EdData II

Task Order 20: Measurement and Research Support to Education Strategy Goal 1: Improved Reading Skills for 100 Million Children in Primary Grades by 2015

Costing Early Grade Reading Programs: An Examination of Various Costs and Issues Around Costing

EdData II Technical and Managerial Assistance, Task Number 20
Contract Number AID-OAA-BC-12-00003
Strategic Objective 3
RTI Project No. 09354.020.006

February 2014

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

Task Order 20: Measurement and Research Support to Education Strategy Goal 1: Improved Reading Skills for 100 Million Children in Primary Grades by 2015

Costing Early Grade Reading Programs: An Examination of Various Costs and Issues Around Costing

Prepared for
Bureau for Bureau for Economic Growth, Education, and Environment (E3)
United States Agency for International Development
Suezan Lee, Senior Education Advisor

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

Table of Contents

List of Tables

Abbreviations

CAII	Creative Associates International, Inc.
COP	Chief of Party
DCOP	Deputy Chief of Party
EdData II	Education Data for Decision Making (project)
EGR	early grade reading
EGRA	Early Grade Reading Assessment
GILO	Girls' Improved Learning Outcomes (project)
HCG	host-country government
K	Thousand
KEYS	Keys to Effective Learning (subcontractor)
M	Million
NGO	nongovernmental organization
R&D	research and development
RTI	RTI International (trade name of Research Triangle Institute)
SSME	Snapshot of School Management Effectiveness
USAID	United States Agency for International Development

1. Introduction and Overview

Substantial improvements in access to basic education have been achieved over the decades during which Education for All has been a global focus. Recently, the development community has turned an eye toward education quality—what children learn once they gain access to school. Particular attention has focused on early grade reading (EGR), which is a fundamental aspect of a high-quality education, not just as an academic *end* (learning how to read), but as a *means to many academic ends* (reading to learn). Over the past few years, early grade reading assessments have been undertaken in a great many developing countries, only to find that development of reading skills in most is rather poor. This being the case, improving instruction and outcomes in early grade reading in these countries has become the focus of many donors and development practitioners. USAID, for example, has put forth the goal of "improved reading skills for 100 million children in primary grades by 2015."[1]

Given this goal, USAID has begun to develop evidence-based EGR programs in a number of countries. As more and more USAID missions ready themselves to launch their own EGR programs, and as others are poised to take their evidence-based EGR programs to scale, the issues of cost and cost effectiveness become a matter of concern, particularly in the present political milieu of reduced U.S. Government spending. Of particular interest to USAID are (a) determining the costs of developing and implementing small- or medium-scale evidence-based EGR interventions, (b) examining the cost-effectiveness of those interventions, and (c) gaining some sense of what it might cost to take an evidence-based program to scale.

This study is a deliverable for an activity put forth in the EdData II Task Order *Measurement and Research Support to Education Strategy Goal 1: Improved Reading Skills for 100 Million Children in Primary Grades by 2015*: "Cost of Early Grade Reading Programs." The costs of five EGR efforts are presented and discussed in this study: three efforts that have already taken place (Egypt, Liberia, and Malawi), one that has been unfolding since 2011 and will continue through much of 2014 (Kenya), and one that started up in late 2012 and is scheduled to be completed by mid-2014 (Jordan). Since much of this costing work was done *post facto*, and because it would have been exceedingly difficult to obtain the necessary data, this activity looked only at U.S. dollar[2] costs: whatever role the host country government (HCG) played in the development of an evidence-based EGR program, the cost associated with that role was not included in this study. Finally, because cost data are proprietary and therefore not likely to be shared among various implementing partners, only those EGR programs that RTI was integrally

[1] USAID. (2011). *Education: Opportunity through learning. USAID Education Strategy 2011–2015.* Washington, DC, p. 1.

[2] All monetary amounts designated by "$" in this report refer to U.S. dollars. While the vast majority of the U.S. dollar costs in this study were borne by USAID, some were covered by the World Bank (Liberia) or the William and Flora Hewlett Foundation (Malawi).

involved in, either as a prime awardee or as a subawardee, were considered; thus the five efforts that are presented in this study.

For this analysis, the researchers deconstructed the work of developing of an evidence-based EGR program into five to seven[3] phases (A through G), which are presented in ***Table 1*** along with a description of the work that needs to be done to carry out each phase.

Table 1. Phases of development for evidence-based early grade reading programs

Phase		Description
A	Development of a viable Early Grade Reading Assessment (EGRA) instrument*	Developing an EGRA instrument from scratch would involve international reading experts, host-country reading and language experts, and local curriculum experts coming together to design it. Local reading materials or grade-level textbooks in the target language (if available) would be used to generate lists of words common to the appropriate age/grade vocabulary. Using those lists, the team of experts would develop the EGRA subtests that could properly test students' reading skills in the host country. The instrument would then be pilot tested to ensure both validity and reliability. This testing would require hiring people to administer the instrument, training them to administer it, and developing all the materials needed to administer it (i.e., printing all the test forms or programming them on an electronic tablet). Then this pilot instrument would have to be administered: The enumerators would go out into the field to test a limited number of students (often 50 to 100). Based on an analysis of the pilot test results, the instrument and the protocols for administering it would be finalized. One cost- and time-saving measure often employed at this stage is preparing several equated instruments at once, for use either as endline assessments or as follow-up tests in subsequent years. The process of using a pre-existing EGRA instrument in the target language and adapting it to a new country would take much less time and effort—usually no more than a week. The resulting instrument would still have to be pilot tested and finalized (as described above). Development and equating of new stories and other subtest items for additional periodic administrations also would involve similar effort and costs.
B	Implement EGRA to generate data needed to conduct a policy dialogue around the need for an EGR intervention*	Based on the objectives of the EGRA survey, the level of disaggregation and required sample size would have to be determined. Once this was done, the enumerators would have to be identified and trained to administer the test (some of these enumerators would likely be the same ones trained and used to pilot test the EGRA instrument). The number of enumerators needed would be determined by the size of the sample and the amount of time available for testing in the field. Once trained and equipped with the testing materials (i.e., the paper instrument or the tablet), the enumerators would be sent out in teams to administer the test. Once the test had been administered, the results would have to be analyzed and packaged for policy dialogue. If paper instruments were used, an additional step of data entry and data base compilation would be required.

[3] Phases B and C took place in some cases but not all. In those instances in which they did take place, their costs were included.

Phase		Description
C	Policy dialogue*	The usual policy dialogue session would involve bringing together education sector officials, policy makers, and other important stakeholders. The results of the EGRA would be shared and often, in a workshop-like setting, discussed with the objective of promoting attention and effort to improving reading outcomes in early grades. Often these policy dialogue events motivate a ministry of education and its key stakeholders to consider introducing an initiative to improve early grade reading.
D	Implement EGRA to generate a baseline in some pilot region	When committing to an early grade reading intervention of whatever scale, it would be absolutely necessary to first conduct a baseline assessment and analysis. Sometimes, the EGRA that contributed to the decision to launch a program could serve as the baseline. If not, because of the grades being targeted, or because of the geographic focus of the program, etc., a new baseline survey would need to be conducted. This would involve administering the assessment to both a representative sample of the control group of students and a representative sample of the treatment group of students. The same procedures described above would then be implemented.
E	Develop an EGR intervention*	To the extent that no appropriate reading curriculum already exists in the country, this phase would involve a number of international and host country reading and curriculum experts coming together to develop, for each language and for each grade level to be included in the intervention: a reading/literacy curriculum, the scope and sequence for the reading/literacy skills included in the curriculum, and a set of teacher and student materials based on the scope and sequence. The materials could include a teachers' manual with daily reading/literacy lesson plans (with varying degrees of programmed instructional content) and student textbooks mapped to daily lessons in the same manner as the teacher's manuals. Additionally, training materials (training manuals, training content and programs, etc.) would be developed for use in training trainers and teachers. Other materials would also be developed to train teacher coaches/support providers and to equip them with tools for observing and providing feedback to teachers and gathering information on implementation of the program.
F	Implement the EGR intervention	All of the materials that were developed for the EGR program would have to be produced and distributed such that every student, teacher, and coach received what the program required. Assuming that a cascade model of training was used, master trainers would have to be identified and trained to train trainers and trainers would have to be identified and trained to train teachers. Then, all of the training would have to take place. Coaches would also have to be identified and trained, but in many cases, they could be the same people who would have served as teacher trainers. If coaches were directly employed by the project implementing the intervention, their salaries would be included as part of implementation. If existing decentralized teacher support personnel were used as coaches, then the salary cost would not be included. In either case, the costs of coaches regularly visiting schools and facilitating teacher interaction would be part of implementation.
G	Implement an EGRA to generate an end-line	The EGRA would have to be administered once again to both the control group of students and the treatment group of students. The work done here (i.e., training, sampling, logistics) would be identical to the work done to carry out the baseline assessment, save for the facts that enumerator training could be less extensive if the enumerators were the same, and instrument development could have been completed in phase A.

*These are fixed costs in that once these activities have been accomplished, they need not be done again, and as such, they need not be paid for again should someone wish to take the EGR program to scale.

When this costing activity was initiated, a very detailed costing template was created to help guide respondents' efforts in acquiring the data.[4] Most people's reactions to the template were that it was far too detailed—that it was virtually impossible to obtain this level of data *post facto*. This being the case, they were subsequently asked to provide the data needed to fill in ***Table 2***.

Table 2. Data collection tool for estimating EGR program costs

	Phase	Technical			Administration	Total
		Labor	Materials	Other		
A	Develop EGRA					
B	Implement EGRA to generate data needed to conduct a policy dialogue around the need for an EGR intervention					
C	Lead policy dialogue					
D	Implement EGRA to generate baseline results in some pilot region					
E	Develop an EGR intervention					
F	Implement the EGR intervention					
G	Implement an EGRA to generate endline results					

The column labels in Table 2 were defined as follows:[5]

- **Technical:** All of the costs associated with the technical work that was done in order to carry out A through G.
  - **Labor:** all the technical *work* that was done to accomplish A through G (short-term technical assistance, long-term technical assistance, office staff who did technical work on A–G, and consultant labor).
  - **Materials:** all the technical EGR materials that were developed and used in the effort (EGRA instruments on paper or tablets; student textbooks, readers, workbooks; teacher workbooks and guides, flashcards, stopwatches; coaches' guides; and training materials).
  - **Other:** the non-personnel and non-materials costs of all technical workshops and training (per diems, travel costs, coffee breaks and lunches, rental

[4] This effort relied heavily on RTI staff who were very familiar with the project from a technical and/or administrative perspective and staff who knew the databases within which costs were accounted for.
[5] As per a USAID senior adviser, this cost structure reflects how USAID disaggregates its costs.

facilities and equipment); non-personnel and non-materials costs of conducting EGRAs (travel costs, per diems, lodging); non-personnel and non-materials costs of implementing the EGR program (transportation costs of the coaches, monitoring and evaluation costs); international and domestic travel, per diem, and hotel costs associated with technical labor.[6]

- **Administration:** all other non-technical costs associated with the effort, including the cost of the project office (if there was one); office staff not doing technical work; office equipment; housing for long-term technical advisors (where applicable); other long-term technical assistance costs where applicable (education allowances, Rest and Recreation allowances, etc.); utilities; office cars; home office staff not making a direct technical contribution to A–G but providing administrative, managerial, and technical backstopping.

2. Costs of Developing Evidence-Based EGR Programs

In this section of the document we offer, by country, a brief overview of the EGR efforts that were examined, the detailed costs of these programs (as per Table 2), the percent distribution of those costs relative to the overall cost, and notes on how the costs were derived.[7]

2.1 Cost Analyses by Country

Egypt

An evidence-based EGR program was developed in Arabic for grade 2 readers under the Girls' Improved Learning Outcomes (GILO) project. The effort unfolded in four of the 27 governorates in Egypt, benefitting 34,930 students and 8,101 teachers, according to the following timeline:

- **June 2008:** Develop the EGRA instrument
- **January–February 2009:** Conduct baseline EGRA
- **June 2009–September 2010:** Develop the EGR program
- **October 2010–April 2011:** Implement the EGR program
- **April–May 2011:** Conduct endline EGRA

The cost of developing and implementing this program amounted to only 12% of the total cost of GILO. These EGR costs are shown in ***Table 3.***

[6] "Other" costs could also contain some labor and some technical materials if these costs were folded into a cost item labeled "training."
[7] In ***Table 8*** we offer a more comprehensive comparative account of the major cost drivers of all of these EGR efforts.

Table 3. Costs of the EGR program in Egypt

Phases	Technical			Technical Total	Administration total	Overall total
	Labor	Other	Materials			
By U.S. dollar value						
Develop an EGRA	101,437	76,880	487	178,804	95,665	274,469
Use EGRA for baseline	150,283	101,629	2,902	254,814	136,333	391,147
Develop EGR intervention	787,343	355,790	0	1,143,134	611,609	1,754,743
Implement EGR intervention	58,929	688,934	279,643	1,027,506	549,745	1,577,251
Use EGRA post-intervention	150,283	101,629	2,902	254,814	136,333	391,147
Totals	**1,248,275**	**1,324,861**	**285,934**	**2,859,071**	**1,529,685**	**4,388,756**
By percentage						
Develop an EGRA	2%	2%	0%	4%	2%	6%
Use EGRA for baseline	3%	2%	0%	6%	3%	9%
Develop EGR intervention	18%	8%	0%	26%	14%	40%
Implement EGR intervention	1%	16%	6%	23%	13%	36%
Use EGRA post-intervention	3%	2%	0%	6%	3%	9%
Totals	**28%**	**30%**	**7%**	**65%**	**35%**	**100%**

Jordan

An Arabic language evidence-based EGR pilot program *is being* developed in Jordan for grades 1–3. It is projected that 10,986 students and 300 teachers will be impacted by the pilot intervention. The effort's timeline is as follows:

- **December 2011:** Adapt the EGRA instrument[8]
- **December 2011–August 2012:** Conduct baseline EGRA (data collection May, 2012)
- **September 2012–August 2013:** Develop the EGR program
- **September 2013–June 2014:** Implement the EGR program

[8] This adaptation exercise gathered nine Jordanian and international experts for five days to review an existing Arabic EGRA instrument and make all the changes necessary to render it specific to Jordan. It is much less expensive to adapt an existing EGRA than to create one from scratch.

- **May 2014–June 2014:** Conduct endline EGRA

Because this effort is still under way, some of the cost calculations in ***Table 4*** were based on actual expenditures, while those costs that have not yet been incurred were estimated from the budget.

Table 4. Costs of the EGR program in Jordan

Phases	Labor	Technical			Technical Total	Administration total	Overall total
		Other	Materials	Subs			
By U.S. dollar value							
Develop an EGRA	17,500	12,500	8,000	32,000	70,000	27,500	97,500
Use EGRA for baseline	20,000	40,000	15,500	62,000	137,500	32,500	170,000
Develop EGR intervention	25,000	77,500	33,000	132,000	267,500	42,500	310,000
Implement EGR intervention	67,500	137,500	46,500	186,000	437,500	95,000	532,500
Use EGRA post-intervention	22,500	45,000	17,000	68,000	152,500	35,000	187,500
Totals	**152,500**	**312,500**	**120,000**	**480,000**	**1,065,000**	**232,500**	**1,297,500**
By percentage							
Develop an EGRA	1.3%	1.0%	0.6%	2.5%	5.4%	2.1%	7.5%
Use EGRA for baseline	1.5%	3.1%	1.2%	4.8%	10.6%	2.5%	13.1%
Develop EGR intervention	1.9%	6.0%	2.5%	10.2%	20.6%	3.3%	23.9%
Implement EGR intervention	5.2%	10.6%	3.6%	14.3%	33.7%	7.3%	41.0%
Use EGRA post-intervention	1.7%	3.5%	1.3%	5.2%	11.8%	2.7%	14.5%
Totals	**11.8%**	**24.1%**	**9.2%**	**37.0%**	**82.1%**	**17.9%**	**100.0%**

Kenya

In Kenya, English and Kiswahili evidence-based EGR programs were developed to be implemented in nonformal, low-cost private schools in Nairobi and in government schools in the Central and Rift Valley provinces. This randomized controlled trial unfolded as per the following schedule:

- **October 2011:** Adapt the EGRA instrument
- **October–December 2011:** Develop the EGR program

- **January 2012:** Conduct baseline EGRA
- **January 2012–October 2013:** Implement the EGR program Years 1 and 2
- **October 2012:** Conduct midterm EGRA
- **October 2013:** Conduct endline EGRA
- **January–August 2014:** Implement the EGR program Year 3

The Years 1–2 interventions in the selected treatment schools benefitted 13,258 grade 1 and 2 students in 2012, 58,893 grade 1 and 2 students in 2013. With the control schools added in Year 3 (2014), PRIMR is estimated to benefit 134,000 grade 1 and 2 students in 2014 in both public and low-cost private schools. The total amount of instruction afforded to these PRIMR students through 2013 was truncated by strikes. As a result, in both 2012 and 2013, the beneficiary students received only 5 months of instruction out of the full academic year (10 months). The estimated costs of developing this evidence-based EGR program through August 2014 are shown in ***Table 5.*** Please note that the costs of all aspects of the effort that had not yet been incurred at the time of this writing were estimated from the 2013/14 work plan and budget, which delineated exactly how things will be spent during this period of time.

Table 5. Costs of the EGR program in Kenya

Phases	Technical			Technical Total	Administration total	Overall total
	Labor	Materials	Other			
By U.S. dollar value						
Develop an EGRA	32,149	0	15,601	47,749	25,042	72,791
Use EGRA for baseline	33,355	320	50,159	83,834	43,967	127,801
Develop EGR intervention	160,258	0	20,119	180,377	94,600	274,977
Implement EGR intervention	431,063	278,908	865,676	1,575,647	826,361	2,402,008
Use EGRA for midterm	25,589	10,381	47,707	83,676	43,885	127,561
Use EGRA for post-intervention	12,054	738	33,416	46,207	24,234	70,441
Totals	**694,467**	**290,346**	**1,032,677**	**2,017,490**	**1,058,089**	**3,075,579**
By percentage						
Develop an EGRA	1.0%	0.0%	0.5%	1.6%	0.8%	2.4%
Use EGRA for baseline	1.1%	0.0%	1.6%	2.7%	1.4%	4.2%
Develop EGR intervention	5.2%	0.0%	0.7%	5.9%	3.1%	8.9%
Implement EGR intervention	14.0%	9.1%	28.1%	51.2%	26.9%	78.1%

Phases	Technical			Technical Total	Administration total	Overall total
	Labor	Materials	Other			
Use EGRA for midterm	0.8%	0.3%	1.6%	2.7%	1.4%	4.1%
Use EGRA for post-intervention	0.4%	0.0%	1.1%	1.5%	0.8%	2.3%
Totals	**22.6%**	**9.4%**	**33.6%**	**65.6%**	**34.4%**	**100.0%**

Liberia

EGRA Plus: Liberia was an English-speaking grade 2 and 3 evidence-based EGR program that was developed in Liberia during the period 2008–2010, as per the timeline shown here:

- **March–June 2008:** Develop the EGRA instrument
- **March–June 2008:** Conduct baseline EGRA
- **June–October 2008:** Development the EGR program
- **October 2008–April 2010:** Implement the EGR program
- **May 2009:** Conduct midterm EGRA
- **May 2010:** Conduct endline EGRA

EGRA Plus benefited 2,825 students[9] and 762 teachers in 60 schools selected from 7 counties. This was a randomized controlled trial with the following design:

- The control group did not receive any EGR intervention (an additional 60 schools).
- In the "full" treatment group (60 schools): Teachers were trained on how to continually assess student performance; teachers were provided frequent school-based pedagogic support, resource materials, and books; and parents and communities were informed of student performance.
- In the "light" treatment group (an additional 60 schools): The community was informed about reading achievement using school report cards based on EGRA results or findings and student reading report cards prepared by teachers. Teachers in these schools received only support on how to assess, track, and communicate student reading results to parents and parent–teacher associations.

The costs of this work are shown in ***Table 6*** on the following page.

[9] Across all 180 schools in the pilot effort—60 of which were full treatment, 60 of which were light treatment, and 60 of which were control schools—there were 8,474 students. For costing purposes, we assumed that these students were evenly distributed among all three sets of schools, resulting in 2,825 students in each set of schools. Since neither those in the control group nor those in the light treatment group benefited from the full treatment, we assumed that only 2,825 did—those in the full treatment schools.

Table 6. Costs of the EGR program in Liberia

Phases	Technical			Technical total	Administration total	Overall total
	Labor	Materials	Other			
By U.S. dollar value						
Develop an EGRA	1,724	11,418	3,806	16,948	2,489	19,437
Use EGRA results for dialogue	5,173	34,253	11,418	50,844	7,466	58,310
Use EGRA for baseline	45,465	7,511	145,283	198,259	26,065	224,324
Develop EGR intervention	27,071	68,203	22,734	118,008	17,328	135,336
Implement EGR intervention	241,390	214,706	639,333	1,095,429	160,852	1,256,281
Use EGRA for midterm	45,465	7,511	114,151	167,127	26,065	193,192
Use EGRA for post-intervention	45,465	7,511	114,151	167,127	26,065	193,192
Totals	**411,753**	**351,113**	**1,050,875**	**1,813,741**	**266,330**	**2,080,071**
By percentage						
Develop an EGRA	0.08%	0.55%	0.18%	0.81%	0.12%	0.93%
Use EGRA results for dialogue	0.25%	1.65%	0.55%	2.44%	0.36%	2.80%
Use EGRA for baseline	2.19%	0.36%	6.98%	9.53%	1.25%	10.78%
Develop EGR intervention	1.30%	3.28%	1.09%	5.67%	0.83%	6.51%
Implement EGR intervention	11.60%	10.32%	30.74%	52.66%	7.73%	60.40%
Use EGRA for midterm	2.19%	0.36%	5.49%	8.03%	1.25%	9.29%
Use EGRA for post-intervention	2.19%	0.36%	5.49%	8.03%	1.25%	9.29%
Totals	**19.80%**	**16.88%**	**50.52%**	**87.20%**	**12.80%**	**100.00%**

Malawi

A Chichewa evidence-based EGR program was developed for grades 2 and 3 as per the following schedule:

- **September 2010:** Adapt the EGRA instrument
- **November 2010:** Conduct baseline EGRA

- **January 2011:** Develop the EGR program
- **January 2011–November 2012:** Implement the EGR program
- **November 2011:** Conduct midterm EGRA
- **May 2012:** Conduct EGRA snapshot of progress
- **November 2012:** Conduct endline EGRA

In the first year of EGR implementation, 46,692 students in two districts benefited. In the second year, 218,177 students from an additional five districts were included, yielding a total of 264,869 students in all seven districts receiving the intervention.[10] Two sets of EGRAs were conducted, one national and one that focused on the students in the two original test-site districts. The overall costs of this effort are shown in ***Table 7***.

Table 7. Costs of the EGR program in Malawi

Phases	Technical		Total technical	Administration total	Overall total
	Labor	Other			
By U.S. dollar value					
Develop an EGRA	98,171	29,198	127,369	11,938	139,306
Use EGRA for baseline	205,183	228,168	433,352	34,179	467,530
Develop EGR intervention	125,597	25,107	150,704	125,449	276,153
Implement EGR Intervention	320,648	217,064	537,712	386,490	924,202
Conduct midterm EGRA	105,754	212,351	318,105	22,847	340,953
Conduct snapshot EGRA	58,560	186,485	245,045	16,455	261,500
Use EGRA for post-intervention	110,975	220,747	331,722	23,860	355,582
Totals	**1,024,889**	**1,119,120**	**2,144,009**	**621,218**	**2,765,226**
By percentage					
Develop an EGRA	3.6%	1.1%	4.6%	0.4%	5.0%
Use EGRA for baseline	7.4%	8.3%	15.7%	1.2%	16.9%
Develop EGR intervention	4.5%	0.9%	5.4%	4.5%	10.0%
Implement EGR intervention	11.6%	7.8%	19.4%	14.0%	33.4%
Conduct midterm EGRA	3.8%	7.7%	11.5%	0.8%	12.3%
Conduct snapshot EGRA	2.1%	6.7%	8.9%	0.6%	9.5%

[10] Malawi has a total of 34 districts.

Phases	Technical		Total technical	Administration total	Overall total
	Labor	Other			
Use EGRA for post-intervention	4.0%	8.0%	12.0%	0.9%	12.9%
Totals	**37.1%**	**40.5%**	**77.5%**	**22.5%**	**100.0%**

2.2 Discussion of Costs

The cost of developing these evidence-based EGR programs ranged from $4,388,756 to $1,297,500, differing by a factor of 3.38. That these costs differ by so much can be attributed to the fact that the programs *themselves* differed, as can be seen by the information presented in ***Table 8***.

Table 8. Key elements of the EGR efforts that most affected technical cost differences among the programs

Program element	Egypt	Jordan	Kenya	Liberia	Malawi
Total cost	**$4,388,756**	**$1,297,500**	**$3,075,579**	**$2,080,244**	**$2,765,226**
Grades targeted	2	1–3	1–2	2–3	1
Languages	Arabic	Arabic	English, Kiswahili	English	Chichewa
Lessons developed	56	160	300	110	80
Period of performance	2010–2011	2012–2014	2011–2014	2008–2010	2010–2012
Treatment group (students)	34,930	10,986	Y1: 13,258 Y2: 58,893 Y3:134,000[11] ∑:206,151	2,825	Y1: 46,692 Y2: 218,177 ∑: 311,561
Estimated size of pilot—distinct beneficiary students as a percentage of the total enrollment for those grade(s)	2.4%	2.4%	6.1%	2.1%	22.7%
Number of materials developed and provided (books, manuals, sets of materials, tools)					
for coaches	1	1	5	12	1
for teachers	7	4	8	7	6

[11] This is a projected value.

Program element	Egypt	Jordan	Kenya	Liberia	Malawi
for students	1	3	4	3	1
Number of training sessions for each: # days session 1, # days session 2, # days session 3, etc.)					
Teachers	Y1:3,3	Y1: 10,5 Y2: 5	Y1:5,3,2	Y1:7,3 Y2:7,3	Y1:3,3,2
Coaches	Y1:3,3	Y1:10,5	Y1:9,4,3	Y1:7,7 Y2:7,7	Y1:3,3,3
Training ratios: teachers per trainer	90	15:1	20:1	24:1	45:1
Coach visits per month per school	0.8	2	2	1	1.2
Time teaching (minutes per day/days per week)	25:5	NA	30:5	45:5	60:5
Total time of EGR instruction offered (months)	7	9.5	14	13.5	NA
Number of EGRAs administered	3	2	3	3	3
1st EGRA					
# schools	59	156	230	176	33
# students tested	2,900	3,063	4,385	2,988	976
Paper/tablet	Paper	Tablet	Paper	Paper	Paper
2nd EGRA					
# schools	59	156	210	175	20
# students tested	1,783	3,063	4,162	2,805	210
Paper/tablet	Paper	Tablet	Tablet	Paper	Tablet
3rd EGRA					
# schools	200	NA	210[12]	175	33
# students tested	1,992	NA	4,222	2,688	1,322

[12] Results of the October 2013 endline were being tallied as this report was being finalized. The exact number of schools ultimately to be counted as part of the endline sample is being determined; this figure is an estimate.

Program element	Egypt	Jordan	Kenya	Liberia	Malawi
Paper/tablet	Tablet	NA	Tablet	Paper	Tablet
Libraries	No	No	Yes[13]	Yes	No
Social mobilization	No	Yes	Yes	Yes	Yes[14]
Administration	$1,529,685	$232,500	$1,058,089	$266,328	$621,218[15]

NA = not applicable.

Of note in Table 8 is the range one can find among a number of aspects of these programs. Some unfolded over the course of just more than a year, while others over two to three years. The number of students impacted in Liberia was only 2,825, while in Malawi 311,561 were impacted. In Egypt only 56 lessons were developed, while 300 lessons were developed in Kenya. In both Jordan and Malawi, eight types of EGR materials were developed/used while EGRA Plus: Liberia developed/used 22. The cost of administration in Egypt was estimated to be $1,529,685, while in Liberia this cost was $266,328. That these programs varied quite widely in a number of ways is why their overall costs vary so much.

Reflecting back on Tables 3–7 (and summarized in ***Table 9*** below), one can see that with the exception of Egypt, the cost of implementing the EGR intervention was the most expensive phase. (In Egypt, it proved to be the second most expensive phase—only 4 percentage points less expensive than developing the EGR intervention.)

Table 9. Percent distribution of total costs across all phases, by country

Phase	Egypt	Jordan	Kenya	Liberia	Malawi
Develop an EGRA	6.0	7.5	2.4	0.9	5.0
Use EGRA for policy dialogue	NA	NA	NA	2.8	NA
Use EGRA for baseline	9.0	13.1	4.2	9.8	16.9
Develop EGR intervention	40.0	23.9	8.9	6.5	10.0
Implement EGR intervention	36.0	41.0	78.1	60.4	33.4
Conduct midterm EGRA	NA	NA	4.1	9.8	12.3
Conduct snapshot EGRA	NA	NA	NA	NA	9.5

[13] Ten books—not a library—were put into every classroom.
[14] These costs were not in the costing analysis.
[15] Recall that this cost excludes all the administrative costs of the prime awardee, CAII, including the DCOP-related costs.

Phase	Egypt	Jordan	Kenya	Liberia	Malawi
Use EGRA post-intervention	9.0	14.5	2.3	9.8	12.9
Total	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**

NA = not applicable.

That the implementation of the EGR program proved to be the most expensive phase for most countries (with an average percent cost of 49.78% of the total percent cost across all five programs) should not come as a surprise, given the costs that are associated with training and the cost of producing and distributing the materials needed by all the students, teachers, and coaches. Of interest is the fact that EGRA instrument development proved to be the least costly activity, with an average of 4.36% of the total percent cost across all five programs.

Looking at the ratio of technical costs to administration costs in ***Table 10*** below, one can see that technical costs on average outweigh administration costs by a factor of 3 to 1. Jordan and Liberia had very low administration costs due largely to the fact that they had no project offices and as such, no COP, no office staff, and no office expenses.

Table 10. Percent distribution of technical costs and administration costs by country

Cost category	Egypt	Jordan	Kenya	Liberia	Malawi
Technical	65	82	66	87	77
Administration	35	18	33	13	23
Technical/Administration	1.86	4.56	2.00	6.69	3.35

When one looks at the distribution of costs across all technical components—labor, materials, and other—one can see in ***Table 11*** that there is a fair amount of variance within each technical category across all the countries. Labor costs in Egypt, Kenya, and Malawi were quite high compared to both Jordan and Liberia, while "other" costs in Kenya and Malawi were quite high compared to Egypt, Jordan, and Liberia. The cost of materials varied quite widely as well. That the labor costs for Egypt, Kenya, and Malawi were relatively high can be attributed in part to the fact that in all three of these countries, both the EGRA instruments and the EGR programs were built from scratch: Arabic, Kiswahili, and Chichewa. By contrast, in both Jordan and Liberia, the instruments and many of the intervention materials were adapted from previously created ones. Kenya's, Liberia's, and Malawi's high "other" cost can be attributed to the fact that they did a lot of training, while Kenya's relatively low materials cost can be attributed to the fact that

they purposely tried to keep these costs down to match or beat what the ministry was spending per textbook.

Table 11. Percent distribution of technical costs across labor, materials, and other by country

Technical cost category	Egypt	Jordan	Kenya	Liberia	Malawi
Labor	43.7	26.1	34.4	22.7	47.8
Materials	46.3	53.4	14.4	19.4	NA
Other	10.0	20.5	51.2	57.9	52.2

NA = not available.

The variability in *overall* costs therefore does not necessarily reflect dramatic differences in the cost-efficiency of the projects, but by and large is an artifact of project design. Since projects are designed in part with the amount of funding available in mind, it would be very unwise to assume that the cost of one project necessarily indicates that another project could have accomplished the same results at a lower (or higher) cost.

To further clarify the costs of these programs, the next section of this document examines project unit costs and then uses those unit costs to extrapolate indications of what scale-up of such programs could cost in each context.

3. Unit Costs and Estimated Scale-Up Costs

Two unit costs are presented in ***Table 12*** below. The first takes the total cost of each program and divides that cost by the number of students who were treated over the time in which those costs were incurred. This yields what is referred to in the table as the *total unit cost*. The second takes the technical cost of implementing the EGR program and divides that cost by the number of students who were treated over the time in which those costs were incurred. This yields what is referred to in the table as the *unit technical cost of implementing EGR*. The scale-up costs shown in Table 12 were derived by taking the second unit cost—the *unit technical cost of implementing EGR*—and multiplying it by the respective national enrollment of the treatment group.[16] For example, if the treatment group was grade 1, the national grade 1 enrollment was used to estimate the scale-up cost.

[16] We used this unit cost in our scale-up calculation because the fixed costs associated with the total unit cost are not included.

Table 12. Results of unit-cost estimation

Cost factor	Egypt	Jordan	Kenya	Liberia	Malawi
Enrollment treatment group	34,930	10,986	206,151	2,825	311,561
Technical cost of implementing EGR pilot	$1,027,506	$437,500	$1,575,647	$1,095,428	$537,712
Overall cost of the EGR pilot	$4,388,756	$1,297,500	$3,075,579	$2,080,068	$2,765,226
Total unit cost	$125.64	$118.10	$14.92	$736.31	$10.44
Unit technical cost of implementing EGR	$29.42	$39.82	$7.64	$387.76	$2.03
National enrollment	1,440,335	463,361	2,200,279	135,397	960,126
Scale-up costs (technical)	$42,369,096	$18,452,616	$16,817,105	$52,501,843	$1,949,157

As one can see, the *total unit costs* vary quite widely, with a low of $10.44 in Malawi and a high of $736 in Liberia (~70 times higher).[17] That these unit costs vary so greatly can be attributed to both the cost of the programs (the numerator) *and* the size of the treatment groups (the denominator).[18] As noted in the previous section, the cost of each program varies by a factor of 3.38, but the size of the treatment groups varies even more so, by a factor of 93.7. As the denominator increases, the fixed costs do not; accordingly, the unit cost is driven downward as this happens. Moreover, as the denominator increases, the non-fixed costs will decrease as some economies of scale are realized—again, driving the unit cost downward as this happens. Had the size of the treatment group in Liberia been much larger, the unit cost would surely have been much smaller.

The scale-up costs range from over $42M to just less than $2M, attributable to both the differing *unit technical cost of implementing EGR* and enrollment. We note that the scale-up costs put forth here are indicative at best. In the first instance, they do not account for the economies of scale that can be realized as one takes a pilot to scale. Nor do they

[17] That Liberia proved to be so expensive on a unit-cost basis can be attributed to a number of factors. By way of example, the cost of printing student and teacher materials during scale-up (which is now being carried out by RTI and FHI 360 under the USAID Liberia Teacher Training Program: 1,600 schools, 2010–2015) was about one-fifth as expensive as during the pilot, because during scale-up, the project team found an international printer that could print the books at a fraction of what it cost to print it locally, inclusive of the shipping costs. In addition, during the pilot, the coach-to-school ratio was 1:4; during scale-up it was increased to 1:12. The point is that with EGRA Plus being one of the first EGR efforts carried out by USAID, a lot of work, and costs, were put into it.

[18] Again, it must be noted that in Kenya, a conscious effort was made to drive program costs down; one explicit focus of the program was measuring cost-effectiveness in preparation for possible scale-up.

account for the cost of addressing the various policy, institutional, systemic, and reform support activities that may have to be undertaken to ensure that what has been taken to scale can be sustained. Finally, what is piloted is often times not what is taken to scale. Most pilots succeed because a lot of resources are directed at them to ensure that they do succeed, but as efforts are made to take them to scale, measures are taken to make them more scalable.

Finally, the program scale-up costs offered in Table 12 do not include the EGRA costs, and as such, they are lower than what they likely would be in actuality. While we could generate a unit cost for implementing an EGRA, the question then would become, what enrollment figure would one multiply it by to project its scale-up cost? It is not as simple as projecting the scale-up cost of implementing the EGR intervention because one need not apply the unit cost to the national-level enrollment—that is, not every child needs to be tested in an EGRA survey, only enough to offer a nationally (or subnationally) representative picture. Since without knowing the details of an appropriate sampling design for each context (including, for example, the levels of disaggregation that would be needed), we do not know what that number is, and thus we do not offer these scale-up costs.[19]

4. Cost Effectiveness

For these programs to be "evidence-based," they need to have been empirically proven to have an effect on the participating students. That "effect" is measured by creating a control group, a set of students who will not be involved in the EGR program; and a treatment group, a set of students who will receive the intervention through the EGR program. Both groups are then tested before the EGR program is implemented, thereby creating a baseline set of EGR test scores; and both groups are tested after the EGR program, creating an endline set of EGR test scores (in some instances, both groups may be tested in the middle of an EGR program's implementation, creating a midterm set of EGR test scores). Both sets of students are tested across a number of EGR skills, which are defined as follows:

- **Letter sound or syllable reading[20]:** students' ability to read the *sounds* of either letters or legal syllables (e.g., ba, ka, tu, mo), depending on the characteristics of the language, from a list of 50–100 such letters or syllables in one minute's time.

[19] Standard practice for EGRAs assesses 400 students per unit of disaggregation, so if one wished to carry out a nationally representative sample with no disaggregation, one would need to assess only 400 students. If one wanted to look at a rural/urban nationally representative sample, then one would need to sample 800 students. If one wanted to discern how each state was doing and there were 20 states, then one would have to sample 8,000 students. What this means is that the scale-up costs of conducting an EGRA could be quite small if all that one feels is needed is a nationally representative sample. It also means that it could become rather expensive if one wanted to have a high level of disaggregation.

[20] Of the languages of assessment in the countries covered in this analysis, two (English and Arabic) are considered to be letter-based; Chichewa and Kiswahili are more syllable-based. The letter-sound and syllable-reading subtests measure similar skills and generally one or the other is selected for the EGRA based on the language characteristics.

Student performance is based on the number of correct letters sounds or syllables that were read during this time.

- **Nonword reading:** students' ability to read nonwords (i.e., made-up words) from a list of 50 one- and two-syllable nonwords (i.e., ga, ot, tog, popo). This is a commonly used methodology for testing students' ability to decode unfamiliar words, as distinct from their ability to sight-read familiar ones. Students are asked to read as many such nonwords as they can in one minute. Student performance is based on the number of correct nonwords read during this time.
- **Word reading:** students' ability to read a list of grade-appropriate words from a list of words presented to them in 10 rows with five words on each row (e.g., go, cat, pat, mama). They are asked to read as many words as they can in one minute. Student performance is based on the number of correct words read during this time.
- **Oral reading fluency:** students' ability to read a narrative story with accuracy within one minute. Student performance is based on the number of correct words read during this time.
- **Reading comprehension:** Students' ability to answer a number of questions about the story they just read. Student performance is based on the number of correct answers that were given.

The impact that these EGR programs had on the treatment students, compared to their respective control students, is shown below for each country in this study except for Jordan where, because the EGR intervention is still being implemented, a post-intervention EGRA has not been conducted. Also shown below are a number of cost-effectiveness calculations. These calculations were derived in the following manner:

- For each reading skill for which the control and treatment groups were tested (i.e., letter sound or syllable reading, word reading), their baseline weighted average scores and endline weighted average scores were obtained.
- The gains (endline scores minus baseline scores for each reading skill) were calculated for each group.
- The gain of the treatment group was then subtracted from the gain experienced by the control group to yield the overall net gain of the treatment group over the control group for each reading skill measured. *This represents the impact made by the EGR intervention*.[21]
- The overall net gain of each reading skill was then factored into the cost-effectiveness calculation using this equation: [22]

[21] If the EGR intervention created a gain of 10 over baseline scores, and the control group that had no intervention had a gain of 3 over baseline, the gain that can be attributed to the EGR intervention is the gain made by the treatment group minus the gain made by the control group: the net gain.

[22] We realize that cost effectiveness values are ordinarily calculated with the cost in the denominator. We have chosen to use the inverse of this value simply because the resulting value makes a bit more intuitive sense for the reader (i.e., $7.32 per person net gain versus 0.14 per person net gain per dollar).

$$CE = \frac{Cost}{Pupil * Net\ Gain}$$

- $Cost^A$ = overall cost of the EGR program
- $Cost^B$ = technical cost of implementing the EGR intervention[23]
- Pupil = total number of students in the treatment group[24]
- Net gain = the overall net gain described above

The two values used for the numerator—$Cost^A$ and $Cost^B$—allow us to calculate the total cost per pupil net gain (using $Cost^A$), presented in the second to last column; and the technical cost of implementing the EGR program per pupil net gain (using $Cost^B$), presented in the last column of the tables for each country.

4.1 Cost-Effectiveness Calculations by Country

Below, these calculations are applied to four of the countries for which we have cost and baseline and endline data: Egypt, Kenya, Liberia, and Malawi.

Table 13 summarizes the data that went into the above-described calculation for Egypt. For example, the gain between baseline and endline in Arabic letter sounds in the control group was 1.6 more letter sounds read correctly per minute. For the treatment group, the

Table 13. Cost-effectiveness calculations for Egypt

EGRA subtest	Weighted averages of all scores						Treatment – control: Net gain	Total cost per pupil net gain	Imp. EGR cost per pupil net gain
	Control group			Treatment group					
	Base	End	Gain	Base	End	Gain			
Letter sound reading (correct sounds per minute)	8.6	10.1	1.6	9.8	28.5	18.7	17.2	$7.32	$2.63
Word reading (correct words per minute)	5.6	7.5	1.9	7.4	15.5	8.2	6.3	$20.07	$7.21
Oral reading fluency (correct words per minute)	8.9	10.9	2.0	11.1	21.1	10.1	8.0	$15.63	$5.62

[23] We assert that the cost effectiveness values generated from this cost are more "comparable" than those generated by $Cost^A$ because they factor out a number of variable costs that would render the cost effectiveness values more "incomparable."

[24] Only these students factor into the cost-effectiveness calculation because they are the ones on whom the intervention is having its effect—the ones who are being treated.

gain from baseline to endline was 18.7 more letter sounds per minute. Therefore, the net gain is the difference between these two, or 17.2 (with rounding). The total number of students in the treatment schools was 34,930, leading to the *total cost* per pupil net gain of $7.32 and a *technical cost of implementing the EGR program* per pupil net gain of $2.63. This can be interpreted as meaning that in Egypt it cost a total of $7.32 for every pupil that experienced a net gain of 17.2 more letter sounds read per minute. Each of the subsequent country tables ***(Tables 14–16)*** is laid out in the same manner.

Table 14. Cost-effectiveness calculations for Kenya[25]

EGRA subtest	Control group			Treatment group			Treat-ment – control: Net gain	Total cost per pupil net gain	Imp. EGR cost per pupil net gain
	Base	End	Gain	Base	End	Gain			
English									
Letter sound reading (correct sounds per minute)	16.6	20.4	3.8	21.6	42.4	20.8	17.0	$0.88	$0.45
Nonword reading (correct words per minute)	5.4	14.4	9.0	7.0	21.4	14.4	5.4	$2.76	$1.42
Oral reading fluency (correct words per minute)	4.0	18.8	14.8	6.8	29.3	22.5	7.7	$1.94	$0.99
Reading comp. (% items correct)	5.8	20.2	14.4	6.6	21.8	15.2	0.8	$18.65	$9.55
Kiswahili									
Letter sound reading (correct sounds per minute)	11.6	26.4	14.8	14.1	45.7	31.6	16.8	$0.89	$0.45
Nonword reading (correct words per minute)	2.8	9.6	6.8	3.5	13.6	10.1	3.3	$4.52	$2.32
Oral reading fluency (correct words per minute)	3.1	14.9	11.8	4.7	21	16.3	4.5	$3.32	$1.70
Reading comp. (% items correct)	14.2	35.6	21.4	15.3	43.3	28	6.6	$2.26	$1.16

[25] The impact data shown in Table 14 are for the intervention that unfolded in 2012. The costs given earlier in Table 4 were for the entire period, 2012–2014. The estimated costs used in the cost-effectiveness analyses had to be the costs incurred in 2012, the start-up period, to account for EGR program development. These 2012 costs (total and technical implementation of the EGR intervention) were estimated by calculating the unit costs (total and technical implementation of the EGR intervention) over the period 2012–2014 and multiplying each one by the enrollment in 2012—the enrollment of pupils for whom the impact calculations are provided.

Table 15. Cost-effectiveness calculations for Liberia

EGRA subtest	Control group			Treatment group			Treatment – control: Net gain	Total cost per pupil net gain	Imp. EGR cost per pupil net gain
	Base	End	Gain	Base	End	Gain			
Letter sound reading (correct sounds per minute)	27.9	40.8	12.9	29.6	47.7	18.1	5.2	$141.60	$74.57
Word reading (correct words per minute)	10.5	18.1	7.6	11.7	30.6	18.9	11.3	$65.16	$34.32
Nonword reading (correct words per minute)	6.0	6.7	0.7	6.3	16.0	9.7	9.0	$81.81	$43.08
Oral reading fluency (correct words per minute)	7.3	17.5	10.2	8.4	28.2	19.8	9.6	$76.70	$40.39
Reading comp. (% items correct)	24.0	43.0	19.0	26.0	64.0	38.0	19.0	$38.75	$20.41

Table 16. Cost-effectiveness calculations for Malawi[26]

EGRA subtest	Control group			Treatment group			Treatment – control: Net gain	Total cost per pupil net gain	Imp. EGR cost per pupil net gain
	Base	End	Gain	Base	End	Gain			
Syllable reading (syllables correct per minute)	0.6	0.6	0.1	0.6	16.2	15.6	15.6	$0.67	$0.13
Word reading (words correct per minute)	0.4	0.2	-0.2	0.4	10.4	10.1	10.2	$1.02	$0.20
Nonword reading (words correct per minute)	0.2	0.2	-0.0	0.2	6.2	6.0	6.0	$1.74	$0.34

[26] The impact results were derived from students in the original two districts. The costs in Table 6 earlier were for the work done in all seven districts over the course of two years (five districts were added in Year 2). The costs associated with just these two districts were estimated by calculating the unit costs (total and technical implementation of the EGR intervention) for the entire effort and multiplying each one by the enrollment in the two districts. These estimated costs were used in the cost-effectiveness calculations.

EGRA subtest	Control group			Treatment group			Treatment – control: Net gain	Total cost per pupil net gain	Imp. EGR cost per pupil net gain
	Base	End	Gain	Base	End	Gain			
Oral reading fluency (words correct per minute)	0.3	0.2	-0.1	0.3	8.1	7.8	8.0	$1.31	$0.25
Reading comp. (% items correct)	0.0	0.0	0.0	0.0	11.0	11.0	11.0	$0.95	$0.18

4.2 Discussion of Cost Effectiveness

To facilitate a discussion on the relative cost effectiveness of these programs, we created a summary table (***Table 17***) that shows the cost-effectiveness values for the "technical cost of implementing the EGR intervention" (the last column in Tables 13–16) vis-à-vis the pupil net gains made in the literacy skill areas tested in all four programs.

Table 17. Summary of cost-effectiveness values[27]

EGRA subtest	Implementation of EGR (technical costs) per pupil net gain				
	Egypt	Kenya (English)	Kenya (Kiswahili)	Liberia	Malawi
Letter sound or syllable reading	$2.63	$0.45	$0.45	$74.57	$0.13
Oral reading fluency	$5.62	$0.99	$1.70	$40.39	$0.25

It is clear that the Malawian effort achieved the same effect (person net gain for syllable reading and oral reading fluency) as the other efforts *for much less money*. This is valuable information in that it gives one reason to look into what exactly unfolded in Malawi (and, at the opposite end of the spectrum, what unfolded in Liberia) to achieve these results. Mathematically speaking, the lower the number in Table 17, the more cost-effective the intervention is, and what drives these cost-effectiveness numbers downward are the following: lower costs, higher enrollment, and higher net gains. As one can see in ***Table 18***, where these three pieces of information are provided for each country (for just oral reading fluency), Liberia had the highest net gain (1.25 times higher than the lowest, and 1.20 times higher than Malawi), but its high cost (11.56 times higher than the

[27] Letter sound or syllable reading, word reading, and oral reading fluency only are offered in this comparative table because they are the only three reading indicators common across all the countries being analyzed.

lowest—Malawi) and very small enrollment (the lowest of the four by a factor 16.53 compared to the highest) both rendered the effort the least cost-effective of the four.

Table 18. Isolation of factors contributing to cost effectiveness

Factor	Egypt	Kenya	Liberia	Malawi
Technical cost of implementing EGR	$1,027,506	$101,333	$1,095,428	$94,790
*Factor higher than **lowest gain***	*10.84*	*1.07*	*11.28*	***1.00***
Enrollment treatment group	34,930	13,258	2,825	46,692
*Factor higher than **lowest gain***	*12.36*	*4.69*	***1.00***	*16.53*
Oral reading fluency net gain	8	7.7	9.6	8
*Factor higher than **lowest gain***	*1.04*	***1.00***	*1.25*	*1.04*

Clearly, a major issue in Liberia was a lack of economies of scale, giving USAID good reason not to design such small efforts in the future. Liberia was also the first USAID intervention designed to improve reading outcomes. Subsequent projects benefited from the Liberia experience in designing their approaches. As noted earlier, we also know that the unit costs of the materials that were used in the Liberian pilot were four to five times more expensive than the unit costs of the very same materials that are being used in the subsequent scale-up effort under the Liberia Teacher Training Program (which will produce its own analysis of the impact of the large-scale application of the model). Additionally, the school/coach ratio that is being used in the Liberian scale-up effort is three times higher than the one used in the pilot (12/1 versus 4/1). But in reducing these costs during scale-up, one wonders what effect it will have on the overall efficacy of the Liberian EGR program. One can posit that the reduced unit cost of the materials will have virtually no effect, but that the increased school/coach ratio could have a serious negative impact on program efficacy.

As useful as these cost effectiveness numbers are, we caution readers about taking them *too* far—that is, putting too much emphasis on them. Inasmuch as the cost figures are incomparable, one could say that these cost-effectiveness figures are equally incomparable. Granted, we have tried to factor out key elements of the costs that render them incomparable by using only the technical cost of implementing the EGR intervention in the cost–effectiveness calculations presented in Table 14, and so they do have a fair amount of comparability. Nevertheless, while one can indeed say that the Malawian effort was more cost effective than the Liberian effort, it is not as though the Minister of Education in Liberia can take the Malawian EGR program, implement it in exactly the same way it was implemented in Malawi, and expect to get the same oral reading fluency impact for 1/100th the cost of the effort that unfolded in Liberia. To do

this sort of thing, a number of low-cost variants would have to be developed in Liberia and their cost-effectiveness values would have to be determined.[28] If two such variants were developed, say, EGR1 and EGR2, and their oral reading fluency cost effectiveness values were determined to be $34.17 and $15.19 respectively (as compared to $40.39 for the original EGR effort), then indeed, the Minister of Education in Liberia could say that by implementing EGR2, the same oral reading fluency results could be achieved for less than half the cost of implementing the original EGR program.

5. Recommendations

If USAID is keen on examining the cost of developing evidence-based EGR programs, it must ensure that the data needed to examine the costs of these programs is readily available. Of all the things that have come out of this activity, one of the most glaring is how hard it is to collect this data *post facto*. Implementing partners account for costs in a variety of ways, depending on both the nature of the contract/agreement and their own proprietary accounting systems. Unless the implementing partners are asked (*forced* by the contract/agreement) to track data at the level of detail needed to do these cost analyses, USAID and its implementing partners will face in the future the very same set of problems we both faced in this activity. This means that USAID should require its implementing partners to track data and costs to the level of detail offered in Table 2. With the one exception of Phase F, implementation of the EGR intervention, we feel that this level of detail will suffice.

For Phase F, the implementation of the EGR intervention, we recommend that a much greater level of detail be captured.[29] Every EGR program has what one might call its distinct recipe—the particular mix of inputs and processes that actually produce the net gains evidenced in the treatment group as compared to the control group. This recipe would include the teacher–class ratios; time spent each day teaching reading; the particular lessons that are being used over the course of the treatment; the number and nature of the materials used by students, teachers, and coaches; the teacher–trainer ratios; the master-trainer-to-trainer ratios; the amount of training time each person gets; the teacher- and/or school-to-coach ratios; the amount of time every teacher spends with a coach over the course of a school year; etc. The particulars of these elements of the EGR program are what ultimately create the impact and as such they must be tracked along with their associated costs.

The reason why it is so important to gather these details for this phase of the EGR programs is that this recipe is what informs the challenge of going to scale. Should it prove to be too costly or physically impossible to take to scale (e.g., the pilot relied heavily on local nongovernmental organizations [NGOs] to do the training and they are

[28] A lower-cost variant might have few materials for the teacher, lower unit costs of the materials, higher teacher-to-trainer ratios, and fewer coaching visits.

[29] A template that can be used to capture these costs is a deliverable of this costing activity and will be made available along with the final report of the study.

either too expensive to take to scale and/or there simply are not enough local NGOs in the country to do what was done during the pilot), then measures may have to be taken to alter the recipe to render it more affordable/doable. Knowing the exact recipe and its associated costs allow one to alter the program in an informed manner and to have some sense of the impact those alterations might have on the overall efficacy of the program.[30] Short of having cost-effectiveness studies in one's hands (see next paragraph), one can alter the recipe of a program in the presence of technical experts and gain some sense of how the alteration(s) will impact program efficacy. For example, reducing the unit costs of the materials used (i.e., printing them on less expensive paper) may, in the view of the reading experts, not have any impact at all on program efficacy, whereas reducing the number of training days by half could, in the minds of these experts, decrease overall efficacy by as much as 50%.

It would help considerably, however, to have real cost-effectiveness information at one's fingertips, and for this to happen, one needs to design and implement cost-effectiveness studies—pilots aimed at looking at the relative cost-effectiveness of various EGR recipes, studies that strive to determine the exact impact that reduced training and/or reduced coaching might have on program efficacy. Should USAID choose to carry out these studies, we recommend that they be designed with scale-up and sustainability in mind—examining the cost effectiveness of models that countries can afford to take to scale, and which the HCG can afford to sustain once taken to scale.

This being the case, our last recommendation is to do the work necessary to examine what it costs to sustain an EGR program once taken to scale and to determine the extent to which the HCG can afford it. Given the right data, computerized, demographically driven enrollment, input, and cost projection models can make it feasible to examine the costs of sustaining an EGR program over an extended period of time, especially if one has the detailed costs of that program's "recipe." Whether the HCG can afford to assume these costs will depend on how much money it currently spends on education and how that money is currently spent. If the government is not spending enough money on education, then measures must be taken to address this issue—It simply isn't possible for an HCG to sustain a viable EGR effort if it lacks the overall funds needed to do so. If there are sufficient funds, and the situation is such that they are not spending those funds wisely, then a computerized model of the type mentioned above could be used to examine how that money is currently being spent and identify areas where it is it not being well spent. This would allow decision-makers to consider taking the measures needed to syphon funds away from those areas where they are not being well spent and channel them toward those areas where they could be better spent (such as sustaining an EGR effort).

[30] This kind of research is being carried out under USAID's PRIMR effort in Kenya.

Annex: Notes on Cost Calculations by Country

Egypt

General

- As indicated earlier, most of RTI's technical "labor," "other," and "materials" costs were built up from a template that was filled in by project staff who knew the program well and had access to some of the detailed data.[31]

Technical costs

- For some technical people, we could obtain only their total labor and expenses for the period in which A–G unfolded, so we had to estimate the distribution of these costs across phases. This was done by people who knew enough about the project that they could estimate the amount of time (and expenses) each of these persons spent on each phase.
- Total amount of time each subcontractor spent on A through G was estimated by the Chief of Party (COP).[32]
- Subcontractors' technical costs were assumed to include labor and "other," but not materials. This was assumed since most of the materials were accounted for within RTI's budget.
- Subcontractors' technical costs were parsed out by phase based on estimates made by project staff who knew the role each subawardee played (one was 100% EGR program implementation [F], another 100% EGRA implementation [D, G], and another 10% EGRA [D, G] implementation and 90% EGR program implementation [F]).

Administration costs

- Since EGR was only a small part of what was done under GILO, we had to estimate the amount of administration costs that could be attributed to this activity. That share of the total administration cost was determined as the share of technical labor costs devoted to the EGR pilot relative to total technical labor costs across the entire project.
- RTI administration costs were distributed among each phase as per the percentage distribution of "technical costs" across each phase and include subcontractors' administration costs derived in the manner noted below.

[31] This template, an Excel spreadsheet, is another deliverable of this costing activity and can be seen by requesting a copy from USAID.

[32] The percent total time that the COP estimated each subcontractor spent on the development of this evidence-based EGR program was as follows: CID Inc. (35%), Infonex (30%), Keys to Effective Learning (KEYS) (20%), and World Education (0%).

- Subcontractors' administration costs were discerned from their total labor and expenses. Since RTI's technical costs and administration costs accounted for 65% and 35% of the total cost respectively, we assumed subcontractors' administration costs were 35% of their total costs.

Jordan

General

- It was estimated from both existing cost records and the budget, by people familiar with the program, that ~50% of the overall cost of the project was focused on reading, and the other 50% on math and the Snapshot of School Management Effectiveness (SSME) survey. Therefore, various total costs were divided by 2 to generate the relevant EGR costs.

Technical costs

- The costs for each phase were estimated based on the period of performance in which they occurred (i.e., technical labor incurred between September 2012 and August 2013 was attributed to the development of an EGR program).
- Costs that have yet to be incurred came from the work plan and the budget.
- Subcontractors' "materials" were folded into the "materials" column while their "other" costs were folded into the "Subs" column. Since the effort in Jordan is a fixed-price contract to RTI with a subcontractor, this was the extent to which these numbers could be discerned.[33]

Kenya

General

- Since this project did more than develop an evidence-based EGR program, the EGR costs had to be estimated as a percentage of the total cost of the project. The COP made the estimation based on his first-hand knowledge of the project, the project's costs, and the project's budget.

Technical costs

- Baseline EGRA and midterm EGRA costs exclude the math assessment and SSME survey that were administered at the same time. It was estimated by the COP that one-third of the total assessment cost was for math and SSME and that an additional one-sixth of the assessment costs (per assessment) was for the

[33] A fixed-price contract requires a contractor to produce a deliverable and/or carry out some piece of work for a particular price. The client does not dictate how that work is done and what the contractor spends money on to get it done, so the awardee does not expend effort on itemized cost tracking for the client. As a result, it can be very difficult to discern specific costs.

specific data requirements associated with the research aspects of the project. These were factored out.

- EGRA endline costs were estimated by the project's Kenya-based staff.
- Totals for EGR implementation were determined to be 75% of the calculated total cost given that it was estimated by the COP that 25% of those costs were for non-EGR activities.
- Labor for "EGR implementation" includes almost $200K for revisions of the materials (reading only). By December 2013 the project had completed five rounds of revisions with USAID funding, the costs of which are included.
- Subcontractor labor costs for coaches in low-cost primary schools ($357,511) are included.

Administration costs

- Administration costs were parsed out by phase according to the percentage distribution of technical costs for each phase.[34]
- As was the case for the technical cost estimates, the administration costs also removed one-third of the total costs for math/SSME and one-sixth for research and development (R&D) activities.
- COP labor was allocated as follows: 50% management (included in Administration); 30% R&D; 20% "EGR development" and "EGR implementation" (of which all labor incurred in 2011 was allocated to "EGR development," anything after was allocated to "EGR implementation").

Liberia

General

- Liberia had the most detailed cost data of the five country programs examined in this report. No estimates had to be made except for the distribution of the administration costs across each phase, which was done, as elsewhere, using the percentage distribution of technical costs across each phase.
- Contrary to the three previous cases, for Liberia, administrative loads (or overhead costs) are not included in the labor, other, and materials costs figures, but instead are folded into the overall administration cost number.
- The World Bank paid for about 45% of the cost of developing the EGR program; USAID paid for the other 55% of this phase of the effort. The World Bank's costs are included in ***Table 6***.
- The finance records had an aggregate cost for implementing all three major EGRAs: baseline, midterm, and endline. Since the training costs for conducting the baseline EGRA were higher than the training costs for the subsequent EGRAs,

[34] If 10% of the total technical cost was attributed to the development of an EGRA, 10% of the total administration costs was assigned to that phase as well.

we could not take this aggregate number and divide by 3.[35] Historical data from these five and other countries in which RTI has administered the EGRA have shown that the training costs in the baseline data account for 40% of the total cost of conducting that EGRA. It was assumed based on data from several EGRAs that the training cost for the midterm and the endline EGRAs in Liberia accounted for only 10% of the total costs of the EGRAs. It was also assumed that the rest of the costs of conducting these EGRAs remained the same for all three EGRAs. Based on these assumptions, the cost of conducting each one of these EGRAs was calculated.

Malawi

General

- It was determined that the "other" costs for EGRA (development and implementation) constituted 22% of the total "other" costs of the project and that those "other" associated with the EGR intervention (development and implementation) added up to 39% of the total "other" costs for the project. These percentages were used to help estimate many of the respective EGRA and EGR costs shown in the table.

Technical costs

- The EGRA costs reflect the cost of doing the national EGRA and the EGRAs that were conducted in the two original test districts. We were unable to separate out the costs of the two sets of EGRAs since they were conducted simultaneously.
- The EGR implementation costs are for the implementations that took place in both Year 1 (in which only two districts were involved) and Year 2 (in which seven districts were involved).
- The non-personnel and non-materials cost of a known EGRA "adaptation workshop" could not be discerned from the records, so an estimate of $5,000 was included for this item based on our knowledge of the workshop costs usually associated with such an activity.
- Local staff support for the intervention was provided through the project-hired Teacher Training Coordinators. Since these staff did work beyond the development of an evidence-based EGR program (e.g., other training activities supported by the project), it was estimated that 70% of their total labor was devoted to the implementation of the EGR program. This estimate was made by technical staff familiar with the project.

[35] The training costs for conducting the baseline EGRA are higher than the training costs for conducting the midterm and endline EGRA because once the enumerators have been trained to conduct the baseline EGRA, their training needs are substantially less to conduct subsequent EGRAs. At best, for subsequent EGRAs, enumerators need only some refresher training, and while some of the enumerators in these subsequent EGRAs may not be the same people as those who conducted the baseline EGRA, the refresher training and their being partnered with an enumerator who did conduct the baseline EGRA provides them with the training they need to conduct the subsequent EGRAs.

Administration costs

- RTI was a subcontractor to Creative Associates International, Inc. (CAII). RTI does not have access to their costs, and although they did not have many in-country costs, they had overhead costs and a Deputy Chief of Party (DCOP) who received a salary along with allowances for himself and his five dependents. These costs, which could be considerable, are not included here as we could not obtain them from CAII.[36]
- For home office labor, we used current (August 2013) salary information to calculate from the hours worked in previous pay periods. However, this calculation may represent salary increases implemented since the period of performance under the contract and may therefore be slightly higher than the actual costs incurred during the project.

[36] The largest variance in this costing exercise is in the administration costs: One project did not have a COP; for another we were unable to obtain the cost of a DCOP. This is why, in the cost effectiveness section of this study, one should focus on the calculations generating the technical cost of implementing the EGR.