form and train governorate EGRP "planning teams," which generally consisted of the Undersecretary for Education (the chief MOE officer in the governorate), the General Arabic Supervisor, and the Primary Education Manager, plus in some governorates training officers. These team members were introduced to EGRP through a video-conference held by the First Deputy Minister that featured the national working group and planning teams from the Project-supported governorates. Subsequently the teams were trained and supported in their own planning, which was to cover EGRP administration, logistics, financing, and training. In some cases, their orientation to EGRP included visits to program pilot sites. Video-conferences between the First Deputy Minister's office and the Governorate planning teams soon became a weekly event.

b.4. Once the governorate EGRP plans were reviewed and accepted, the next step was training the governorate and district trainers. The first round of this was conducted in a central location and lasted several days. It covered not only reading instruction, but also student centered learning, supervision and coaching, use of the library, and methods in the training of trainers. This led to more governorate and district cadre training in the governorates, followed by teacher training in the districts. Altogether, tens of thousands of 1st and 2nd grade teachers were trained (precise numbers are in question).[10] In conjunction with the training, trainer and teacher materials were distributed by the Project – about 150,000 manuals and teaching aids in all.[11]

b.5. ***Institutionalizing EGRP.*** Whereas scaling-up was one strategy to assure EGRP sustainability, institutionalization of EGRP components was another. Project staff in partnership with MOE experts worked feverishly to prepare EGRP methods and materials for take-over by permanent bodies. At the same time USAID/Egypt's Contracting Officer's Technical Representative (COTR), Hala El Serafy, partnered closely with MOE managers to create management and political openings. Two breakthroughs came 2012 in the areas of learning

[9] To accommodate GILO's expanded role in supporting EGRP scale-up, USAID extended the completion date of the Project three times: from Sept 2010 to Sept 2011, from Sept 2011 to 2012, and from Sept 2012 to March 2013. These were all no-cost extensions; additional Project resources were made available for EGRP by shifting allocations within the Project, taking advantage of favorable changes in the exchange rate, and improving Project efficiency (e.g., scaling down on staff and stretching existing resources further – moves similar to those in the Ministry).

[10] The 2012 USAID/Egypt GILO Fact Sheet indicates that 35,000 grade 1 teachers were trained in EGRP, whereas an MOE PowerPoint cites 58,000 (see Shahinaz Eldesouki Abdelhady Turk, *Early Grade Reading: Egypt*, Cairo: Ministry of Education, Sept. 2012). This difference reflects, at least in part, a divergence in how many hours were considered to constitute being trained with USAID having a stricter definition for inclusion of beneficiaries in its TraiNet reporting system.

[11] Just below 50,000 each of five different EGRP materials were distributed to teachers by the Project. See next section about the highly successful distribution process.

materials and teacher training. Early primary school textbooks had already been under revision before then, but just at the right strategic moment the Project was able to convince the curriculum center to align the grade 1 reader with EGRP teaching methods. The grade 2 readers are also being aligned with EGRP and will be released in 2013.

b.6. Concerning teacher training, the Ministry (with Project and USAID support) sought to get EGRP teacher training methods and materials accredited by the nation's Professional Academy of Teachers (PAT). In 2012, after a rigorous review, PAT accredited the 1st grade EGRP training program, training and teachers' manuals, and EGRP reference materials and assessment methods. It also incorporated EGRP training into its regular program for teachers and teacher trainers. New primary school teachers who are not yet licensed are now required to take the EGRP training program (along with other courses) as part of the certification (licensing process). In addition, teacher trainers can become PAT-certified by going through a 24-hour EGRP training course. This, among other things, makes it possible for them to be paid additional income for being EGRP trainers. When the textbook and accompanying trainers' manuals are developed for grade 2, PAT accreditation will be sought for that also.

b.7. What is not guaranteed is a program of continuous teacher training/refreshment and coaching, which will require continued political will and budget support (at all levels, given the decentralization movement).[12] Also, domestic and expat reading experts are quick to point out that the early grade reading program in Egypt is not yet fully developed. Besides the need to continue the work on grades 2 and 3 currently in the pipeline, it is crucial that EGRP be more fully integrated into primary school curriculum (see the **Possible Future Directions** section as the end of this report).

b.8. Another kind of institutionalization is the promulgation of policies and regulations supportive of EGRP, which has included: a) a regulation that phonics be taught for 25 minutes per day in early primary classrooms; b) a decree that allows mentor teachers posted in one school to give coaching on EGRP to teachers in others; and c) the requirement that Arabic Language supervisors at the primary school level use an EGRP observation form in assessing teacher performance in reading instruction and giving feedback.

b.9. Finally, the EGRA (assessment) has been partially institutionalized in the sense that it can be implemented by local assessment specialists, as shown by its having been administered in Cairo and Beheira governorates in 2011 without external involvement. This does not mean that EGRA has become fully institutionalized in the Ministry but that the capacity to independently administer the tool now exists.

C. Evidence of Scale-up Success

[12] After the drafting of this report a new Ministerial decree was published establishing an EGRP Unit at the headquarters of the Ministry of Education to further the professional development of teachers in early grade reading and math (Ministerial Decree No. 28, dated 22/1/2013).

c.1. There is a widespread perception within and outside Egypt that the scaling up of EGRP has been a success of unprecedented proportions. Hard data confirmation of this general view is not as easy to come by. However, two recent (2012) GILO studies do shed some light on the subject. The first was a telephone survey of scale-up implementation. The second was GILO's use of a rating scale – along 23 dimensions of EGRP implementation – to compare how effectively the 27 governorates have been establishing the system.

Telephone survey[13]

c.2. This was based on the assumption that a school's receipt of EGRP materials made EGRP *operational* there. Since material distribution was, like training, done through the cascade model, the proportion of schools becoming EGRP operational *in this sense* was a good indicator of whether scaling-up at the national level truly occurred.

c.3. The survey was conducted in all 23 non-Project governorates. Two *idara* were selected at random in each (except for three small governorates where only one was selected) and in each *idara* three schools were randomly selected. Of the 129 schools selected, telephone contact was made with 111.

c.4. The survey sought to find out which of five grade 1 resource materials had been received by the schools. The criterion established for the school having become operational was the receipt of *at least three* of the five kinds of resource materials. This was considered reasonable, since the different materials were generally shipped at different times.

c.5. The results of the survey, conducted by the Project staff, were as follows: of the 111 schools reached, 100 (90.1%) had received three or more of the EGRP resources materials. Most of the rest had received at least two kinds of materials; almost none had failed to receive any EGRP materials. From this point of view and given the representativeness of the sample, scaling-up reached almost all of the country's schools. Given that distribution was done by cascade, this finding suggests excellent system functioning in getting the system "out there."

Implementation ratings by governorate.

c.6. During the first half of school year 2012-2013 the Project staff undertook an assessment of EGRP implementation effectiveness by rating all 27 governorates on 23 dimensions of program management.[14] These ratings were done mostly for operational reasons – to help the Project identify locations that required implementation support. The dimensions were at five organization levels: *Muderiya* Planning Team; *Muderiya* Cadre Team; *Idara* Cadre Team; Follow-Up and Supervision; and Teachers. Ratings could range from zero to three (high). Results are presented in Table 1 below, and in Figure 2 (following page), the latter contrasting the ratings of high and low performing governorates. (See Annex F for a more detailed analysis.)

c.7. Looking at column "c," the red shadings (showing those items rated 2.5 or above) were mostly at the *muderiya* level; the *Muderiya Cadre Team* being especially highly rated.

[13] For a full report, see *GILO Quarterly Report #19* (October-December, 2012).
[14] For details, see *GILO Quarterly Report #18* (July-September, 2012).

Proceeding down the column, the *idara* teams received moderate ratings, and the local groups (supervisors and teachers) received mainly low to moderate. These findings seem to illustrate the limitations of the cascade system, with quality reductions following the spill-over from one level to the next. The most vulnerable feature of scale-up implementation appears to be the training sequence, as seen by the fact that training (or "training adequacy" – see teacher level) was not highly rated at any level.

Table 1. Ratings of Governorates on EGRP Implementation dimensions: all, high- & low-rated

(ratings on individual indicators range from 0 to 3 [high])

(a)	(b) Indicators	(c) All governorates	(d) High-rated governorates	(e) Low-rated governorates	(f) Low-High Difference
Planning team	Planning team is stable	2.7	2.9	2.6	0.3
	Team members know their roles	2.5	3	1.9	1.1
	Role in implementing training plan	2.4	2.8	1.7	1.1
	Role in social marketing	1.8	2.4	1.1	1.3
	Ability to face challenges	2.0	2.9	1.6	1.3
	Cooperative; coordinates with training team	2.3	2.6	1.8	0.8
Muderiya cadre team	Team formation is stable	2.8	2.9	2.6	0.3
	Participated in all basic training	2.9	3	2.9	0.1
	Represented by all Idaras	2.6	2.6	2.5	0.1
	Members have adequate training skills	2.2	2.8	1.6	1.2
	Members play their roles in training others	2.5	3	1.8	1.2
Idara Cadre team	Team formation is stable	2.8	2.9	2.7	0.2
	Participated in all training in Muderiya	2.4	2.8	1.8	1
	Represented by supervisors and teachers	2.5	2.8	1.9	0.9
	Has clear concepts about program applications	2.1	2.6	1.4	1.2
	Has good teacher training skills	2.2	3	1.4	1.6
Follow up and supervision	Arabic lang supervision represented in planning team	3.0	**3**	3	0
	All supervisors received EGRP training	2.2	**2.6**	2	0.6
	Supervisors follow up classrooms seriously	1.7	**2.3**	1.1	1.2
	Teachers see supervisors as very positive	1.6	**2.1**	0.9	1.2
Teachers	All teachers received training on EGRP	2.4	**2.7**	2.1	0.6
	Role in implementing training plan	2.1	**2.8**	1.6	1.2
	Teachers believe in projects; express results	1.6	**2.1**	1.2	0.9
	Average Total Score	53.6	62.3	43.2	19.1

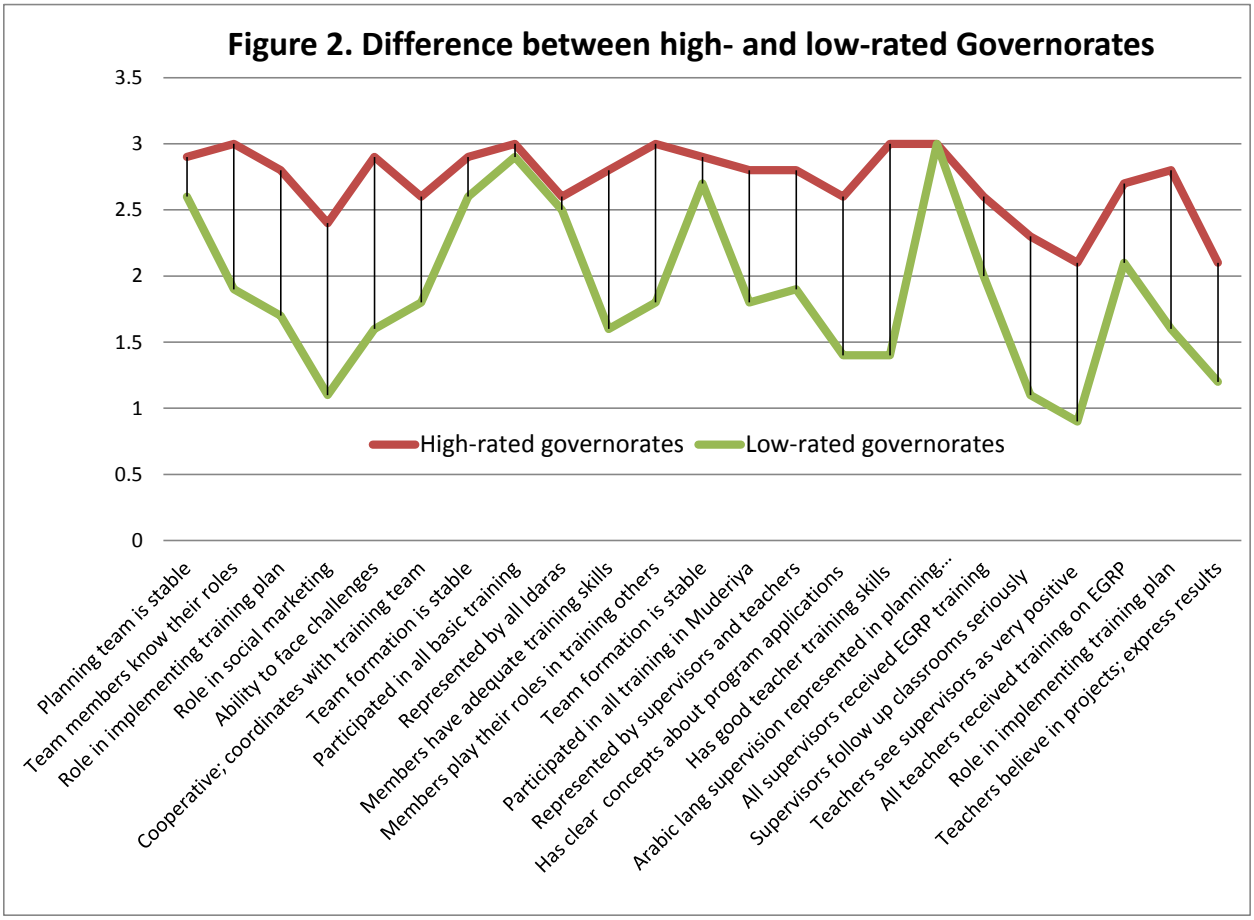
* Low rated governorates were (with composite score): Qalubia (36); Beheira (39); Luxor (40); So. Sinai (42.3; Assuit (43); Ismailia (45); Suhag (46); Matrough (48) and New Valley (49).

**High rated governorates were: No. Sinai (65.5); Qena (64); Port Said (64); Beni Suef (63.5); Aswan (63); Kafr el Sheikh (60); Suez (59.5); Cairo (59).

c.8. Variations in overall implementation effectiveness across governorates were also revealed, as noted in the Table 1's footnotes. Composite rating scores (across the 23 dimensions) ranged from **36** in Qalubia governorate to **65.5** in North Sinai. Surprisingly, the highest rated governorate (N. Sinai) was not one where EGRP was piloted; in fact, only two of

four pilot governorates (Qena and Beni Suef) were rated among the top eight, another testimony to the scale-up's ability to get the system "out there."

c.9. The differences between the highest- and lowest-rated governorates are visible in columns (e) and (f) of the Table – where the grey shadings show differences of more than one point (dark grey more than 1.5), and in Figure 2. The one *Muderiya Planning Team* rating that was low in general, "social marketing," was particularly bad in the low-rated governorates; also low-scoring governorates got relatedly poor ratings on training, especially the *idara* level, where all high-performing governorates received maximum scores (3.0) and low-performing governorates averaged scores below 1.5. Finally, big differences were found on the supervision dimensions, especially supervisor's follow-up and attitudes.



c.10. **In sum**, the two studies conducted by GILO about EGRP dissemination point in somewhat different directions. The phone survey on materials distribution provides strong evidence that the cascade system did get the system out to schools throughout the country. The rating study is more nuanced: it showed strong system functioning at the governorate level, with subsequent fall-off at the next two levels of the cascade. This is natural and to be expected with a cascade system, especially one that requires the acquisition of complex new teaching and mentoring skills. Also, these ratings were for the scale-up's first year, leaving open the possibility that the cascade worked better in the second. More encouraging for scale-up managers is that some governorates that have not been intensively supported by the Project (North Sinai, Port Said,

and Aswan) have shown remarkably strong implementation effectiveness, even at the district and school level. Where the implementation is relatively weak, it is mainly in the areas of teacher training and supervision, which presumably can be dealt with through continuous refreshers, and, if possible, targeted "trouble-shooting" support by national and/or *muderiya* trainers.

Perceptions of key informants

c.11. This Case Study's two-week window for live in-country data collection made possible the conduct of 20 key informant interviews and the observation of 32 classrooms and libraries within 13 schools in four governorates. School selection was made according to criteria agreed upon by GILO, USAID/Cairo, and the Case Study investigator. Care was taken to represent a wide variety of geographical areas; thus the governorates visited were: Cairo (Capital), Alexandria (Mediterranean); Damietta (Nile Delta), and Beni Suef (Upper Egypt). Within the governorates, the goal was to select two districts, one urban and one rural, however, given the mainly urban nature of Cairo and Alexandria only urban districts were sampled there. Within the districts a relatively high-performing school and a lower-performing one were selected. Only at the end of the mission did the GILO rating study of EGRP implementation (see c.6-10 above) come to light, so it was not possible to use the ratings in the selection of governorates to be visited. As it turns out, the governorates selected were all in the mid- to high-effectiveness range.[15] *The fact that none of the low ranked governorates were visited during the mission means that there is likely to be a positive bias in the observation results. This probable bias was attenuated somewhat by the design to visit rural districts in two of the governorates and to select both high- and low-performing schools in each district visited. Nevertheless, the observation results are treated in this report with caution.*

c.12. Interviews with key informants (see Annex A: EGRP-Egypt. Persons Interviewed) revealed high levels of enthusiasm for the Program at all levels, and virtually no criticism (see Testimonials in Textbox 1). Officials at the USAID office in Cairo were amazed at how the modest $40,000 allocated to EGRA instrument adaptation had grown into a national early reading program in just 2-3 years. National government officials expressed satisfaction with the intervention, in part, because of its scientific underpinning, and also its growing popularity among management and teachers. GILO indicated satisfaction with the strong collegial relationships that had grown up around the intervention, both within its team/sub-contractors and "across the aisle" with Government of Egypt (GOE) officials and specialists. All expressed amazement at how quickly the intervention spread nationally ("like wildfire," according to one informant). This occurred despite strong initial skepticism on the part of RTI, which expressed concern about the adequacy of existing resources. *The head of the national government team was determined to make the needed resources available, and, in the end, was able to do so, partly by drawing upon volunteerism (dedicated people donating time) and local resources from muderiyas, idaras, and schools (and their boards of trustees). Finally, in virtually all group discussions, managers and teachers spoke of improved student learning outcomes: higher*

[15] Visited governorate rankings on the EGRP implementation scale (across the 27 governorates) were: Beni Suef (4th); Cairo (8th); Damietta (10th) and Alexandria (16th). There were no governorates from the lowest third on this scale.

reading rates, lower dropout, improved school attendance. Unfortunately, however, up to now, there is no hard evidence for this.

Textbox 1: Testimonials from the Field

"We never dreamed of such a successful program." Barbara Toye-Welsh, GILO Director.

"EGRP is something beautiful…" Idara management team member at school meeting, ElWehda El Arabia Primary School, Cairo.

"We parents never dreamed of a reading program like [as good as] this." School board member at group discussion in Salah Salem District, Alexandria.

"Now 100% of my students are learning to read – not just the best 30% as before. I never thought that was possible. With much less effort than before, my results are much better." First grade teacher, Nazla Aqfahas Primary School, Beni Suef.

Classroom observations

c.13. Direct observations of reading instruction were made in 30 classrooms (about 60% grade 1 and 40% grade 2) and two libraries. Teachers were observed presenting their regular reading lessons. In most cases they were given advanced warning, but in some cases not. In all grade 1 classes observed, even in those without advanced warning, teachers were observed to be presenting EGRP lessons using Program materials (manuals, flip charts, visual aids – most locally constructed) and, in most places, official textbooks. The lessons generally followed scripted lesson plans, but there was much teacher creativity in the details (e.g., some teachers introduced a new letter with a small drama; some used a guessing game; some visual aids, etc.).

c.14. Second grade lessons, which were expected to cover word knowledge and reading comprehension, were less well-developed since an EGRP-aligned grade 2 textbook had not yet been developed. Thus, the comprehension lessons observed were mostly about predicting reading content from existing textbook pictures.

c.15. In most cases students in both grades were grouped into clusters of about 5-8 students; group processes were common as were individual responses to questions (the student standing or going to the board). All observed grade 1 classes proceeded through the small steps of letter recognition, phonics awareness, and word construction (from phonemes). In virtually all cases a very high proportion of students were actively engaged (80-90%), a proportion rarely seen.

c.16. Active learning strategies were also applied, but these were not as widespread as the phonics strategies. Observed combinations of phonics and active learning strategies were as follows:

****

- Phonics plus consistent active learning: 60% of classrooms observed;
- Phonics plus some active learning: 35% of classrooms observed;
- Phonics plus no active learning: 5% of classrooms observed.

****

c.17. Where "some active learning" was used, there was still much use of teacher-centered instruction (lecturing) and group responses (students calling out answers). In the few "no active learning" classrooms, there was nothing but teacher-centered instruction and choral responses.

c.18. Although this was not a representative sample of classrooms, it was varied enough (in both quality and geographic location) to confirm earlier survey results (see above) that EGRP methods had been extended widely across the education system despite the use of a cascade teacher training system. Also, although there has been no systemic study of teacher adoption of EGRP methods, the observations left a clear impression that the teaching of phonics and use of EGRP materials in the classroom are widespread. Given the short time of EGRP implementation in most locations (less than one year), it is not surprising that many teachers have not yet fully mastered the teaching of phonics with an active student learning approach. *What is surprising is how far and how wide the system has travelled in one year.*

EGRP in Pilot schools

c.19. By design, only one of the 13 Case Study schools was a pilot school in the original GILO-supported program. Three of its classrooms – two grade 1 and one grade 2 – and the library were visited during the last day of Case Study field work. In all three classrooms, EGRP was implemented according to the established protocol, using phonics combined with consistent active learning strategies. Also apparent were good *classroom management techniques* (one of the non-EGRP strategies covered by GILO in its pilot schools). Teachers seemed very confident using EGRP, as did the students, largely accounted for by the fact that pilot schools have been implementing EGRP for two academic years, not one. Teachers who are now in grade 2 previously taught in grade 1 and were trained in EGRP during the intervention's first year (2010-2011). Also, notable was the effective use of the library in enhancing reading skills, the librarian being skillful in linking the words and topics in books to the sounds and words covered in the grade 1 and 2 phonics lessons. *In short, the pilot school revealed a maturity and confidence that it is hoped the other schools can reach after two years in the program through continuous coaching, refresher training, and targeted support from informed supervisors.*

c.20. This being said, it was not the case that the best teaching of all was observed in the Beni Suef pilot school. The best EGRP teaching, as good as any this observer has seen anywhere, was observed in a primary school in a low income area of Cairo (grade 1) and in a rural area ("Watermelon Town") in coastal Damietta (also grade 1). *The fact that a one-year dissemination effort using the cascade system could have produced such excellence in teaching at the end of the line in these locations is a real tribute to the system and its managers.*

Observed student reading skills

c.21. So far no EGRA has been conducted of students in scale-up schools and, as far as could be determined; no other learning outcomes data have been systematically collected. Thus, except for the glowing anecdotal evidence, there is no proof that children are learning to read better under EGRP scale-up. During the Case Study a quick assessment of grade 2 reading skills was made in three schools: one urban and one rural school in Damietta, and the EGRP pilot school in Beni Suef (also rural). In all three schools, a handful of students (4-5) were chosen at

random and asked to read a passage from the grade 2 textbook. The same numbers of volunteers were invited to read the passages. In the two rural schools, most randomly-chosen readers were able to read some of the passage, but not fluently, whereas all of the volunteers read it fluently. In the Damietta urban school, most of the randomly chosen students also read fluently. *No clear conclusion can be drawn from these impressions given the tiny sample size, but some encouragement can be taken that they do not reflect the EGRA 2009 results, in which more than half of grade 2 respondents could not read a single word in isolation.*

D. Factors Contributing to and Constraining Success

d.1. Many explanations have been given for why EGRP in Egypt was received so enthusiastically and why the country was able to disseminate it so quickly. Below is a summary of the most recently heard factors. At the end are also some factors that constrained program success.

d.2. ***Bipartisan endorsement of the model*.** The EGRP instructional model is based on the use of phonics for the teaching of reading. At every turn during the Case Study, managers and implementers commented on how consistent this is with previous, time-tested approaches to learning Arabic, approaches that had been abandoned in recent decades, much to the detriment of reading skills acquisition. The difference between the EGRP approach and the traditional one is that EGRP also uses progressive student-centered pedagogy. This also makes it attractive to the progressives. Thus, two camps that are often at odds with one another, traditionalists and progressives, have found EGRP not only acceptable, but compelling. To be sure, there was some dissonance at the beginning, for example, among those who preferred phonics to be taught in a rote manner, or "whole language" (anti-phonetics) partisans, *but these voices are now almost completely silent, as Egypt goes ahead in virtually unanimous endorsement of the EGRP model.*

d.3. ***Scientific underpinnings of EGRA/EGRP.*** One of the features of EGRA that made it initially so attractive to Egyptian policy makers was its firm grounding in cognitive science.[16] Scientific grounding was further strengthened in Egypt as policy and program design features were pinned to EGRA 2009 and 2011 outcomes and the results of the grade 1 textbook analysis. Moreover, many Arabic curriculum specialists became convinced of the need for reform in reading instruction after a series of rigorous classroom observations they undertook jointly with GILO specialists. In addition, prominent national and Middle Eastern Arabic language experts were brought into the analysis and decision-making mix. Finally, MOE decision-makers were impressed by the professional credentials of technical and managerial staff that RTI brought to the Project. *All of this created for EGRA/EGRP in Egypt an aura of scientific credibility and rigor.*

d.4. ***A well-led, skillful and committed GILO team and effective RTI backstopping.*** RTI fielded Project staff, consultants, and set of subcontractors (CID Inc.; Keys to Effective Learning; Infonex) that became known for their technical competence, commitment, and flexibility (in the

[16] Senior MOE policy makers were particularly receptive to scientifically-grounded reforms at this time, since many of them, including Dr. Reda, were prominent academics with which a late Mubarak Minister of Education had decided to fill his cabinet.

words of one informant, "willing to stretch and be stretched into uncharted territory.") Key from the beginning was the staff's technical and diplomatic skill in reaching out to Ministry officials and local experts. Just as the Ministry took the Arab Spring "revolution" as a call to rebuild education, the GILO team sensed an opportunity to contribute to a higher cause.

d.5. ***Early MOE engagement and ownership.*** USAID and the Project staff and consultants all mentioned their insistence on engaging Egyptian professionals and managers in the EGRA/EGRP design and implementation process from the beginning, with the view of eventually turning it all over to the government. In fact, many EGRP design features, including the order of introducing letters of the alphabet, are based on the preferences of local language specialists. Hand-over of expertise has also been important, e.g., the 2012 conduct of EGRAs in Cairo and Beheira was almost completely in the hands of Egyptian assessment specialists. This has made for an almost total MOE ownership of the program, a fact that supported the MOE's national scale-up efforts and the likely future sustainability of the program. At the governorate level, it now seems crucial that *Muderiya Planning Teams* were assembled and given much authority over how EGRP was scaled up in their domains.[17] *Commenting about MOE ownership, one GILO manager remarked: "They learned the EGRP song and then began to sing it by themselves*."[18]

d.6. ***A stable central leadership team and ministry champion.*** It is difficult to overestimate how crucial it was to the success of the EGRP piloting and scale-up that the GOE assigned the effort to a young and enthusiastic "working group" consisting of committed professionals who stayed with the effort during almost its entire duration. This included, among others, Ms. Shahinaz Dessouki (Director of Basic Education), Ms. Hanaa Kassem (Expert in Arabic Language), and Dr. Ahmed Shalaby (Advisor Arabic Language). Even more crucial was the guiding hand of the Deputy Minister and (eventual) EGRP champion, Dr. Reda Abu Sree. *It was their unity, commitment, and vision that rallied the GOE behind the project and made the effort a true partnership.*

d.7. ***Young energetic mentors.*** The record is not entirely clear on this. In group discussions during school/district visits a constant refrain was how crucial district coaches and mentors (often recruited from among the best teachers) were to the success of EGRP dissemination. Repeatedly during these visits the enthusiasm for EGRP among the young protagonists was apparent through both their teaching and their comments afterwards. However, the governorate rating data (see Section C above) tends to show poor local training as one of the factors behind below average EGRP implementation. A fair conclusion would be that where young and committed mentors have been raised up, the system has performed particularly well. According to EGRP guidelines local coaches were uniquely recruited on the basis of their talent, not seniority. Qualifications included their being no more than 45 years in age, specialists in Arabic, and strong in communications skills. Their principal roles were as members of their own school-based teacher training unit, but a recent MOE regulation allowed them to coach EGRP in

[17] This is one factor for which there is some ambiguity: Many informants claimed the *Muderiya* Planning Teams were a key to success, but in governorates where implementation was weak, the rating study points to some weaknesses in the planning team as a reason.

[18] Taken from a GILO staff member interview, November 6, 2012.

other schools as well. Many have had their "credentials" certified by completing the PAT course for EGRP trainers. Comments from older teachers during the Case Study suggest that these coaches' youth does not stand in the way of their credibility. *The widespread deployment of this cadre has produced a large, sustainable, semi-professional training force*.

d.8. ***Quick results; teacher testimonials.*** One point that came up in meetings in all 13 schools is how quickly children began to improve in their reading skills under EGRP. This is confirmed by the results of the EGRA 2011, which showed dramatic improvement after only about 4-5 months of EGRP teaching and learning. Teachers attribute this to EGRP's simple step-by-step approach, which students can grasp quickly. Project reports are full of teacher testimonials about how much they and their pupils enjoy the system and how quickly they learn to read with it.[19] From the group discussions, it is clear that many of these features are also appreciated by parents. *These successes and testimonials have clearly contributed to the positive reputation enjoyed by the system*.

d.9. ***Volunteerism and local resources.*** Perhaps partly driven by its positive reputation, EGRP implementation has enjoyed high levels of voluntary support from implementers and supporters, including the contribution of time and resources. According to Dr. Reda this is one of the factors that made it possible to disseminate the system nationally, against all odds. The huge amount of training for national scale-up was able to go forward largely because teachers were willing to forgo the normal levels of compensation for meals and transportation. At the school level, teachers volunteered their own time and resources in creating AV materials and buying candy treats for students as rewards. Likewise, local MOE offices and school boards used their own funds to purchase hard-to-come-by student materials and exercise books.[20] Once start-up costs are covered, however, and the system moves towards institutionalized, it is likely that such high levels of volunteerism will not be required, nor even justified (especially given low teacher salaries and pockets urban and rural poverty). Furthermore, sustainability will become an issue as the novelty of EGRP begins to recede.

d.10. ***A skillful and politically adept USAID country team.*** Many informants point to the skillful behind-the-scenes strategizing by USAID/Egypt's COTR Hala El Serafy, who cultivated good will connections with a string of changing Education Ministers, and supported timely institutional-building moves, such as PAT's endorsement of EGRP training, and the revision of early grade language textbooks. *The good working relationship between MOE's First Deputy Minister and the USAID/Egypt office is one of the factors that supported rapid national EGRP scale-up*.

d.11. ***Factors constraining success*.** In the drive to establish EGRP as a national program, a decision was made not to precede the nation-wide scale-up with a national-level EGRA, which

[19] See, as an example, Project Quarterly Reports #14 and #17 which contain teacher testimonials revealing rapid student successes, printed in both English and Arabic.

[20] Elements of the decentralization system that GILO has helped to pioneer have the potential to strengthen local funding support for innovations like EGRP. These elements allow more local choice in what to fund, and recent actions in Alexandria (described in Project Quarterly Report #17) show how some districts are making use of these elements to allocate additional funds to EGRP.

could have laid down a national baseline for reading skills. Without such a baseline it is difficult to maintain the scientific rigor that so far has made EGRP so attractive to the professional community (a good baseline is needed to show the extent to which EGRP – in various parts of the country – is indeed improving reading skills). The situation is still redeemable – EGRA could be administered to grade 3 students in the first half of 2013 (those who have not yet experienced the EGRP system), with testing again two years later to reveal program-induced changes. This window of opportunity will need to be seized if the country wants to keep up its scientific rigor.[21]

d.12. Another threat to scientific rigor is the lack – so far – of good empirical data on teacher adoption and use of EGRP methods. In the design of the system, a teacher observation instrument was created and, according to Case Study informants is being used in supervision. However, so far there has been no attempt to aggregate the data from that or other instruments to document the extent to which teacher instructional behavior has actually changed in the direction of EGRP norms. Such an analysis would be useful for both policy making and trouble shooting.

d.13. Another constraining factor is class size, which in Egypt is well above the acceptable norms. In the 30 classrooms visited for the Case Study, students per classroom varied between 35 and 52. The former is already high for the kind of individualized and student-centered instruction used in EGRP. Numbers above 45 make this very challenging. (In some cases students were packed so densely into classrooms that they had to climb atop each other's desks in order to move to the front of the class for individual reporting or blackboard work.)

d.14. Finally, a strong feeling exists among the teachers in the 23 non-Project governorates that they have not yet received enough training to reach mastery in the system. (A reduced number of training days is one of the compromises that came with national scale-up.) This is being dealt with through refresher training at least twice a year, and it is hoped that such training will continue well after GILO's contract ends in March of 2013.[22]

d.15. ***Summary***. Egypt's success in scaling-up its fresh and innovative early reading program, unprecedented on the global scene in its scope and rapidity, is attributable to set of factors and forces that intersected in a unique and fortuitous manner. As EGRA was adapted for use with the Arabic language and Egyptian language specialists came to accept the need for change in the way reading was taught (in many ways moving back to an older "tried and true" system), a consensus was born, but the ground for innovation was also prepared by earlier projects – some

[21] Whereas a national EGRA was not clearly on the agenda during Case Study field work in November 2012, one has subsequently been set into motion through another contract mechanism covering a national sample of 3rd graders (as of yet not touched by EGRP) drawn from 25 of 27 governorates, and with data collection scheduled to begin in early March of this year.

[22] Subsequent to the drafting of this report a Ministerial Decree was issued (22/1/2013) establishing a permanent EGRP Unit at the Headquarters of the Ministry of Education, which is charged with continuing EGRP training. The first head of the Unit will be Hanaa Kassem, Arabic Language specialist, and one of the original EGRP Working Group members. Also, USAID has just announced a follow-up project to GILO (starting in 2013), which will provide additional EGRP support to the government.

USAID supported -- and an Arab Spring appetite for change and national building. This combined with effective management and commitment by both a core of committed EGRP "champions" in the Ministry and in the USAID and Project teams, which worked together and made sure of eventual full Egyptian ownership. *It is unlikely that this fortuitous combination factors and forces can/will be replicated in other countries, although many lessons can be learned from what has been transpired in Egypt's EGRP scale-up*. This is the subject of the next section.

E. Lessons Learned

E.1 ***For innovation managers and implementers.*** The Case Study has identified numerous lessons related to *innovation management and implementation* that have either already been articulated by key informants or are constructible from the material uncovered. An initial list is the following.

a. ***Science matters.*** It was the solid grounding of EGRA in cognitive science that made the EGRP attractive to the eventual Egyptian champions of the system; this plus the continued scientific rigor of the 2009 EGRA, pre-implementation classroom observations, and the textbook analysis in Egypt helped to turn the improvement of early grade reading into a movement in the country. Continued scientific credibility in Egypt and other parts of the world will depend on the conduct of a sample-based, nation-wide EGRA in 2013 and its use as evidence that the scaled-up intervention has made a difference.[23]
b. ***Bipartisan endorsement of the model (consensus) is a powerful motivator.*** Egypt was fortunate in that two language camps, the more traditional phonics advocates and the more progressive student-active-learning specialists, both endorsed the model used in EGRP. This consensus allowed the work to go forward with enthusiasm in both camps and without substantial opposition. Consensus of this sort can also be reached where it is less automatic, but often requires a long process of consensus building.
c. ***True partnerships generate energy.*** The GILO team and its USAID sponsors made sure to engage Egyptian education managers and language experts from the beginning, in both design decisions and program implementation. As ownership shifted fully to the Egyptian teams and managers (national and governorate), a lot of positive energy was generated that allowed all parties to share the burden and to take credit for successes.
d. ***Innovations with simple steps can catch on quickly.*** The EGRP was deeply grounded in learning theory, but not particularly complicated in its steps, building from simple letter recognition through phonemic awareness to word recognition and reading comprehension. Teachers find the steps easy to follow (easier than before) and that students catch on quickly. This has made the system popular among teachers, students, and parents.
e. ***Early and easily demonstrable gains get attention and generate enthusiasm*****.** Those using the system have found (at least based on anecdotal evidence) that their students' reading abilities have improved quickly, and this has affected other outcomes such as

[23] As mentioned in other places, this was undertaken subsequent to the drafting of this report.

student behavior and attendance (reduced dropout). Being able to assert marked improvement in reading skills has given teachers and schools an increased sense of accomplishment and feeds enthusiasm about the system.

f. ***Shifting from a teacher- to a student-centered pedagogy takes time****. Most observed teachers were able to adopt the phonics teaching approach to reading, but many did it in the traditional teacher-centered manner, suggesting that changing teacher styles takes time and requires continuous professional support and coaching.*

g. ***Successful pilots can be brought to scale with the right kind of leadership, ownership, and commitment.*** This program shows that successful pilots can be scaled-up to the national level, with no more than reasonable and expected levels of quality reduction. This could not have been done, however, without widespread consensus over the need and way to improve early grade reading, a political-moral mandate to reach all primary school children, and dynamic leadership on the part of Egyptian managers and scientists, full system ownership (facilitated by a wise contractor), and strong commitment by all parties.

h. ***Radical positive changes can occur "beneath the radar" even during times of instability and conflict on the main stage.*** One of the most surprising contextual features of EGRP scale-up in Egypt is that it occurred during a period of extreme national turmoil and uncertainly. Ironically, this may have been one of the factors contributing to success, since program champions wanted to make sure that the program was implanted and institutionalized before any political/organization shifts occurred that might threaten to undo it.

i. ***Widespread scaling-up is possible even when funds are tight if the government can increase efficiency through increased volunteerism and local buy-ins.*** This is one of the dividends of a popular innovation. As start-up costs are covered and the system moves towards institutionalization, nationally and locally, such contributions will be less necessary and difficult to sustain in any case, as the novelty of the intervention wears off.

j. ***A cascade model of dissemination presents both strengths and limitations: without it EGRP could not have been disseminated so rapidly and inclusively, but with it came the consequence of regional disparities, as some localities were better prepared to take up scaling-up challenges than others.*** Such limitations can be addressed as long as they are recognized up front and proactively compensated for. This would entail national and governorate officials becoming fully aware of where implementation falls below acceptable standards (using feedback mechanisms like GILO's rating system) and providing resources for trouble-shooting and remediation.

E.2. ***For Other Countries/USAID Support Programs.*** This detailed and systematic study of the Egyptian Case has generated some insights that might be helpful for other countries or USAID as they consider expanding work in this field. Following are some ideas.

a. The scientific rigor of EGRA/EGRP is an asset that must be safeguarded as advocates, managers, and implementers think about/work on cross- or within-country dissemination.

b. A viable model to consider for introducing and/or scaling up EGRAs and interventions is the one followed by Egypt (not necessarily by design but more by default). Egypt implemented a modestly-funded program at the beginning and then, relying on skillful implementers/contractors and local USAID managers, nurtured its growth. The Egyptian program grew within and in response to the ever changing political/educational environment, moving briskly at times to seize opportunities, but also gradually growing local ownership and commitment, so when a chance did arise for scaling-up, substantial political will and domestic funding (including volunteerism) had materialized.
c. No single country, even one as successful as Egypt, can be used the prototype of how to scale-up an effective early grade reading program, given the unique combination of factors and forces in each country. However, countries should be able to learn lessons from one another for use within their own contexts.
d. An exciting and immediately successful learning system may be more effective in increasing school attendance than more conventional methods to increase access. Such methods, for example, using social promotions or family incentives may be able to bring reluctant children to school *but cannot hold them* if instruction is weak and little is learned. Experience with EGRP has shown (at least anecdotally) that where students enjoy learning and feel a sense of accomplishment attendance soars.
e. Macro-level instability and fragility is not always detrimental to the success of reform effects – as long as there is a stable corps of visionary managers at the operational level, supportive and sensitive contractors and Agency managers, and some space to innovate under the radar. Once the government of Mubarak fell, even though chaos ensued in the wide political realm, the educational bureaucracy felt liberated and charged with a mandate to move into new realms of nation building.
f. Effective training and real professional growth of teachers are the keys to creating improved student learning outcomes in low-performing educational systems. From the beginning of EGRP it was clear to GILO that in Egypt, as elsewhere, the educational system's main asset and best change agent was the classroom teacher. That's why most of the Project's funds were spent on teacher training and teacher coaching efforts. As a result the teachers, especially those whose vision and abilities were ignited by the program, did become the agents of change, through creating teacher materials, inventing clever ways to activate children, speaking out on the benefits of the system, and mentoring others. The best trainers in the program have turned out to be young, energetic teacher/coaches who have found a calling to help others improve their teaching.

F. Possible Future Directions

f.1. Evidence gathered in this Case Study has established that Egypt's new EGRP has been nationally disseminated with a high degree of success and is well on its way to being institutionalized. Yet, given its recentness, it is fair to say that it has not yet taken root: continued work and further model development is needed in order to assure that. What follows are some considerations concerning future directions that became apparent during this mission. Many of them are already acknowledged by GOE and USAID officials; some are already underway.

f.2 ***Sustainability.*** Given that EGRP national scale-up is now entering its second calendar year, and GILO is preparing to close in March of 2013, the question on everyone's lips is "Is the program sustainable?" The consistent answer to that during the Case Study field work was, from the top of the Ministry to the bottom (classroom teachers and parents), an emphatic "yes!" Certainly the MOE is planning to maintain it as a national priority, and given the institutionalization of many of its features (see Section II above); much of it is now part of the "establishment." It is possible, however, to view two dimensions of the system: operational and development. The government does seem poised to maintain program operations and to cover operational costs, many of which have been pushed down from the national to the governorate, district, and local (school/community) levels, in the spirit of decentralization.[24] However, development costs are often harder for governments to maintain. These are costs that are not in routine budgets, but instead, at least in Egypt, are often covered by international financial assistance. It seems less likely that Egypt will be able to cover with domestic funding all of the development costs needed to fully develop the early grade reading system. *For additional development tasks, it is likely that the nation will need to seek external funding. In response, USAID, as the leading international support of basic education in Egypt, would do well to keep the momentum going, at least until the system is fully developed and fully institutionalized.*[25]

f.3. ***Further model development.*** There was much discussion during the mission of the remaining model development tasks for EGRP in Egypt. Some of the tasks are near completion or in the pipeline: This included producing a revised (EGRP-consistent) textbook for grade 2 (scheduled to appear in early 2013) and related teacher routines ("scripted lessons"). On the drawing board is also planning for grade 3 EGRP (textbooks and teacher routines). There has also been serious talk of a sample-based national EGRA in 2013 (e.g., for 3rd graders – see above), which, given national self-reliance in early reading assessment management, could be administered by the national assessment agency.[26]

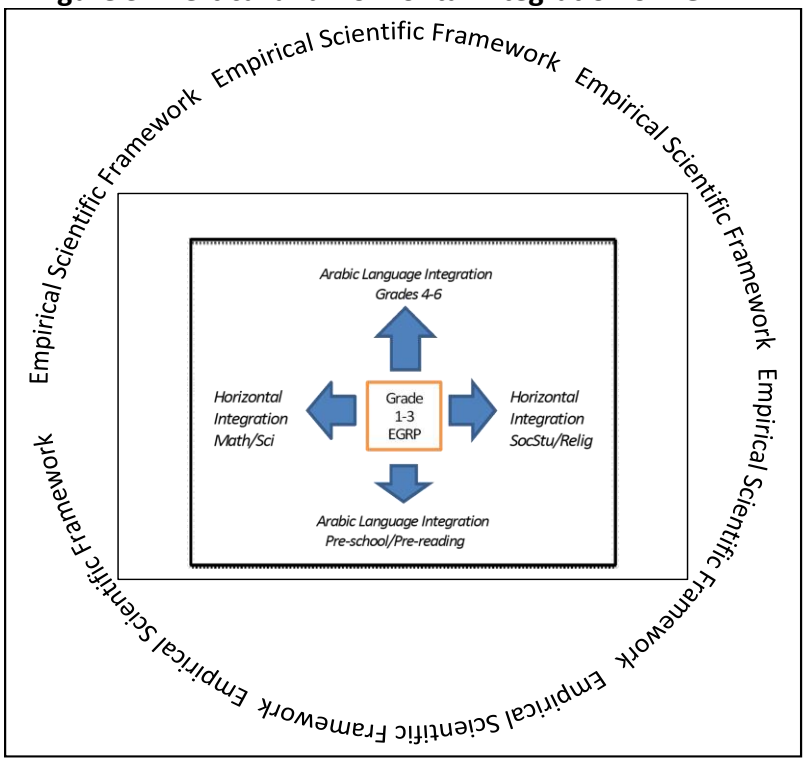
[24] One good sign of the Government's determination to strengthen EGRP operations was the establishment of a permanent EGRP unit in the MOE by Ministerial Decree in January 2013 (see footnote #23).

[25] When this report was first drafted it was unclear as to whether USAID would field a continuation project, but USAID's Business Forecast now lists a follow-up project in Egypt called *Core Skills and Organizational Reform in Education*, which is to "provide technical assistance to the MOE in the nationwide early grade learning program."

[26] After this report was drafted, a program was set in motion to conduct a national, sample-based EGRA with USAID technical support in March-April 2013 covering grade 3.

f.4. More into the future there is a need for continued EGRP model development, for example, link early grade reading to "pre-reading" programs at the pre-school level and to more advanced literacy education in upper primary grades (*vertical integration*), as well to develop ways of connecting improved early reading skills to other parts of the curriculum (e.g., math and social studies), *horizontal integration. All of such extensions would need to be considered "research and development" ventures, in order to preserve the strong scientific basis for expanded model development. See Figure 3 for a way of visualizing this.*

Figure 3: Vertical and Horizontal Integration of EGRP



f.5. ***Improved system efficiency.*** Continued government financing of costs of EGRP maintenance may require that even greater system economies be found (recalling that recent scaling-up was, in part, made possible by improving the cost-effectiveness of training). It is not clear how long the government can rely on volunteerism and local contributions. Their continuation may depend on how successful government is in maintaining a pioneering attitude among practitioners – involving them in continuous system improvement *resulting in tangible improvements in teaching conditions/rewards and student outcomes.* Another promising avenue is greater reliance on e-learning modes of training and professional refreshment. CDs and DVDs have already been produced containing EGRP materials and video-taped lessons. These can be made available in computer/internet-based training programs, which can be accessed individually or by groups. These can be accompanied by computer-mediated assessment systems, discussion groups, and forums (all already being developed and piloted). *Such e-learning formats can result in vastly improved training program efficiencies, but care must be*

taken to assess the extent to which e-learning can/does actually leverage changes in teacher understandings and teaching behaviors. Again, this is a question that requires the maintenance of an empirical scientific attitude.

f.6. ***Strengthening monitoring and evaluation.*** During EGRP's piloting and scale up the Project and the Ministry were adept at gathering feedback from the field and using it to shape or reshape the intervention. Examples of this are the supervisor's use of the EGRP observation form and the rating study of EGRP implementation per governorate (reviewed in Section B2 above) used by the GILO to identify locations of faltering implementation. It is important that such efforts be continued and expanded as means of identifying high and low system performance and designing interventions to overcome the latter.

f.7. ***Connections to higher education.*** So far most of the teacher training undertaken for EGRP implementation is in-service in nature (including that provided by PAT). A new generation of teachers is coming through Faculties of Education who also need to be exposed to EGRP during the pre-service years. This will require much greater interaction between the program and institutions of higher education. One of the key MOE EGRP Working Group members is embarking on a new doctoral program with the intention to use EGRP as her dissertation research laboratory. *EGRA/EGRP should be seen in Egypt as a virtual goldmine of research opportunities, in fact, as one of the best labs for early grade reading assessment and instruction in the world.*

ANNEXES

ANNEX A: EGRP - Egypt Case Study. Persons Interviewed

USAID/EGYPT
Ms. Hala El Serafy, COTR, USAID Cairo
Ms. Evelyn Rodriguez-Perez, (former USAID Team Leader in Cairo, now USAID/Peru)

GILO/RTI INTERNATIONAL
Dr. Amber Gove, RTI Washington Director Teaching and Learning
Dr. Barbara Toye-Welsh, GILO Staff (Outgoing Chief of Party)
Dr. Samir Shafik, GILO Teacher Training Specialist and subsequent Chief of Party
Dr. Sylvia Linan-Thompson, EGRA/EGRP Reading and Training Consultant
Ms. Michelle Ward-Brent, RTI GILO Home Office Technical Manager
Ms. Joseph DeStefano, RTI Headquarters
Dr. Robert LaTowsky, GILO sub-contractor Infonex
Ms. Diane Prouty-Harris, GILO Teacher Professional Development Specialist (former)
Dr. Laila Iskandar, GILO, subcontractor CID, Inc.
Dr. Ashraf Maher, GILO, subcontractor CID, Inc.
Dr. Essam Assad, GILO, subcontractor CID, Inc.

MINISTRY OF EDUCATION AND OTHERS
Dr. Reda Abu Serie, Ministry of Education (formerly First Deputy to the Minister)
Dr. Alaa Sabra, Vice Chairman, Professional Academy of Teachers
Dr. Ramadan M. Ramadan, Chairman, Professional Academy of Teachers
Ms. Shahinaz Dessouki, former MOE Director of Basic Education (now Vice-Minister in charge of Cairo governorate schools)
Ms. Hanaa Kasseim, Ministry of Education, Expert Arabic Language
Dr. Ahmed Shalaby, Ministry of Education Advisor Arabic Language
Dr. Isobel Coleman, Council on Foreign Relations

ANNEX B. EGRP – Egypt Case Study. Schools Visited

Governorate	District	School
Cairo	Abdeen	Qalaly Primary School
		El Wehda El Arabia Primary School
Alexandria	Salah Salem	Salah Salem School
	A Gomrok	Abi El Dar Daa School
Damietta	Faraskaar	Al Tarha Primary School
	Damietta	Atef Sharaf El Din School
		Ibn Khaldoun School (midtown)
	Kafr El Bateekh	Al Awamer School
		School 2
Beni Suef	Ihnasia	Al Refia School
		Mayanna Girls' Primary School
	Somasta	Yehia Galal Primary School
		Al Asakra New School (GILO-Pilot)

Annex C. Main Features of EGRA 2009 and EGRA 2011

Feature	EGRA 2009	EGRA 2011
1. Skills assessed	Based on general EGRA features, yet modified in Egypt according to the nature of the Arabic language in the country and suggestions from Egypt Arabic specialists; covered the standard EGRA components: a) letter naming fluency; b) syllable reading fluency; c) word reading fluency; d) unfamiliar word reading fluency; e) oral reading fluency, and f) reading comprehension.	Same as EGRA 2009.
2. Students assessed	2900 students in grades 2, 3, and 4 in each of 59 schools—29 primary schools in the first cohort of project-supported schools in 11 districts of Fayoum, Minya, and Qena governorates in Upper Egypt, plus an equivalent sample in 30 control schools (10 in each of the 3 governorates from districts different from project districts yet similar in important ways).	The sample of schools was the same as that in EGRA 2009, except that 3 of the 2009 schools were dropped (2 control schools in Fayoum and 1 project school in Qena). Only grade 2 students were assessed to create a comparison group with EGRA 2009 grade 2 students (one of the two grades experiencing EGRP interventions). The total numbers were 574 grade 2 students from project schools and 635 from control schools (compared to 444 and 465 second graders in EGRA 2009).
3. Selection of students	Students in the 3 grades were selected at random from their classrooms (from class lists) from project and control schools. Substitutions (few) were made by taking the next student on the list. Girls and boys were equally represented in the project and control schools.	The second graders in the sampled project and control schools were selected at random from class lists, with substitution for absent students made as in 2009. Girls and boys were equally represented in the project and control schools.
4. Administration of assessment	Forty (40) enumerators (data collectors) and 7-10 assessment team leaders were selected from among volunteers and staff of leading NGOs in education and recent university graduates in the Minya governorate (GILO-selected and supported; not MOE). The same team of enumerators worked in all 3 governorates.	Half of the enumerators were taken from among those who implemented the 2009 EGRA; the other half were MOE staff taken from the 3 governorates. The assessment teams were thus a mix of experienced (2009) enumerators and a new MOE partner.
5. Training of administrators	Enumerators received 4 days of training from GILO (one EGRA expert plus translator). The trainee/trainer ratio was about 20-1.	GILO trained the enumerators and assessment team leaders for 3 days. Those having experience with EGRA 2009 received a total of 7 days of training, while those from the MOE received 3 days. GILO provided the trainer (a single expert). The trainee-trainer ratio was about 20:1.
6. Administration process	The assessment was administered by trained enumerators 1 pupil at a time, pupils being shown stimulus materials and the enumerator recording their responses. This process lasted about 1/2 hour per student.	Same as for 2009.
7. Assessment materials	Provided by GILO	Provided by GILO

8. Duration of assessment	EGRA 2009 was administered over 3 weeks during January and February at the rate of about 5 schools per day (20 Minya schools in 4 days; 19 Fayoum in 4 days; and 20 in Qena in 4 days).	EGRA 2011 was administered during April and May (a 3-month delay due to closing of schools because of instability). Ten (10) schools were covered per day (given the need to assess only in grade 2), meaning the assessment took 2 days in each governorate.
9. Cost of the assessment	The costs of EGRA 2009 were covered by GILO. The cost elements were as follows: Enumerator fees, $14,800 ($400 per unit); transportation, $1,345 ($88 per enumerator); accommodations, $8,509 ($851 per person needing accommodations); assessment materials, $970 ($0.33 per subject); tea breaks, $545 ($15 per enumerator). This totals approximately $9 per pupil tested.	The costs of the EGRA 2011 were again covered by GILO, in part to keep the test secure. The cost elements were the same as in 2009, however since half of the enumerators were local MOE officials the outlay for fees, transportation (inter-governorate), and accommodations were reduced—by 40% for transportation and 50% for the rest; also given the need to spend only two days in the field per governorate (compared to 4 in 2009) the above costs were again reduced by 50%. The breakdown was as follows: enumerator fees, $2,960; transportation, $336; accommodations, $2,127; assessment materials, $200 ($0.33 per subject); tea breaks $136. This totals just under $5 per pupil tested.
10. Management/analysis of data	Every test item was scored by 2 people to enhance accuracy and then entered into Excel spread sheets, which were randomly checked for errors. Any errors found were corrected. Calculations of subtest scores were made according to benchmarks recommended by the reading consultant and were checked by RTI statisticians for accuracy.	Same as in EGRA 2009. Cross-sectional comparisons were made on sub-test scores, comparing 2009-2011 differences for both the intervention and control schools.

ANNEX D. EGRP - Egypt Case Study. Essential Features of EGRP Interventions or Program

Case study question or topic	EGRP intervention that was first piloted	EGRP expansion in the 4 GILO-supported governorates	EGRP intervention taken to scale nationally
1. How was the content of the intervention (program) developed?	EGRP was developed cooperatively between GILO and its consultants (e.g., early reading and Arabic language specialists) and MOE specialists (eventually becoming the EGRP Working Group), based on reviews of existing Arabic language curricula and teachers' guides, the negative results of EGRA, insight from EGRA and other sources into the positive relationship between students' ability to read syllables (letter sounds) and their decoding unknown words and reading passages, and a critical review of the grade 1 Arabic language textbook. The program, designed to be piloted in GILO schools, included phonics, phonemic awareness, word lists, decodable sentences and stories, and comprehension routines. GILO staff and sub-contractors worked in cooperation with the MOE Working Group in developing teacher/supervisor/coach training packages and routines.	This intervention was based on the pilot one, using the same collaborative process, but was modified to align the program with the new grade 1 textbook, which covered the Arabic alphabet over the school year in alphabetical order.	The content of the EGRP intervention for the 23 non-GILO governorates was the same as for the GILO-supported governorates.
2. What materials were needed to deliver the program to the students?	Initially there were no formal training/teaching materials, so teacher-made materials were used. Eventually (by Dec. 2010) teachers' manuals, CDs, and other supporting resources were produced and widely distributed. These were developed by GILO and its subcontractors in consultation with the MOE EGRP Working Group. Teachers continued to be encouraged to create their own teaching aids.	Grade 1 and 2 teachers all received GILO-created training materials (see below). Grade 1 teachers received the grade 1 teachers' guidebook, flipbooks (2 each), and CDs of learning materials and DVDs of video-taped teaching routines; EGRP student worksheets, which were not yet available in large numbers for students, had to be copied by teachers, schools, school boards, or districts. Materials for grade 2 were created to strengthen word knowledge and reading comprehension in conjunction with the (yet to be revised) grade 2 textbook.	The same as for the program in 4 GILO supported governorates, namely: Early grade reading manual and flipbooks (for teachers) with CDs of materials and a DVD of video-taped teaching routines; EGRP student worksheets, which were not yet available in large numbers for students, had to be copied by teachers, schools, school boards, or districts. (A study conducted by GILO in Nov. 2012 among 111 schools selected randomly from all governorates indicated that 90% of schools had actually received at least 3 of 5 kinds of EGRP grade 1 materials sent to schools.)

3. What was the nature of the teacher training (# of teachers trained; trainer-trainee ratios, nature of trainers (GILO or MOE staff or both)?	During school break in 2010 teachers in the GILO schools in the 3 project governorates were trained an average of 6-8 days in intervention routines including the use of specific learning aids. Trainers came from GILO staff/subcontractors and professionals from the MOE EGRP Working Group. Number of teachers trained? Grades? Trainee-trainer ratios? GILO staff (sub-contractors) spent a lot of time in schools doing on-the-job training (modeling lessons and working with teachers on practice teaching sessions).	Training of teachers was conducted by GILO staff/sub-contractors in the 4 GILO-supported governorates. During the period of August 2011 to September 2012 GILO trained approximately 9,800 grade 1 and 2 teachers for 18-24 hours; another approximately 2,500 teachers received 1-17 hours of training. School administrators also received training (approximately 500 of them), but about 40% only received 1-11 hours.	Teacher training proceeded via the cascade system, wherein GILO trained governorate trainers and many district (*idara*) trainers, who then trained teachers. In the first year of EGRP expansion (2011-2012) only grade 1 teachers were trained (5 days of initial training and 2 days refresher training during mid-year break); in the second year grade 2 teachers were trained and grade 1 teachers were given refresher training. GILO trained about 3,156 *idara* trainers in mid-2012; these trainers delivered "countless" training workshops to teachers in the 23 non-GILO governorates. (According to the then Director of Basic Education, Shahinaz Dessouki, the total number of grade 1 and 2 teachers trained in 2011-2012 for EGRP was 58,000 and 56,000 teachers respectively.)
4. What materials were used to support the training (who developed, per trainee ratios)?	At first there were no training materials for EGRP. However, by December 2010 (see above) teacher manuals and materials, plus CDs containing additional teaching aids, were available and used for training.	Trainers and managers (supervisors and principals) received Trainers' Manuals for Grade 1 EGRP; Orientation Manual for Supervisors and Principals; Overview and Introduction to Early Grade Reading; facilitator and participant manuals on Integrating Early Grade Reading with the Library (also for librarian); and e-learning materials available online (although currently to a few pilot schools only).	Training of Trainers' Manual for Grade 1 EGRP; Orientation Manual for Supervisors and Principals; Overview and Introduction to Early Grade Reading; set of facilitator and participant manuals on Integrating Early Grade Reading with the Library; and e-learning materials available online (although not yet widespread).
5. What was the nature of coaches' (and others') training?	The coaches and supervisors received the same EGRP training as the teachers, but, in addition, a generic package for coaching and mentoring. Also, when GILO staff was in schools doing on-the-job training and coaching, school principals and supervisors were invited to observe EGRP sessions.	Coaches were trained in EGRP processes as well as in coaching and supervising, including the following: what are coaching and supervision (including roles and responsibilities for each); how to do coaching and supervising; the use of the Classroom Observation Form; and how to give constructive feedback. During 2011-2012 about 1,200 supervisors were trained by GILO, 63% for 24 hours or more.	GILO trained 3,156 *idara* trainers, many of whom also acted as coaches during the implementation of the intervention. In addition, GILO trained over 500 supervisors (more than 60% for 24 hours or more), who also provided coaching and feedback.

6. What was the nature of the support work that went on (coach-teacher ratio; # of visits to a teacher per month by coaches; the duration of visits, etc.)?	The MOE's system was for teachers to be given support at the school level through School-Based Teacher Education Units, which during early EGRP were operational in some schools/districts and not others. Also, supervision was facilitated through the use of a classroom observation tool (generic at first, but later modified specifically for observing Arabic instruction). During the first year of EGRP implementation, GILO staff maintained close contact with schools, visiting them, on average, once a week during the first phase and once in two weeks during the second.	Coaches and teachers plan lessons together, develop learning resources, and share feedback and reflections (not confirmed by data). According to the Teacher Cadre Law there are to be 40 teachers per supervisor (compliance unclear), and coaches are to be taken from among excellent teachers (usually senior), mainly selected to work in their own schools (one per school. This results in a ratio of 4-6 teachers per coach), but coaches sometimes also help in other schools. Coach/supervisor visits are expected to last one period. Feedback is mainly through the Classroom Observation Form.	In the national scale-up, coaches and supervisors are expected to visit classrooms with the same frequency as those in the GILO-supported schools. There are no-nationally aggregated data to indicate the extent to which this is taking place. A Ministerial decree was issued in 2012 that allowed classroom teachers, who had been trained as cadre/coaches, to visit classrooms in other schools for mentoring. Case study discussions in the *idaras* visited indicated that coaching is occurring in those *idaras* (but how much is undetermined). Many informants mentioned that the key to EGRP implementation success has been the work of EGRP cadres.
7. What were the costs and unit costs of the various factors in the intervention?	Data not available	Data not available	Data not available.
8. How effectively did teachers employ EGRP intervention strategies?	Most teachers picked up the system quickly and enthusiastically, but some did not. Those who did not pick up the system right away were given extra support, but those who could not pick it up at all were (recommended to be) moved out of early grade classrooms.	Teacher performance data are being collected routinely by supervisors (using official observation forms), but these data have not been aggregated across *idara* and *muderiya* so there is no general picture.	Same as GILO-supported.

ANNEX E: Summary of ERGA 2009 and 2012 Results

(Source: USAID Egypt, *Improved Reading Performance in Grade 2*: GILO-Supported Schools vs. Control Schools. Results of the Early Grade Reading Assessments (EGRAs): 2009 Baseline and 2011 Post-Intervention, Cairo, 2011.)

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

EGRA Measures	Mean Scores, INTERVENTION Schools		% Change	Mean Scores, CONTROL Schools		% Change
	2009	2011		2009	2011	
Syllable Reading	9.76	28.47	192%	8.55	10.1	18%
Word Reading	7.35	15.5	111%	5.56	7.45	34%
Oral Reading Fluency	11.09	21.14	91%	8.92	10.93	23%

Table 2

Comparing EGRA Results: 2009 and 2011

Improved Syllable Reading

Intervention versus Control Schools: Primary Grade 2

	Intervention Schools			Control Schools		
	2009	2011	Change	2009	2011	Change
27+ Correct Syllables	12.6%	51.2%	+ 38.6%	12.0%	13.5%	+ 1.5%
16-26 Correct Syllables	14.0%	15.9%	+ 1.9%	8.0%	10.7%	+ 2.7%
1-15 Correct Syllables	25.0%	21.8%	- 3.2%	28.4%	30.7%	+ 2.3%
NO Correct Syllables	48.4%	11.1%	- 37.3%	51.6%	45.0%	- 6.6%
Mean Number of Correct Letters	9.76	28.47		8.55	10.10	

Table 3

Comparing EGRA Results: 2009 and 2011

Improved Word Reading

Intervention versus Control Schools: Primary Grade 2

	Intervention Schools			Control Schools		
	2009	2011	Change	2009	2011	Change
25+ Correct Words	8.8%	24.2%	+ 15.4%	7.1%	10.6%	+ 3.5%
16-24 Correct Words	9.7%	12.2%	+ 2.5%	5.8%	6.6%	+ 0.8%
1-15 Correct Words	32.9%	42.2%	+ 9.3%	31.8%	31.0%	- 0.8%
NO Correct Words	48.6%	21.4%	- 27.2%	55.3%	51.8%	- 2.5%
Mean Number of Correct Words	7.35	15.50		5.56	7.45	

Table 4

Comparing EGRA Results: 2009 and 2011

Improved Oral Fluency (Passage Reading)

Intervention versus Control Schools: Primary Grade 2

	Intervention Schools			Control Schools		
	2009	2011	Change	2009	2011	Change
45+ Correct Words	4.7%	19.7%	+ 15.0%	4.1%	7.6%	+ 3.5%
21-44 Correct Words	15.1%	21.1%	+ 6.0%	11.4%	12.6%	+ 1.2%
1-20 Correct Words	35.8%	38.5%	+ 2.7%	33.8%	33.4%	- 0.4%
NO Correct Words	44.4%	20.7%	- 23.7%	50.8%	46.5%	- 4.3%
Mean Number of Correct Words	11.09	21.14		8.92	10.93	

ANNEX F. Detailed Analysis of EGRP Scale-up Implementation Ratings by Governorate

(Use in conjunction with Table 1 and Figure 2 in text.)

1. Looking first at Table 1, column "c" shows how the governorates were rated overall on the various dimensions. The average dimension rating was 2.3; and the range went from 3.0 (all governorates getting the highest rating) to 1.6. Overall, nine dimensions received high ratings (between 2.5 and 3.0, *shaded red*); ten were rated in the middle range (2.0 to 2.4, *shaded orange*) and only four were low (below 2.0, *shaded yellow*). Looking across levels, it is at the level of *muderiya* (especially *muderiya* cadre) that most of the high ratings were found (e.g., 4 of 5 dimensions for the *muderiya* cadre were so rated). Dimensions at the *idara* level mainly received ratings in the middle range. At the lower (field) levels where teachers and supervisors operate, the ratings were all but one in the low to middle range. These findings seem to illustrate the limitations of the cascade system, with quality reductions following the spill-over from one level to the next.

2. These data were also used in the Case Study to describe variations among the governorates and to identify the indicators that were the most powerful in discriminating between high and low performers. The theoretical maximum for ratings across the 23 items was 69, and of course, the theoretical minimum was 0. The actual (observed) maximum was 65.5 (North Sinai) and the minimum was Qalubia (36). The mean across the 27 governorates was 53.6 and the standard deviation was 8.4. The medium was 56.5. The distribution was fairly normal, with nine governorates sitting more than ½ standard deviation (SD) below the mean, and eight more than ½ SD above. In this analysis, the nine were grouped together as the low-rated governorates, and the eight on the other end of the curve were grouped as the high-rated (see the table for the names of the governorates in each group). As shown in Table 1 (columns "d" and "e"), the average mean score for the low-rated group was about ten points below the overall mean (43.2) and for the high-rated group about ten points above (62.3). One surprising fact is that the four governorates where EGRP was piloted were not all represented in the high-rated group – Qena and Beni Suef were, but not Fayoum and Minya, which were in a four way tie for 9th place. Also, three of the top five governorates in aggregate rankings were *not* GILO pilot governorates. Column "F" in the table shows the difference between mean ratings for the two groups, with differences greater than 1 shaded in gray (dark gray for the one difference greater than 1.5).

3. The differences are also represented graphically in Figure 2. The shaded items in the table's column "f" and the gaps between the red and green lines in the figure both show considerable difference between the two groups in the planning team's effectiveness. Gaps are particularly wide on "Role in implementing training plan," "Role in social marketing," and "Ability to face challenges," which are all seen as important elements of program leadership at the governorate level. There are also two notable gaps in *muderiya* Cadre Team, which in general saw high ratings, namely, under "Team members have adequate training skills" and "Team members play their roles in training others." This seems to indicate that GILO was not able to create solid *muderiya* training teams in all governorates. This lack is traceable to the next lower level, where the low-rated governorates also show weaknesses in *Idara Cadre Team* understanding of "program applications" and in *their own training skills* (the item with the biggest gap in ratings – 3 for high-rated governorates verses 1.4 for low). Finally, in the areas of effectiveness where all governorates fell short – supervision and teacher performance – low-rated governorates were particularly weak on supervisors "seriousness" with classroom follow-up and supervisors not rated as "positive" by teachers, and in teachers' perceptions that they had acquired "adequate skills."

ANNEX G. EGRP - Egypt Case Study. Scope of Work (from RTI)

The aim of this Case Study is to provide a detailed account of the EGRP that unfolded between 2008-2012 and the context within which it happened to enable USAID and its partners to better understand what took place and why, and ultimately to be better able to design and implement equally, if not more, successful programs in other countries. It is also intended to outline the essential steps that enabled the MOE to take EGRP to scale so rapidly and identify what sacrifices were made in terms of quality.

To accomplish this aim, the following tasks are proposed:

1. Gather and review all relevant GILO project documents: technical reports, quarterly reports, annual reports, EGRP materials, etc.
2. Interview relevant GILO staff: Barbara Toye-Walsh (Chief of Party) and other project staff in Egypt; and Michelle Ward-Brent (Home Office Technical Manager) in Washington, D.C.
3. Interview relevant MOE staff: Ms. Shahinaz Dessouki (Director, Basic of Education) and other staff; and former First Deputy to the Minister of Education, Dr. Reda Abu Sree, and other key people in the Ministry of Education, such as Arabic Language Specialists: Mrs. Hanaa Kassem and Dr. Ahmed Shalaby, and a sample of MOE local personnel (e.g., teachers, trainers, planning team members, master trainers and supervisors).
4. Interview relevant USAID staff: Hala El Serafy (Contracting Officer's Representative) and Evelyn Rodrigues-Perez (former Education Team Leader).
5. Outline the characteristic features of the first (pilot) EGRA exercise:

a. What EGR skills were assessed; what students were assessed: number by grade, gender, school, district, and governorate; how the students that were assessed were selected; who carried out the assessment (GILO, MOE, or other); how they were trained; hours trained; trainer-to-trainee ratio; who they were trained by (i.e., GILO staff or MOE staff); how the assessment was administered; who printed the assessment materials; cost of the assessment materials; duration over which the assessment took place; distances that had to be travelled to get to the schools; cost of various aspects of the assessment; who cleaned the data; who entered the data; who analyzed the data; etc. The aim here is not so much to delineate the details of a comprehensive EGRA exercise, but to know what exactly was done, by whom, and at what cost to GILO and the MOE, with an eye on the feasibility of taking the EGRA model to scale.[3]

b. Describe *relevant* aspects of the political economic (inside and outside the MOE, inclusive of larger GOE and USG/USAID issues) and education contexts within which this took place—aspects that impacted positively / negatively on the effort, aspects that might have been impacted by the effort. The various forces that may have shaped or been shaped by this effort, and the impact(s) made.

6. Outline the essential features of the EGR intervention that was first piloted

a. The content of the EGR intervention; how it was developed; the materials needed to deliver the program to the students; the cost of the materials; who bore the costs; the nature of the teacher training that went on: number of teachers trained, trainer-trainee ratios, nature of the trainers (GILO staff or MOE staff), materials used to support the training, who developed the materials, cost of the materials, materials-per-trainee ratios; the nature of coaches' (and others') training that went on (see details of teacher training); the nature of the support work that went on: the coach-teacher ratio, the

number of visits made by a coach to a teacher(s) on a monthly basis, the duration of these visits, etc. The aim here is to get a detailed understanding of the EGR intervention—the EGR recipe that was used and the various unit costs of that recipe—to generate the results that were acquired, again with an eye on assessing the feasibility of taking *this* EGR model to scale.

b. See 5b.

7. Outline the essential features of the assessment that was done to assess the above EGRP intervention, of the EGRP intervention to take to scale regionally.
8. Outline the essential features of the EGR intervention that was taken to scale regionally (Phase II)

9. Outline the essential features of the EGR intervention that was taken to scale nationally

10. Identify and describe key aspects of the overall EGRP institutionalized by the MOE/GOE (e.g., changes in the textbook curriculum that have resulted from the EGRP, materials that have been formally adopted by the MOE, training programs that have been approved by PAT, etc.). Identify what made it possible for the MOE to scale up and institutionalize aspects of EGRP, lessons learned, weaknesses, and areas for improvements.

11. Outline the essential steps that enabled the MOE to take EGRP to scale rapidly, identify what sacrifices were made in terms of quality and what did not work or could have been done differently.

12. Review and assess what data may exist, such as from an in-class teacher observation instrument, to try and assess to what extent teachers were successfully using the new instructional practices.

13. Arrange all of the above material together into a coherent and comprehensive Case Study that highlights the major factors behind the EGRPs' successes, lack of successes, etc., offering an *honest and objective account of the overall effort* and a set of guidelines for how best to design, implement, and scale-up effective early grade reading initiatives, based on the lessons learned from this Case Study.

Deliverables

• Prior to departure for Egypt, a draft outline of the proposed field methods will be submitted to RTI and USAID for comment. The consultant will receive a phone briefing from USAID's AME Contracting Officer's Representative (COR) Mitch Kirby at the outset.

• While in country, the consultant will share completed sections of the report with RTI and USAID on an ongoing basis for comments and feedback and will maintain contact with the USAID COR for updates and guidance.

• At the end of the trip, the consultant will produce a penultimate Case Study for additional review by RTI and USAID upon return.

• The consultant will also prepare a debriefing presentation for sharing with USAID that will highlight the important findings of this Case Study.

• The final Case Study, not to exceed 40 pages of formal text, excluding the Executive Summary. Annexes are not included in the page count.