



USAID
FROM THE AMERICAN PEOPLE



SYNTHESIS OF FINDINGS AND LESSONS LEARNED FROM USAID-FUNDED EVALUATIONS

Education Sector, 2013-2016

March 2, 2018

This publication was produced for review by the United States Agency for International Development. It was prepared for the E3 Analytics and Evaluation Project and for the Reading and Access Evaluation Project by Management Systems International, A Tetra Tech Company.

SYNTHESIS OF FINDINGS AND LESSONS LEARNED FROM USAID- FUNDED EVALUATIONS

EDUCATION SECTOR, 2013–2016

Contracted under

AID-OAA-M-13-00017
E3 Analytics and Evaluation Project

and

AID-OAA-M-13-00010
Reading and Access Evaluation

Prepared by:

Thomaz Alvares de Azevedo, Team Leader
Jeff Davis, Goal 1 Lead (Reading)
Clare Ignatowski, Goal 2 Lead (Workforce Development and Higher Education)
Nitika Tolani, Goal 3 Lead (Education in Crisis and Conflict)
Sean Kelly
Rosemary Taing
Kristen Potter
Andrew Trembley

DISCLAIMER

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

ACKNOWLEDGMENTS

This report has been a large collaborative effort to which many individuals have contributed their time and expertise. The study team would like to thank Bhavani Pathak of USAID's Bureau for Economic Growth, Education, and Environment (E3) for her leadership as the Contracting Officer's Representative of the E3 Analytics and Evaluation Project, and Benjamin Sylla for his leadership as the Contracting Officer's Representative of the Reading and Access Evaluation Project.

USAID/E3's Office of Education has been very supportive of this study, and we reserve a special thanks to the following people who patiently worked with the study team to identify the topics of interest and guiding questions for each Education Strategy Goal: Rebecca Rhodes and Nathaniel Haight (Goal 1: Reading), Nancy Taggart (Goal 2: Workforce Development), Gary Bittner and Susan Werner (Goal 2: Higher Education), and Nina Weisenhorn (Goal 3: Education in Crisis and Conflict). We also thank Anna Roberts, Brooke Estes, Eirini Gouleta, Grace Lang, Lauren Greubel, Nina Papadopoulos, Olga Merchan, Patricia Viala, Yolande Miller-Grandvaux for their feedback to the report. Finally, we thank Elena Walls for her support and guidance through the entire process.

Thomaz Alvares de Azevedo oversaw this study. The primary authors of the review of the findings and lessons learned associated with each Education Strategy Goal were: Dr. Jeff Davis (Goal 1), Dr. Clare Ignatowski (Goal 2), and Dr. Nitika Tolani (Goal 3). Sean Kelly contributed extensively to the analysis for all three goals, with a focus on Goal 2. Dr. Kristen Potter and Andrew Trembley supported Goal 1 and Rosemary Taing supported Goal 3. In addition, the study greatly benefited from the support of project managers Jeremy Gans and Haley Barry. The study team also thanks Marie-Louise Orsini, Lexi Boiro and Zuhayr Ahmed for their support in addressing many critical needs.

CONTENTS

Acknowledgments	ii
Acronyms	vii
Executive Summary	x
Objective.....	x
Study Selection.....	x
Methodology.....	x
Key Findings.....	xi
<i>Education Strategy Goal 1: Reading</i>	xi
<i>Education Strategy Goal 2: Workforce Development and Higher Education</i>	xii
<i>Education Strategy Goal 3: Education in Crisis and Conflict</i>	xiii
Conclusion.....	xv
Introduction	1
Background.....	1
Study Objectives and Intended Audiences.....	1
Methodology	2
Study Selection.....	2
Evaluation Quality Assessment.....	3
Selection Criteria.....	3
Synthesis Process.....	4
Objective 1: Synthesis of Findings and Lessons Learned Related to USAID	
Education Strategy Goal 1 (Reading)	7
Evaluation Sample, Context, and Limitations.....	7
Intervention Modalities.....	9
Findings.....	12
1. <i>Learning Outcomes</i>	12
2. <i>Support for Classroom Instruction</i>	16
3. <i>Teacher Training</i>	19
4. <i>Community Engagement</i>	23
5. <i>Systems and Policy Strengthening</i>	25
6. <i>Crosscutting Themes</i>	28
Objective 2: Synthesis of Findings and Lessons Learned Related to USAID	
Education Strategy Goal 2 (WfD and Higher Education)	32
Evaluation Sample, Context and Limitations.....	32
Intervention Modalities.....	34
A. <i>Workforce Development Activities</i>	35
B. <i>Workforce Development for Violence Prevention and Countering Violent Extremism</i>	37
C. <i>Higher Education Activities</i>	38
Findings.....	38
1. <i>WfD: Responsiveness to Labor Market Demands</i>	38
2. <i>WfD: Entrepreneurship Integration</i>	41

3. WfD: Access for Marginalized Groups.....	43
4. WfD: Policy and Systems Strengthening.....	44
5. WfD: Responsiveness to Needs and Dynamics in Conflict-Affected Regions.....	46
6. WfD: Measurement Issues.....	48
7. HE: University Extension Services.....	52
8. HE: Entrepreneurship Promotion.....	54
9. HE: Access for Marginalized Youth.....	54
10. HE: Policy and Systems Strengthening.....	55
11. HE: Responsiveness to Needs and Dynamics in Conflict-Affected Regions.....	56
12. HE: Measurement Issues.....	56
13. Crosscutting Themes.....	58

Objective 3: Synthesis of Findings and Lessons Learned Related to USAID

Education Strategy Goal 3 (EiCC).....	65
Evaluation Sample, Context and Limitations.....	65
Intervention Modalities.....	67
Findings.....	68
1. Strengthened Institutional Capacity to Provide Services.....	69
2. Violence Prevention and Countering Violent Extremism.....	74
3. Improved Programming.....	82
4. Safer Learning Environments for Children and Youth.....	85
5. Equity.....	88
6. Social and Emotional Learning.....	90
7. School Dropout Prevention.....	92
8. Crosscutting Themes.....	94

Recommendations	101
Goal 1	101
Goal 2	103
Workforce Development.....	103
Higher Education.....	104
Goal 3	105
Cross-Goal Recommendations	107
Lack of Impact Evaluations.....	107
Lack of Clear Description of Key Information in Evaluation Reports.....	107
Lack of Cost-Effectiveness Analyses.....	108
Gender	108
Disability.....	108
Information and Communications Technology.....	108
Innovative Finance.....	109

Conclusion.....	109
------------------------	------------

Annex 1: Statement of Work.....	111
--	------------

Annex 2: Evaluation Reports Reviewed By Activity and Education Strategy	
Goal	117

Annex 3: Goal 1 Supporting Information	120
---	------------

Annex 4: Goal 2 Supporting Information	124
---	------------

Annex 5: Goal 3 Supporting Information	130
Annex 6: 2017 Revised Version of the Evaluation Quality Assessment Tool	134
Item Description and Source	137
Annex 7: References	151

List of Tables

Table 1: Office of Education Minimum Quality Criteria.....	4
Table 2: Goal 1 Intervention Components	120
Table 3: Learning Outcomes.....	120
Table 4: Support Classroom Instruction.....	121
Table 5: Teacher Training.....	121
Table 6: Community Engagement.....	122
Table 7: Policy and Systems.....	122
Table 8: Crosscutting.....	123
Table 9: Types of WfD Services	124
Table 10: WfD Skills Training.....	124
Table 11: WfD Employment Services.....	125
Table 12: WfD Entrepreneurship Promotion.....	125
Table 13: WfD Systems and Institutional Change	126
Table 14: Types of Violence Prevention and CVE WfD* Services	127
Table 15: Violence Prevention and CVE WfD Skills Training.....	127
Table 16: Violence Prevention and CVE WfD Entrepreneurship Promotion.....	128
Table 17: Violence Prevention and CVE WfD Employment Services.....	128
Table 18: Higher Education Activities.....	129
Table 19: Goal 2 Range of Enrollment and Completions for Activities	129
Table 20: Interventions by factors related to individual radicalization and recruitment.....	130
Table 21: Intervention Applications of USAID’s Core Body of Program Principles (for Analysis, Planning, and Design).....	130
Table 22: Intervention Applications for Operational and Management Responsiveness	131
Table 23: Goal 3 Evaluations Addressing Specific EICC Priority Areas and Crosscutting Topics	132
Table 24: Needs Assessment by Outcome, Data Collection Methodology, Engagement of Local Actors.....	133

List of Figures

Figure 1: Study Process	6
Figure 2: Percentage of Evaluations Reviewed Under Goal 1 By Region, Crisis and Conflict Status, and Country Income.....	7
Figure 3: Percentage of Evaluations Reviewed Under Goal 1 By Evaluation Type and Evaluation/Implementation Phase.....	8
Figure 4: Percentage of Evaluations Reviewed Under Goal 1 By Principles of Quality	8
Figure 5: Goal 1 Intervention Modalities.....	11
Figure 6: Goal 1 Types of Learning Outcomes.....	13

Figure 7: Percentage of Evaluations Reviewed Under Goal 2 By Region, Crisis and Conflict Status, and Country Income.....	32
Figure 8: Percentage of Evaluations Reviewed Under Goal 2 By Evaluation Type and Evaluation/Implementation Phase.....	33
Figure 9: Percentage of Evaluations Reviewed Under Goal 2 By Principles of Quality	33
Figure 10: Goal 2 Intervention Modalities Related to Economic Growth and Employment	36
Figure 11: Goal 2 Intervention Modalities Related to Violence Prevention and CVE.....	37
Figure 12: Goal 2 Intervention Modalities For Higher Education	38
Figure 13: Percentage of Evaluations Reviewed Under Goal 3 By Region, Crisis and Conflict Status, and Country Income.....	65
Figure 14: Percentage of Evaluations Reviewed Under Goal 3 By Evaluation Type and Evaluation/Implementation Phase.....	66
Figure 15: Percentage of Evaluations Reviewed Under Goal 3 By Principles of Quality.....	66
Figure 16: Goal 3 Intervention Modalities	68
Figure 17: Interventions by Factors Related to Individual Radicalization and Recruitment	75

ACRONYMS

ADS	Automated Directives System
ALSE	Assessment of Learning Outcomes and Social Effects
ApaL	Aprender a Leer Project
ASU	Arizona State University
AWDP	Afghanistan Workforce Development Program
BE ²	Building Evidence in Education
BELT	Basic Education, Literacy, and Technical-Vocational Education and Training
CARSI	Central America Regional Security Initiative
CBE	Community-Based Education
CBSI	Caribbean Basin Security Initiative
CEA	Cost-Effectiveness Analysis
CECA	Community Education and Conflict Assessment
CIES	Comparative and International Education Society
CLP	Community Livelihoods Project
CSR	Corporate Social Responsibility
CV	Curriculum Vitae
CVE	Countering Violent Extremism
CWPM	Correct Words Per Minute
DEC	Development Experience Clearinghouse
DRC	Democratic Republic of Congo
E3	Bureau for Economic Growth, Education, and Environment (USAID)
EBD	Education de Base
ECCN	Education in Crisis & Conflict Network
ED	Office of Education (USAID/E3)
EESA	Education Emergency Support Activity
EGRA	Early Grade Reading Assessment
EiCC	Education in Crisis and Conflict
EIG	Education for Income Generation
ERC	Education Resource Center
ESP	Effective Schools Program
ETP	Education Transformation Project
FADCANIC	Foundation for the Autonomy and Development of the Atlantic Coast
FORSATY	Favorable Opportunities to Reinforce Self-Advancement for Today's Youth

GAO	Government Accountability Office
HE	Higher Education
HEEAP	Higher Engineering Education Alliance Program
HEI	Higher Education Institution
HICD	Human and Institutional Capacity Development
HVAC	Heating, Ventilation and Air Conditioning
ICT	Information and Communications Technology
IDP	Internally Displaced Person
IIE	Interethnic Integration in Education
IIEP	Interethnic Integration in Education Program
IRC	International Rescue Committee
IT	Information Technology
ITB	<i>Institut Teknologi Bandung</i> (Bandung Institute of Technology)
KTI	Kenya Transition Initiative
L3	Literacy, Language, and Learning
LETS	Learning Environment Technical Support
LGBTI	Lesbian, Gay, Bisexual, Transgender, and Intersex
LHC	Learning in a Healing Classroom
LMIS	Labor Market Information System
M&E	Monitoring and Evaluation
MOE	Ministry of Education
MoEST	Ministry of Education, Science, and Technology
MSI	Management Systems International
NEI	Northern Education Initiative
NGO	Nongovernmental Organization
NYU	New York University
OCA	Organizational Capacity Assessment
OPEQ	Opportunities for Equitable Access to Quality Basic Education
ORF	Oral Reading Fluency
OVC	Orphans and Vulnerable Children
PAGE	Partnership for Accountable Governance in Education
PRIORITAS	Prioritizing Reform, Innovation, and Opportunities for Reaching Indonesia's Teachers, Administrators, and Students
PTA	Parent Teacher Association
QRP	Quality Reading Project

RARA	Reading and Access Research Activity
RCT	Randomized Controlled Trial
RERA	Rapid Education and Risk Analysis
SACCO	Saving and Credit Cooperative
SCOPSO	School Community Partnerships Serving Orphaned and Vulnerable Children
SDPP	School Dropout Prevention Pilot
SEL	Social and Emotional Learning
SES	Socio-Economic Status
SMC	School Management Committee
SYLP	Somalia Youth Livelihoods Project
TVET	Technical and Vocational Education and Training
UNA/FCA	National University of Asuncion's School of Agricultural Sciences
UNRWA	United Nations Relief and Works Agency for Palestine Refugees
USAID	United States Agency for International Development
USC	University of Southern California
WfD	Workforce Development
WSRP	Whole School Reading Program
YBLP	Youth Business Leadership Project
YEGRA	Yemen Early Grade Reading Approach
YEPS	Youth Engagement to Promote Stability
YMEP	Yemen Monitoring and Evaluation Project

EXECUTIVE SUMMARY

Objective

The Office of Education in the United States Agency for International Development’s Bureau for Economic Growth, Education, and Environment (USAID/E3) commissioned a study to assess the quality of USAID-funded evaluations in the education sector from 2013 to 2016 and, based on a subset of these evaluations that met minimum quality standards, to synthesize findings and lessons learned for topics related to the Agency’s 2011-2015 Education Strategy. This study’s objective is to provide the Office of Education with a review of findings and lessons learned that may contribute to future USAID programming in three areas: (i) education quality with a focus on early grade literacy; (ii) workforce development and higher education; and (iii) access to education, especially in crisis and conflict environments.

Study Selection

The Office of Education set the following inclusion criteria for evaluations to be reviewed in this study:

- USAID-funded evaluations of education interventions;
- Performance and impact evaluations (and a small number of research studies that did not evaluate a specific intervention);
- Relevant to the Education Strategy;
- Published between 2013 and 2016;
- Single, latest published report (in case of reports for multiple phases of an evaluation);
- Evaluation reports from multiple countries (in case of a multicountry education intervention); and
- Meet minimum evaluation quality criteria for inclusion in the synthesis phase.

Methodology

This study originally included 92 evaluation reports, with the final list of evaluations vetted by the Office of Education. Thirty-six experts from 21 organizations volunteered their time and expertise to review the quality of evaluations, with MSI providing additional experts to complement the team of reviewers. Two experts reviewed each evaluation, and met after their independent reviews to discuss discrepancies and reach harmonized responses. Based on responses to the evaluation quality assessment, the Office of Education set the standard for inclusion in this synthesis. Any evaluation for which the response to one of the four questions below was “no” was excluded from this synthesis:

- Is the methodology appropriate for answering posed study questions? [all evaluations]
- Does the counterfactual meet standards of rigor? [impact evaluations]
- Were the assessments conducted in such a way that the results are generalizable to the population of students reached through the activity? [impact evaluations and quantitative performance evaluations (outcome-focused)]
- Is there a transparent connection between the study questions, findings from the data, and the conclusions and recommendations, and is the report structured to present findings clearly and objectively? [all evaluations]

Sixty-nine of the 92 evaluation reports (75 percent) progressed to the synthesis phase. In collaboration with the Office of Education, the study team identified topics of interest and synthesis questions to serve

as avenues for inquiry under each Education Strategy Goal. The synthesis section for each Education Strategy Goal provides background and context for the evaluation sample and the intervention modalities followed by key findings for topics of interest to the Office of Education and the body of evidence from which the findings were drawn.

Key Findings

Education Strategy Goal I: Reading

The review synthesized results from 23 evaluations covering 21 Goal I activities. Thirteen of the activities mainly focused on early grade reading. The other eight activities had some focus on early grade reading but mainly covered interventions with orphans and vulnerable children and initiatives with ministries of education in areas such as information systems, planning, and training. Key findings include:

- All activities that focused on early grade reading contained some combination of four main intervention areas: classroom instruction, teacher training, community engagement, and policy and systems.
- Reading interventions in almost all activities evaluated had positive effects on student scores, though the magnitude of the gains was generally small.
- Students generally made more progress in oral reading fluency than in reading comprehension. The progress in comprehension was difficult to judge since the measurements were usually lacking in reliability due to the small number of items on the subtask.
- Activities with the lowest baseline scores tended to have the most difficulty showing improvements, perhaps due to the need to develop reading skills for the non-readers. Similarly, the activities with the highest baseline scores also had difficulty showing improvements, perhaps due to ceiling effects for fluent readers.
- Boys generally outperformed girls on most activities, though girls often made progress on closing achievement gaps from baseline scores, especially in the early primary grades.
- New reading materials were featured in most of the activities. Classroom use of the materials was sometimes hampered by production delays and lack of full distribution to all schools.
- Most activities successfully implemented teacher in-service workshops. Evidence of the quantitative relationship between the workshops and student learning was unclear.¹
- Some district-level teacher coaching and in-school mentoring models showed promise, but most experienced implementation difficulties due to staffing, training, and logistical issues.
- Community engagement models focused on parent teacher associations and school management committees, with a small amount of quantitative evidence from activities showing only minimal effects on student learning.
- Promotion of reading outside of the school day was implemented in a small number of activities. Some limited results showed positive effects of extra-curricular reading on learning outcomes.
- Support for ministry of education monitoring and evaluation systems was generally unsuccessful, as were interventions using data in reforming policies around reading.
- Interventions in crosscutting areas such as learners with disabilities, innovative financing, and IT in classrooms were found only in a small number of activities, and evidence on those interventions was largely non-existent.
- Capacity-building and scaling-up interventions were included in most activities, but the resulting ability of countries to sustain and expand interventions was often either not sufficiently measured or lacking in success.

¹ Correlations between changes in student learning and changes in teacher behavior were inconclusive.

Education Strategy Goal 2: Workforce Development and Higher Education

The review synthesized results from 26 evaluations covering 30 Goal 2 activities. Key findings are divided between workforce development and higher education, as summarized below.

Workforce Development

- USAID activities are providing the holistic mix of skills that global research has called for to ensure success of youth in rapidly changing labor markets, including technical skills, soft skills, basic literacy, and work-readiness skills.
- All evaluated workforce development activities pair skills training with employment services (e.g., job placement, internships, career networking) or entrepreneurship support.
- USAID activities are responding to global research and analysis of national labor market conditions that point to the need for more focus on entrepreneurship.
- USAID workforce development activities for violence prevention and countering violent extremism (CVE) targeted marginalized youth: either young people who have dropped out of school and are unemployed, or in-school youth who are at-risk of engagement with gangs or networks espousing violent extremism.
- Nearly all activities included institutional capacity building and private sector partnerships (aspects of systems strengthening) alongside service delivery. However, the body of evidence reviewed suggests efforts at systems strengthening are not as strategic as they could be; this would require deeper understanding of incentives, disincentives, and information flows across the workforce development ecosystem.
- Although sustainable livelihoods and workforce participation are thought to be promotive factors for stability and peace, few activity designs made explicit linkages between workforce development components and violence prevention outcomes. The body of evidence reviewed suggests the need to better understand links between employment and violence prevention and countering violent extremism.
- For the vast majority of evaluations reviewed, it was impossible to render an objective judgment as to whether the activity results were positive or negative due to weak outcome measurement of employment and self-employment.
- Most evaluations reviewed did not capture input/output measurements. While not all evaluations were designed to answer the question of the overall outcome of the intervention, there is nonetheless a marked weakness in the collection and analysis of program enrollment data against program completion.
- Considering the large scale of many of the workforce development activities, often involving tens of thousands of youth and activity budgets over \$20 million (for example, in Liberia, Afghanistan, and Nepal), it is surprising that there are not more impact evaluations of workforce development activities. Without additional impact evaluations, it will be difficult to understand whether programming is truly having an effect, and which interventions work best for which populations in which contexts.
- The body of evidence reviewed suggests some impressive gains for women, though little strategic focus on gender. Workforce development activities, especially those in violence prevention and CVE, did not highlight gender-based targets or interventions.

Higher Education

- Extension services are expanding the reach of universities to contribute to country development. Although activity monitoring did not always provide a clear picture of the results

of services from the end-user perspective, this is a promising growth area for USAID higher education programming.

- Higher education students benefited from entrepreneurship promotion that was integrated into the academic curriculum as specialized training. There was one case of faculty research focused on entrepreneurship.
- The bulk of higher education evaluations reviewed focuses on professionalization of faculty. Scholarship activities were in the minority of those reviewed.
- Systems strengthening for higher education activities focused on improving the quality of instruction and promoting gender equality at all educational levels, as opposed to larger-scale institutional or policy reforms.
- Higher education interventions rarely focus on conflict prevention and stabilization. Higher education activities did take place in countries experiencing violent extremism, such as Pakistan, Lebanon and Kenya; however, the activities' theories of change and designs did not focus on stabilization or violence prevention.
- In the evaluations reviewed, there was little or no evidence of efforts to objectively measure educational or research quality beyond beneficiary self-report or satisfaction surveys. Nor was there evidence of the use of labor market surveys to increase the knowledge within higher education institutions to better respond to country skills gaps.
- Although there is little expectation of sustainability for scholarship activities without USAID funding, most evaluations reviewed of activities designed to modernize pedagogy, curriculum, and research programs revealed an interest in sustainability. However, there were few efforts to measure progress toward sustainability of these outcomes across higher education institutions and systems.
- Evaluations reviewed show mixed attention to gender. Only about half of the evaluations reviewed for higher education activities disaggregated results by gender, but about a quarter had extensive gender-focused objectives.

Education Strategy Goal 3: Education in Crisis and Conflict

The review synthesized results from 28 evaluations covering 18 Goal 3 activities. For this study, the Office of Education instructed the study team to categorize evaluations to Education Strategy Goal 3 thematically as access to education and not limit them geographically only to crisis and conflict environments. Key findings include:

- The most frequently referenced methods to improve institutional capacity building were skills training to improve pedagogical approaches and school management. Public-private partnerships were the least explored but were noted to have strong potential to improve access to education for learners in conflict and crisis settings.
- Violence prevention and CVE activities are complex and uniquely tailored to diverse political, social, cultural and economic contexts. The most effective and sustainable violence prevention and CVE activities were cross-sectoral in approach.
- Improved education programming that reflects changing contexts or beneficiary needs is due in part to immediate use of assessment or evaluation data. Situational or rapid analyses of conflict/crisis settings best represent the subset of studies that attempt to capture—in real time—an honest and accurate picture of how a conflict or crisis is interacting with learners' access to basic education.
- Safer learning environments for children and youth are best maintained through community engagement on the importance of education and mobilization (e.g., community members as volunteer teachers), curricular enhancements on topics such as conflict sensitivity, gender-based violence and interethnic integration in schools, and infrastructure improvements. Practically all

evaluations reviewed under Goal 3 pointed to the significant impact that internal, external and environmental threats posed to creating and maintaining safe learning environments for children and youth.

- Equity was most often defined by how vulnerable or at risk a child was, and was most often addressed through cross-sectoral interventions. Vulnerabilities were broadly defined and included geographic location, internally displaced person (IDP) status, gender, whether a child came from a broken home, whether a child was an orphan or came from a marginalized group, disability status, sexual orientation and socio-economic status (SES).
- Community-based education activities had a significant impact on reducing inequities in access to education, as they focus on providing education to learners who cannot attend classes at mainstream schools for a variety of reasons. In crisis and conflict settings, CBE is particularly suitable approach for providing access to female learners where travel to and from schools can pose threats to their safety.
- The review found that evidence on social and emotional learning (SEL) impacts in crisis-affected settings remains limited, without conclusive findings on improved student performance or personal development.
- School dropout prevention interventions included academic and social support combined with additional enrichment activities for at-risk students and changes in teacher practices to improve student attitudes and behaviors. These shifts in student attitudes and behaviors translated to increased student engagement and ultimately reduced school dropout. The School Dropout Prevention Pilot Program (SDPP) demonstrated varying success in increasing teaching capacity (defined as effectiveness in dropout prevention, teachers' sense of responsibility, and self-efficacy), school administration capacity, and student outcomes (i.e., attendance, dropout and grade promotion).
- Gender is not yet a key influencer in the design, implementation, and evaluation of education interventions. Most of the evaluated Goal 3 activities analyzed results with gender disaggregation but did not explicitly design their interventions with a gender-responsive theory of change.
- Disability, defined broadly, was rarely examined in the reviewed evaluations of Goal 3 interventions. Only 5 of the 28 evaluations reviewed under Goal 3 made any reference to disabilities. Disability was typically mentioned (if at all) as one risk factor under a broader umbrella of vulnerabilities related to addressing equity through the intervention. Increased access to education for students with disabilities was often due to a multipronged approach at local and national levels to raising awareness about the rights and needs of children with disabilities among parents, teachers and communities, and ensuring an enabling education policy environment.
- Fourteen of the 28 evaluations reviewed under Goal 3 made some mention of information and communications technology (ICT) in the activity design, but the evaluation reports contained few details on the effectiveness or impact of ICT on access to education and other outcomes. USAID should systematically evaluate the value added of ICT as a means to improve access to education in crisis and conflict settings. In particular, the usefulness of ICT-based approaches versus face-to-face trainings and interventions should be examined.
- Innovative financing mechanisms are key to ensuring education program sustainability. Eleven of the 28 evaluations reviewed mentioned innovative financing to some extent, usually as a recommendation for future consideration, to enhance activity sustainability. The cost-effectiveness of public-private partnerships to support access to education in crisis and conflict settings is unclear and should be investigated.
- Most Goal 3 evaluations reviewed made some mention of possible barriers to activity sustainability given the contexts in which the activities were implemented (unstable, low-resource areas targeting high-risk children and youth). The predominant barrier facing Goal 3

education activities is financial sustainability, followed by lack of government and community ownership of the intervention.

Conclusion

This study provides a synthesis of findings and lessons learned for topics of interest provided by the Office of Education and associated with USAID's three Education Strategy Goals. Sixty-nine evaluations that met the Office of Education's quality criteria for inclusion in the syntheses were reviewed (therefore, activity-level information may not be exhaustive) and the study only considered information provided in these evaluation reports (so activity information might be incomplete). Given that the 2011 Education Strategy required USAID missions to align programming with the Strategy by 2013, this study only reviewed evaluations published after the alignment. The realignment likely affected the implementation of the activities that were realigned, as well as the evaluations of these activities. Also, as the contracting mechanisms used to support the Education Strategy Goals were awarded in the 2014-2015 period, relevant findings and lessons learned from activities aligned to the 2011 Education Strategy will continue to unfold until at least 2019.

INTRODUCTION

Background

More than five years after instituting the Agency’s Education Strategy, the Office of Education in the United States Agency for International Development’s Bureau for Economic Growth, Education, and Environment (USAID/E3) commissioned a study to synthesize findings and lessons learned from evaluation reports on topics of interest related to each Education Strategy Goal. Recent research² has shown that reviews of education outcomes reported in evaluations arrive at different conclusions about “what works.” This is partly because there is so much heterogeneity across the studies that modest variations in the inclusion criteria lead to different conclusions, and partly because the classes of interventions are often not consistently defined. Thus, for this report the study team closely collaborated with the Office of Education to set the inclusion criteria for the evaluation reports reviewed and, in addition to synthesizing findings and lessons learned, examined the intervention modalities described in the evaluation reports.

Recent research by the Research on Improving Systems of Education (RISE)³ have suggests that although adequate designs have been used for estimating causal impacts of various education activities, the resulting body of evidence has been too inadequately theorized to be of direct use in formulating effective action. That is, just producing more systematic reviews of impact evaluations⁴ is unlikely to add up to a research agenda. This focus on effective use of evidence aligns with the Office of Education’s interest in this study being a more qualitative exploration, with synthesis questions treated as avenues for inquiry into topics of interest—in contrast to treating the questions as fixed research questions to be answered with a more quantitative meta-analysis.

Study Objectives and Intended Audiences

This report synthesizes findings and lessons learned from USAID-funded evaluations reviewed on topics of interest to the Office of Education under the three USAID Education Strategy Goals for the 2011-2015 period (extended until December 2017):⁵ (1) improved reading skills for 100 million children in primary grades; (2) improved ability of tertiary and workforce development (WfD) activities to generate workforce skills relevant to a country’s development goals; and (3) increased equitable access to education in crisis and conflict environments for 15 million learners.

All evaluations synthesized in this study met minimum quality standards set by the Office of Education for inclusion in the synthesis phase of this study. The evaluation quality assessment was the product of the first phase of this study that involved a collaborative process with reviews by experts from the international education community (who volunteered their time and expertise). The resulting syntheses

² Evans, David K., and Anna Popova. "What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews." *The World Bank Research Observer* 31, no. 2 (2016): 242-270.

³ Pritchett, Lant. *The Evidence About What Works in Education: Graphs to Illustrate External Validity and Construct Validity*. RISE Insights, June 2017.

⁴ For instance, for a list of recent systematic reviews on improving learning in developing countries, see: Kremer, Michael, Conner Brannen, and Rachel Glennerster. "The challenge of education and learning in the developing world." *Science* 340, no. 6130 (2013): 297-300; Krishnaratne, Shari, Howard White, and Ella Carpenter. "Quality education for all children? What works in education in developing countries." New Delhi: International Initiative for Impact Evaluation (3ie), Working Paper 20 (2013); McEwan, Patrick J. "Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments." *Review of Educational Research* 85, no. 3 (2015): 353-394.

⁵ USAID, *Extension of the USAID Education Strategy to December 2017*.

of findings and lessons learned identify the intervention modalities being funded by USAID and contextualize the body of evidence behind them.

The objectives of this study are to provide the Office of Education with a review of findings and lessons learned that may contribute to future USAID programming related to: (i) education quality with a focus on early grade literacy; (ii) WfD and higher education; and (iii) access to education, especially in crisis and conflict environments.⁶ The primary audiences for this study are the Office of Education, USAID mission staff, and implementing and country partner organizations that plan and deliver education and WfD activities and related support services.

The study team hopes that the information in this report will also enrich USAID's programming discussions about how to tackle Goal 4 of the Sustainable Development Goals on ensuring inclusive and equitable quality education and promoting lifelong learning opportunities for all,⁷ and the U.S. Reinforcing Education Accountability in Development Act (HR 601) that establishes USAID's mandate to: (1) respond to the needs and capacities of developing countries to improve literacy and other basic skills; (2) strengthen educational systems, expand access to safe learning environments (including breaking down barriers to basic education for women and girls), and support the engagement of parents in their children's education; (3) promote education as a foundation for economic growth; and (4) monitor and evaluate the effectiveness and quality of basic education activities in partner countries.⁸ The study team also hopes the Office of Education will find the information in this report useful when setting the Agency's next education strategy.

A team led by Management Systems International (MSI) conducted this study across two mechanisms: The E3 Analytics and Evaluation Project (implemented by MSI in partnership with Development and Training Services, a Palladium company, and NORC at the University of Chicago), and Reading and Access Evaluation (implemented by NORC with MSI as a subcontractor). Annex I provides USAID's statement of work for this study.⁹

METHODOLOGY

Study Selection

The Office of Education requested that evaluation reports published between 2013 and 2016 and related to at least one of the three Strategy Goals be included in this study. This is consistent with the [Implementation Guidance to the 2011 USAID Education Strategy](#) which set the target date for alignment of field Mission activities with the new strategy as the beginning of FY13.¹⁰

Based on the guidance provided in the [USAID Evaluation Policy](#), no single evaluation design should be privileged over others; selection depended on considerations about the appropriateness of the evaluation design for answering the evaluation questions, cost, feasibility, and the level of rigor needed to inform specific decisions.¹¹ Thus, this study includes findings and lessons learned from both impact evaluations

⁶ For this study, the Office of Education instructed the study team to categorize evaluations related to Education Strategy Goal 3 thematically (access to education) instead of geographically (crisis and conflict).

⁷ Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators, Annex III.

⁸ READ Act (HR 601).

⁹ Annex I provides the statement of work for the E3 Analytics and Evaluation Project's component of this study, focusing on Goal 2. The Reading and Access Evaluation project's statement of work for this study is nearly identical, replacing references to Goal 2 with Goals 1 and 3.

¹⁰ USAID, *Implementation Guidance*.

¹¹ USAID, *Evaluation Policy*.

and performance evaluations. At the Office of Education’s request, the study team also reviewed a limited number of research studies that did not evaluate a specific USAID-funded intervention.

If an activity was evaluated at multiple time points (e.g., baseline, midterm, final), the study team included only the latest published report. If the activity being evaluated was implemented in several countries, the study team included reports for each of the countries. Only evaluations the Office of Education considered relevant for the Education Strategy were included, and all evaluations reviewed were funded by USAID.

During the first phase of this study, 92 evaluation reports were reviewed.

Evaluation Quality Assessment

During the evaluation quality assessment phase, each evaluation report was reviewed by two experts and scored using an evaluation quality assessment tool developed by the study team in collaboration with the Office of Education. This tool was based on the framework prepared by the United Kingdom’s Department for International Development and produced by the Building Evidence in Education (BE²) working group¹² that identified seven principles of quality: the conceptual framing of the study; its openness and transparency; the robustness of the methodology; the cultural appropriateness of the tools and analysis; and the validity; reliability; and cogency of the report. Co-reviewers then compared their scorings before recording a final consensus response for each item in the tool. This “Assessment of the Quality of USAID-Funded Evaluations: Education Sector, 2013-2016”¹³ report presents the results for the quality assessment of the 92 evaluation reports reviewed. The expert volunteers provided feedback on the evaluation quality assessment tool, and upon the completion of the review process the tool was revised. Annex 6 contains the final version of the evaluation quality assessment tool.

During the evaluation quality assessment, 36 volunteer experts from 21 organizations conducted reviews for this study. MSI complemented the volunteer reviewers with eight staff and six consultants.

Selection Criteria

The Office of Education set minimum quality standards for evaluations to progress to the synthesis phase of this study: the evaluations had to meet all four criteria listed in Table I below.¹⁴ These criteria used questions related to the three quality principles proposed by BE²: robustness of the methodology, validity, and cogency. These principles mostly overlapped with the principles suggested in the [BE²](#)

¹² BE² is a donor working group started in 2012 that includes 30 member organizations, led by USAID, the United Kingdom’s Department for International Development (DFID), the World Bank and multiple United Nations agencies. It promotes evidence to inform policy and make programming decisions, and build common standards on how to assess evidence from education evaluations.

¹³ Thomaz Alvares de Azevedo, *Assessment of the Quality of USAID-Funded Evaluations: Education Sector, 2013-2016* (USAID, January 2018), http://pdf.usaid.gov/pdf_docs/pa00srwl.pdf.

¹⁴ The expert reviewers were given the option to respond “yes,” “partial,” or “no” to all these items except for the item about the counterfactual, which was applied only to impact evaluations—it allowed only a “yes” or “no” as response. Evaluations for which the response was “no” to any of these questions did not progress to Phase 2 (synthesis).

[Assessing the Strength of Evidence in the Education Sector](#)¹⁵ guidance note to be lifted from individual studies for categorizing the quality of the body of evidence.¹⁶

TABLE 1: OFFICE OF EDUCATION MINIMUM QUALITY CRITERIA

Principle	Item	Options	Applicability
Robustness of the methodology	Is the methodology appropriate for answering posed study questions?	yes/partial/no	All
Robustness of the methodology	Does the counterfactual meet standards of rigor?	yes/no/not applicable	Impact evaluations
Validity	Were the assessments conducted in such a way such that the results are generalizable to the population of students reached through the activity?	yes/partial/no/not applicable	Impact and performance evaluations (outcome-focused)
Cogeneity	Is there a transparent connection between the study questions, findings from the data and the conclusions and recommendations, and is the report structured to present findings clearly and objectively?	yes/partial/no	All

Out of the 92 evaluations reviewed, 69 evaluations (75 percent) met the Office of Education quality standards for inclusion in the synthesis phase.

Synthesis Process

In collaboration with the Office of Education, the study team identified topics of interest under each Education Strategy Goal and synthesis questions to serve as avenues for inquiry. The synthesis was structured as follows:

- An overview of the evaluation sample and context, intervention modalities, assessment of the strength of the body of evidence, and limitations.
- Key findings and lessons learned for the topics of interest for the Office of Education and summary of the body of evidence from which the key findings were derived.

¹⁵ Building Evidence in Education (BE²) Steering Committee. *Assessing the Strength of Evidence in the Education Sector*. 2015.

¹⁶ BE² suggests that robustness of methodology (rigor), validity, and reliability be used to determine the quality of the body of evidence. Instead, to determine which evaluations to select for the synthesis phase, the Office of Education requested that the question about the conclusion validity under cogeneity be used, replacing the question about reliability.

THE OFFICE OF EDUCATION TOPICS OF INTEREST

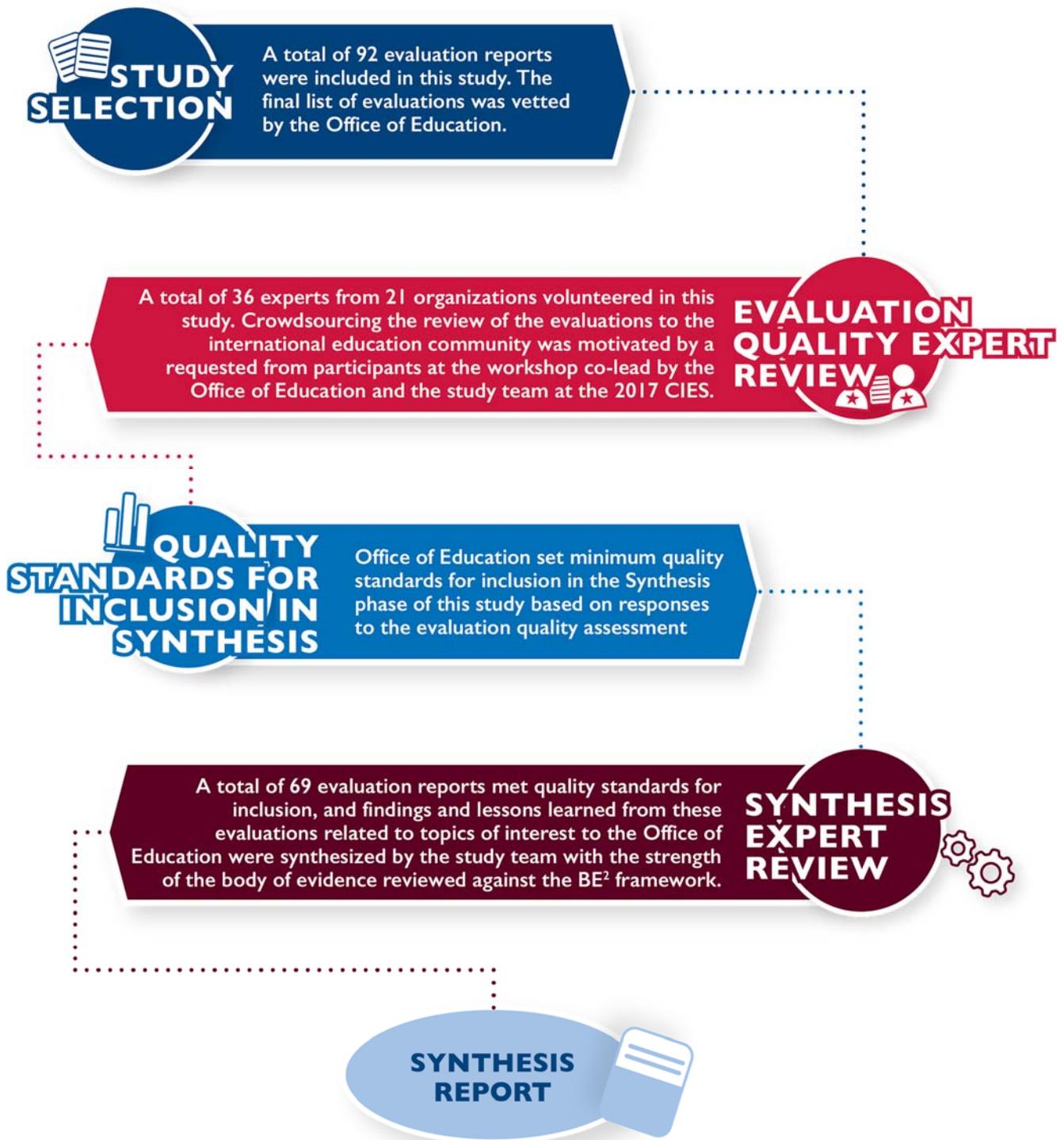
Goal 1: Reading – Learning outcomes, support for classroom instruction, teacher training, community engagement, and policy and systems strengthening.

Goal 2: Workforce Development and Higher Education – Responsiveness to labor market demands, university extension services, entrepreneurship, access for marginalized groups, policy and systems strengthening, responsive to needs and dynamics in conflict-affected regions, and measurement issues.

Goal 3: Education in Crisis and Conflict – Strengthened institutional capacity, violence prevention and countering violent extremism, improved programming, safer learning environments for children and youth, equity, social and emotional learning, and school dropout prevention.

Crosscutting Themes – Gender, disability, information and communications technology, innovative financing, and scaling up/sustainability.

FIGURE 1: STUDY PROCESS



OBJECTIVE I: SYNTHESIS OF FINDINGS AND LESSONS LEARNED RELATED TO USAID EDUCATION STRATEGY GOAL I (READING)

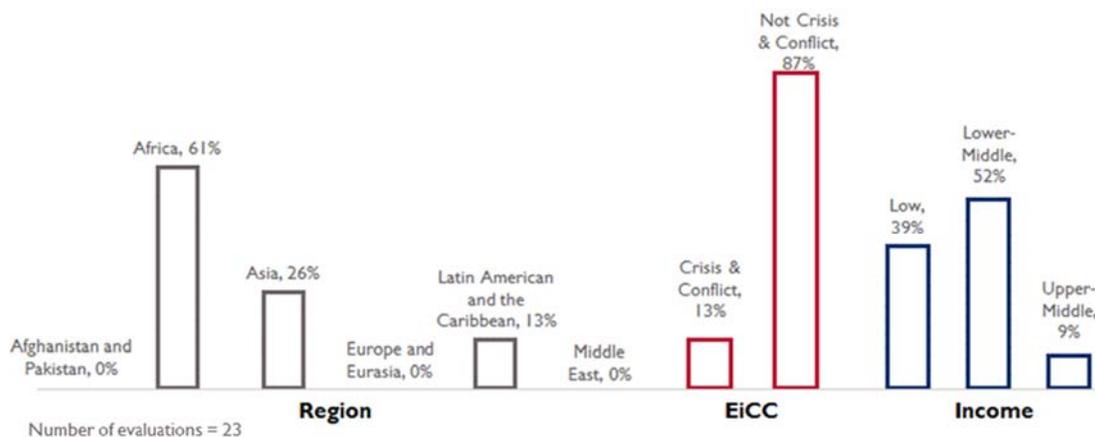
Goal I programming addresses the high-level Education Strategy goal to “improve reading skills for 100 million children in primary grades.” This review synthesizes the results for 23 evaluations reviewed covering 21 activities related to Goal I.¹⁷ Overall, this constitutes a medium-sized body of evidence.

Evaluation Sample, Context, and Limitations

All Goal I related evaluations reviewed in this study were funded by USAID and met the Office of Education criteria for evaluation quality. These included performance and impact evaluations published between 2013 and 2016.

Although USAID reading activities are implemented globally, there are gaps in the evidence, as no evaluations from Afghanistan and Pakistan, Europe and Eurasia, or the Middle East were conducted during the reviewed period and passed the Office of Education evaluation quality check. Furthermore, more than half the evaluations analyzed (61 percent) took place in sub-Saharan Africa. The reasons for this imbalance in geographic distribution of the evaluations that met the criteria for the review would be worthy of further study.

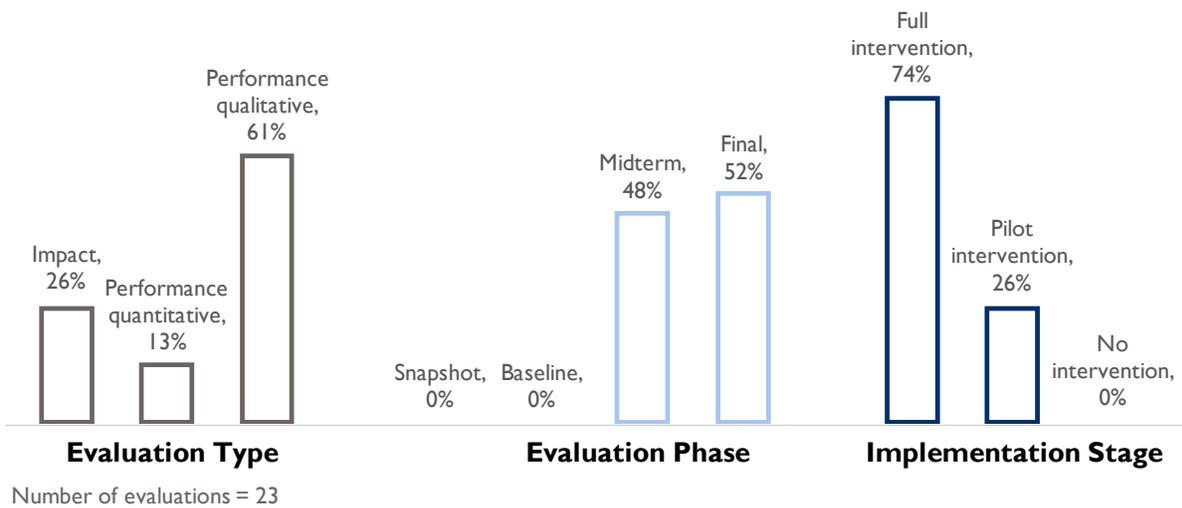
FIGURE 2: PERCENTAGE OF EVALUATIONS REVIEWED UNDER GOAL I BY REGION, CRISIS AND CONFLICT STATUS, AND COUNTRY INCOME



Most of the evidence came from qualitative performance evaluations (61 percent) and about half (52 percent) were final evaluations.

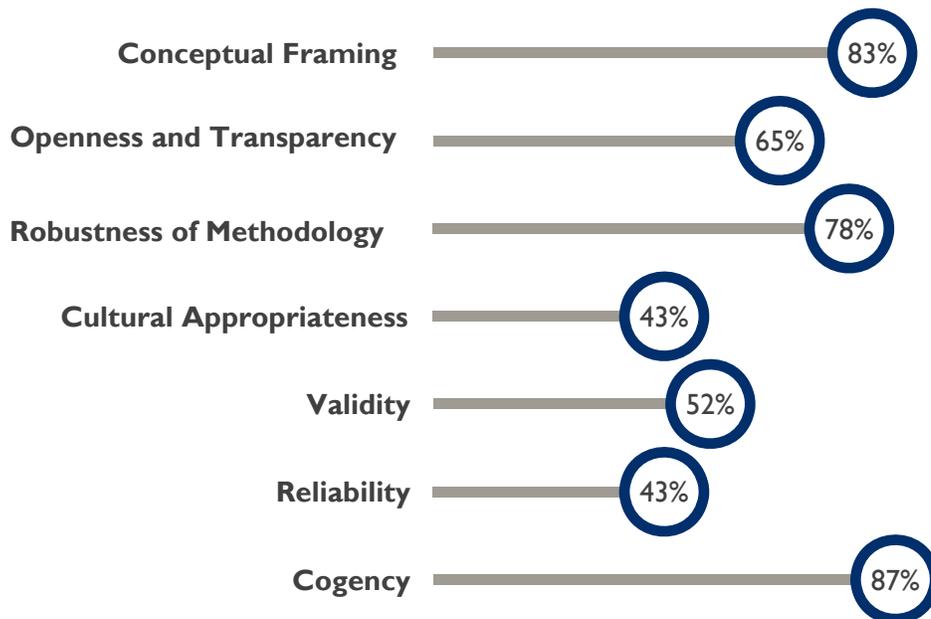
¹⁷ Activities could include more than one report in case they were implemented in multiple countries or in case both impact and performance evaluation reports have been produced and met the Office of Education standards for inclusion in this study.

FIGURE 3: PERCENTAGE OF EVALUATIONS REVIEWED UNDER GOAL I BY EVALUATION TYPE AND EVALUATION/IMPLEMENTATION PHASE



Based on the expert judgments of the seven principles of quality, cogency (87 percent) was rated as “adequate” most frequently, while cultural appropriateness and reliability (43 percent each) were rated as “adequate” the least. Results on Figure 4 indicates that despite the Office of Education’s minimum quality standards for inclusion in the synthesis, the general quality of the body of evidence is only moderate.

FIGURE 4: PERCENTAGE OF EVALUATIONS REVIEWED UNDER GOAL I BY PRINCIPLES OF QUALITY



USAID's success in promoting best practices for improving children's reading skills under the education strategy allowed for some generalizations in the synthesis of the findings. However, there were the following difficulties in obtaining information on the effects of intervention components from the evaluation reports reviewed:

- Since each of the activities in the evaluation reports reviewed had combinations of at least two different components of intervention models – i.e., from among classroom instruction, teacher training, community engagement, and policy and systems – the activity designs did not allow for separating out component effects.
- Each component had sub-components – e.g., within classroom instruction, activities had some combination of pedagogy, time on task, and materials – so gathering quantitative evidence on the effects from activities related to sub-components was equally difficult to obtain.
- Evaluations reviewed were often able to identify the intended dosage of interventions by activities and collected a variety of information on fidelity of implementation, but the information was not clear or consistent enough to permit analyses relating actual dosage to intervention effects across activities.
- Information was available on student learning outcomes from most activities involving teaching and learning, with consistent measurement methods used for most activities, but it was not possible to capitalize on this information to evaluate the effects of intervention components due to the issues described above.
- Given these activity and evaluation designs, qualitative studies were the only avenues for gathering and analyzing information on the relative benefits of different components, e.g., anecdotal information showed that interventions related to classroom instruction and teacher training were more effective in improving children's learning than interventions in community engagement and policy and systems.

Large-scale quantitative information on the relative (and more cost-effective) effects of the components of the intervention packages would be highly valuable for programming purposes. The synthesis findings suggest the following recommendations for obtaining additional quantitative information on interventions from evaluation studies:

- Develop better theories of change in activity designs with comprehensive descriptions of how and why a desired change in a particular context is expected to happen, so that evaluations can be designed to determine whether that change has occurred.
- Continue to improve on the application of methods already in place to accurately and consistently measure student learning outcomes, especially grade two students' reading skills at baseline, midline, and endline at the same time point in the school year, so that outcome measures will be similar across all activities.

Promote comparability of results across activities by using standardized methods for calculating effect sizes as an important step in generalizing results across activities, given the contextualized packages of intervention components and dosages in the field, to make more informed programming decisions.

Intervention Modalities

The foundation of the Goal 1 activities is learning outcomes in reading. The activities that focused on early grade reading outcomes contained some combination of the four main intervention areas aimed at

the following: support for classroom instruction,¹⁸ teacher training,¹⁹ community engagement,²⁰ and policy and systems.²¹ See Tables 3-7 in Annex 3 for the range and frequency of types of interventions, classroom instruction, teacher training, community engagement modality, and policy and systems change.

While there were 13 activities that focused on early grade reading (sometimes coupled with early grade math), other activities were also classified as falling under Goal I in this study. These include three activities that focused on orphans and vulnerable children (OVCs) but included reading components along with five activities that targeted specific interventions involving Ministries of Education. Note it is possible that some of the remaining activities measured learning outcomes but did not include the measurements in the reports since the evaluations focused on other activities.

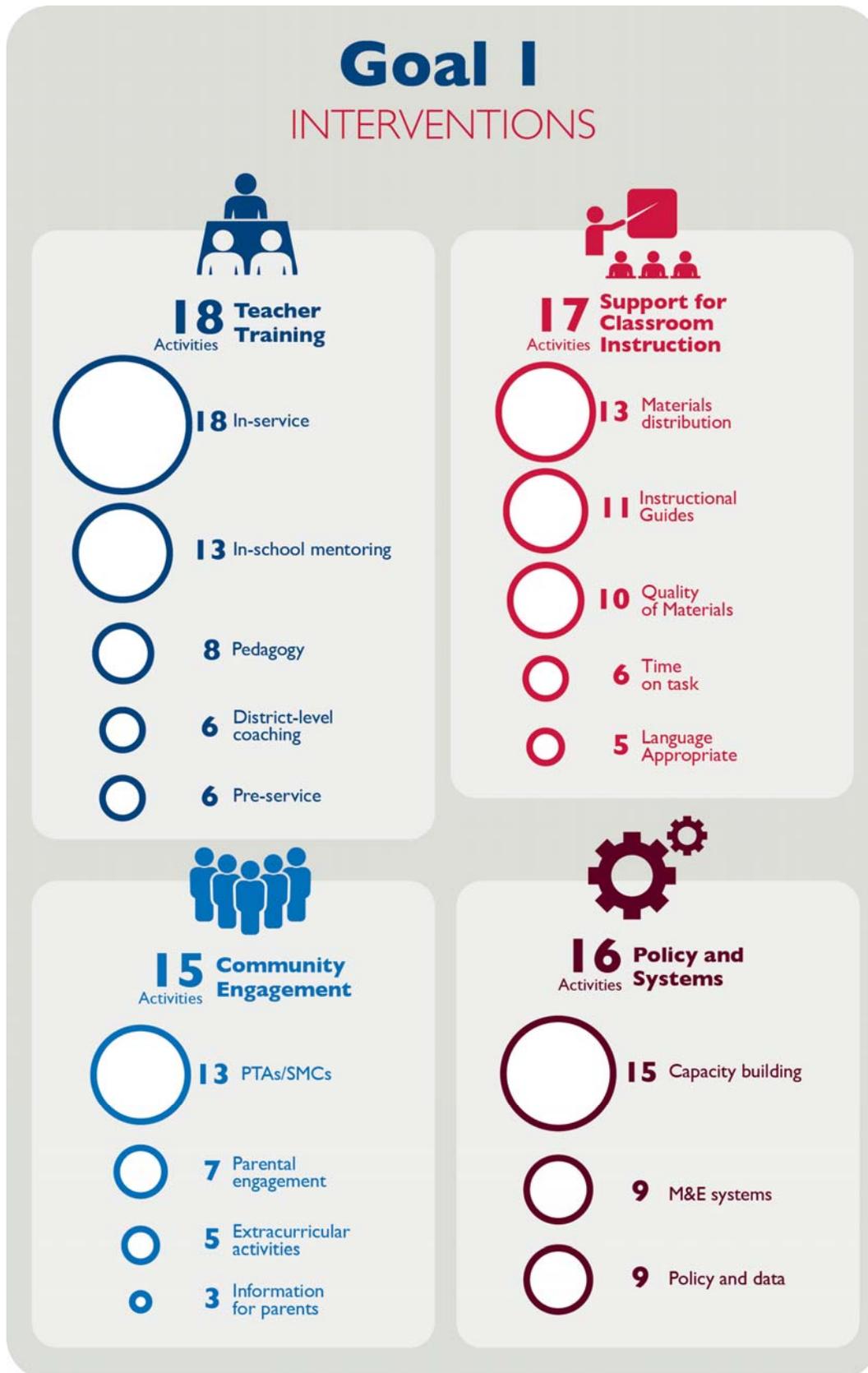
¹⁸ Support for classroom instruction includes interventions in quality learning materials, multiple languages, and time-on-task.

¹⁹ Teacher training includes pre-service, in-service, in-school coaching/district-level mentoring, and pedagogy.

²⁰ Community engagement includes parent-teacher associations (PTAs), school management committees (SMCs), parental support at home, extracurricular activities, and sharing student performance information with parents.

²¹ Policy and systems include monitoring and evaluation (M&E) systems, use of policy and data, and general capacity building.

FIGURE 5: GOAL I INTERVENTION MODALITIES



Findings

The lessons learned from this in-depth review are categorized by their intervention components and by crosscutting issues. Guiding questions, developed in conjunction with USAID, served as a starting point of inquiry. Other important themes arose organically during the evaluation review and are interspersed throughout the report.

I. Learning Outcomes

Guiding Questions

Are interventions leading to improved reading outcomes in early grades? Are there differences in disaggregated results by factors such as student gender, location (urban/rural), socio-economic status, language, and region? Which reading skills are most related to improved oral reading fluency and reading comprehension?

Results Summary

The review found that reading interventions on many of the activities had positive effects on student scores. Often, however, the practical effects of the interventions were small. The countries with the lowest baseline scores generally had the most difficulty showing improvements. Some of the activities in the sample did not measure student learning. In addition, questions that required quantitative analysis of intervention components, i.e., classroom instruction or teacher training, could not be answered due to combining the components into packages for the activities.

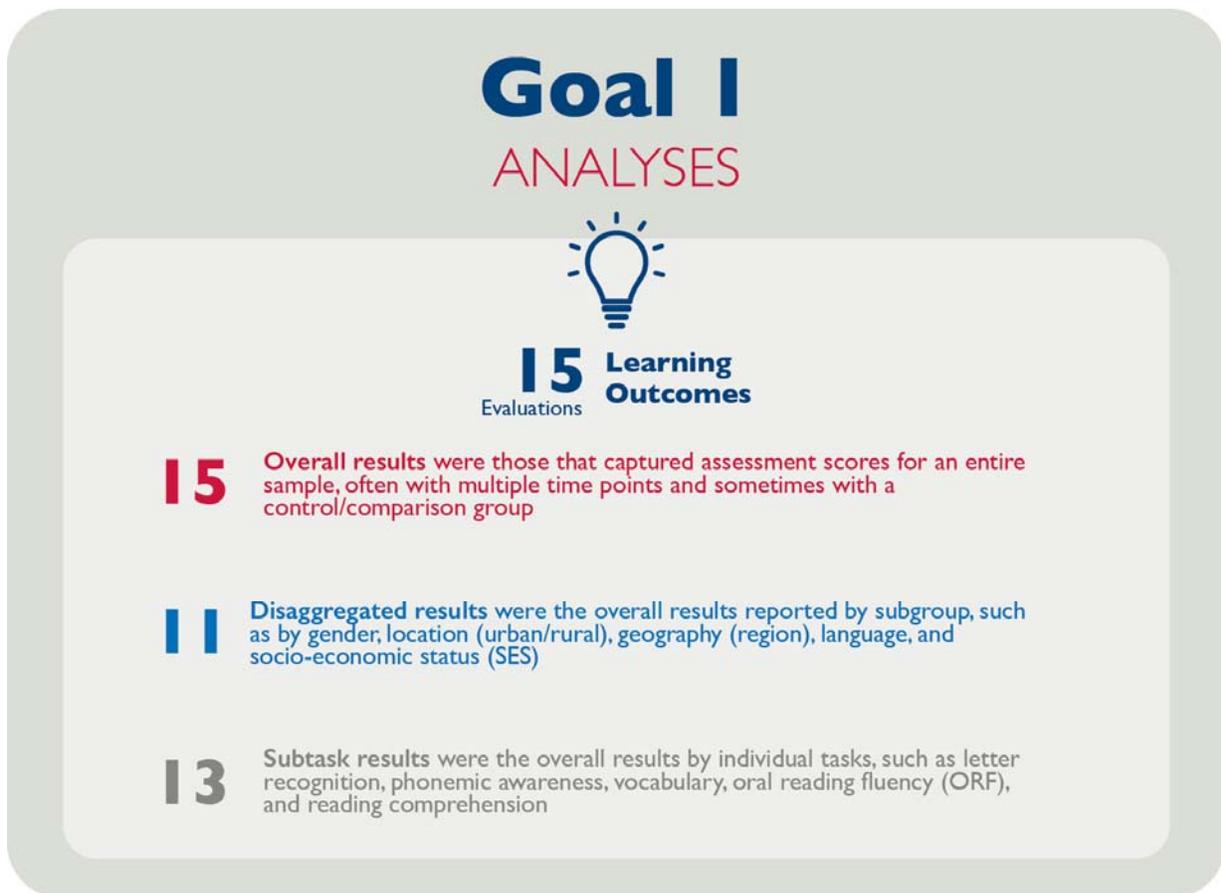
- *For learning outcomes, there were three types of analyses: overall,²² disaggregated,²³ and by subtask.²⁴ See Table 2 in Annex 3 for the range and frequency. See Figure 6 below for a summary.*

²² Overall results were those that captured assessment scores for an entire sample, often with multiple time points and sometimes with a control/comparison group.

²³ Disaggregated results were the overall results reported by subgroup, such as by gender, location (urban/rural), geography (region), language, and socio-economic status (SES).

²⁴ Subtask results were the overall results by individual tasks, such as letter recognition, phonemic awareness, vocabulary, oral reading fluency (ORF), and reading comprehension.

FIGURE 6: GOAL I TYPES OF LEARNING OUTCOMES



A. Overall Results

- Most reading interventions had significant, positive, albeit small effects on reading scores.
- Activities with low and high baseline scores had difficulty showing improvements.

The Early Grade Reading Assessment (EGRA) was used to measure reading outcomes in slightly over half the activities (11 of 21), with the Jamaica Education Transformation Project (ETP) implementing an EGRA-like instrument and three others employing national assessments without measuring ORF or other EGRA subtasks. Of the 15 activities with learning outcomes, 10 had pre- and post-test results. Eleven activities reported their results in raw (or percent correct) scores, and 10 used performance categories for reporting.²⁵ ORF data showed that many activities had statistically significant gains, but the practical effects were mostly small (see below). Most reviewed evaluations did not report percentages of students meeting reading benchmarks and targets.

Although reading outcomes in target schools across the evaluated activities showed a general trend toward positive effects on student reading skills, the practical effects, in terms of either correct words per minute (CWPM) gains or effect sizes, were often small. There was also a general trend toward better results when the baseline showed medium levels of ORF; the countries with low or high starting points often found it more difficult to show gains. Countries with the

²⁵ For impact evaluations, the analytical methods were “difference in difference” in approximately three-quarters of activities and “difference of means” in the remaining quarter. Eight activities evaluated results with control/comparison groups.

high percentages of nonreaders were often unable to move large percentages of students from their zero scores. Substantial gains in reading comprehension were difficult for most activities to attain. Language of instruction made positive differences in some cases where activities implemented interventions in local languages.

In Philippines Basa (Read Philippines), the baseline ORF results were relatively high on average, and the midline showed positive gains with a small effect size (.09). For the Kyrgyz Republic Quality Reading Project (QRP), the intervention students had medium levels of ORF at baseline, and the endline showed positive gains with a medium effect size (.24). In the Rwanda Literacy, Language, and Learning Initiative (L3), there were low to medium ORF levels at baseline and medium effect sizes at midline (.19 to .29, depending on grade level). A recent meta-analysis of ORF results suggest that effect sizes over 0.25 standard deviations as substantial.²⁶ In Malawi EGRA, the levels were very low at baseline, and gains were small at endline (only 1 CWPM). In the Mozambique Aprender a Ler Project (ApaL), results were positive, but the percentage of students with zero ORF scores remained high (53 percent). In Kyrgyz Republic QRP, differences in reading comprehension were small and insignificant. In Ghana Partnership for Accountable Governance in Education (PAGE), the activity reached its target for gains in ORF, but only in Ghanaian languages, not in English.

Activities with large percentages of low-achieving students at baseline, such as Malawi EGRA and Mozambique ApaL, showed significant ORF gains, i.e., as a percentage of the starting point. However, even with the significant gains, fluency levels remained low, so the practical gains were small. Activities with more high-achieving students—such as Indonesia Prioritizing Reform, Innovation, and Opportunities for Reaching Indonesia’s Teachers, Administrators, and Students (PRIORITAS) and Philippines Basa—were able to achieve gains, but the effect sizes were small. Some activities with higher percentages of middle-achieving students, such as Dominican Republic Effective Schools Program (ESP), had greater practical gains. Reasons for the lack of large gains by the activities with high-achieving students were not provided in the reports. In activities with relatively large percentages of both low- and high-achieving students, the small practical gains could have been due to floor and ceiling effects. This issue merits further study.

B. Disaggregated Results

- *Evaluations reviewed differed in how results were disaggregated.*
- *Boys generally outperformed girls, but girls made progress in closing achievement gaps.*
- *Urban students had higher baseline scores, but rural students showed greater performance improvement.*
- *Higher SES and instruction in local languages were linked with higher reading scores.*

Reading outcome data were disaggregated by gender in almost all relevant evaluation reports, and the scores for the boys were usually higher than those for the girls, particularly in the upper primary grades (e.g. grades 3 and 4). Smaller differences were often found for the scores in the lower primary grades (e.g. grades 1 and 2). In many activities, the girls in the lower grades were able make progress closing gaps in performance, such as in Kyrgyz Republic QRP and Mozambique ApaL. It appears that activities with interventions supporting girls’ education were generally successful in improving scores, though the issue needs further study.

Activities differed on whether and how results were disaggregated. Almost half (10 of 21) of the evaluations reviewed did not report disaggregated findings at all. For those reports that showed

²⁶ Graham, Jimmy, and Sean Kelly. *How Effective Are Early Grade Reading Interventions? A Review of the Evidence*. Washington, DC: World Bank, 2018.

disaggregated data, the most common categories were gender, urban/rural setting, geographic location (region/province), language, and SES. For instance, the Indonesia PRIORITAS evaluation disaggregated reading outcomes by gender, urban/rural setting, and geographic location (province or region), while Djibouti Projet AIDE and Kyrgyz Republic QRP only disaggregated by gender and location. Activities such as Philippines Basa and Rwanda L3 disaggregated by gender, language, and SES.

The activities that disaggregated their data tended to show differences in results for some of the groups.

Gender differences in learning outcomes were apparent in most of the evaluations reviewed, with boys' outcomes generally higher than those of girls. In fact, in only three activities were the girls' ORF levels higher than those of boys, including both Philippines activities—Basa and Whole School Reading Program (WSRP)—and Rwanda L3. None of the evaluations reviewed provided reasons for these findings.

Some of the gender gaps narrowed from pre-test to post-test on many of activities, with higher reading gains by girls than boys in the lower primary grades, though the gains in boys' outcomes tended to rise more in the upper primary grades. For example, in Kyrgyz Republic QRP, there was some narrowing of gaps in lower primary, with girls making higher gains than boys, but not in upper primary, where there were large and statistically significant effects in favor of boys. In Mozambique ApaL, the gains in scores for grade 2 female students were higher than those for grade 2 male students, but the situation was the reverse in grade 3. These results indicate that the USAID classroom interventions tended to have a better impact on girls in the lower primary grades; however, none of the reviewed evaluation reports provided reasons for the differences in gains between boys and girls at the different grade levels.

SES differences were studied in only one evaluation, Philippines Basa. The report showed that SES was correlated with reading scores. Reasons provided included a link between higher reading scores and SES-related factors such as a reading-rich environment, better educated mothers, and higher economic/job status.

The language of instruction generally showed differences on activities that measured outcomes in multiple languages. For the Kyrgyz Republic QRP, there were stronger gains in the Kyrgyz language than in Russian, though the small sample size of the Russian speakers may not have permitted the detection of effects. As mentioned earlier, for Ghana PAGE, the gains for mother tongue instruction were higher than those for English. For Nigeria Reading and Access Research Activity (RARA), reading score gains by students receiving instruction in a local language were also higher than those for English.

Finally, the urban or rural setting tended to show differences. Students in urban areas often had higher baseline scores than those in rural areas; however, students in the rural areas often had higher gains at midline and endline. For instance, with Mozambique ApaL, the students in urban schools had higher scores, but they showed little ORF gain between baseline and midline, whereas the rural students had substantial gains. In Kyrgyz Republic QRP, the gains for rural students in grade 2 were much higher than for students from urban areas. These results are consistent with the overall trend that rural students with lower scores at baseline tend to gain more by endline; however, the students with lower scores are generally not able to completely close the gap by endline.

C. Subtask Results

- *Students made more progress in ORF than in reading comprehension.*
- *Comprehension measurements generally lacked reliability.*

Most of the evaluations reviewed (13 of 21) provided information on subtasks of the EGRA tools that they administered. Only two of the evaluations that reported learning outcome results did not include information on their subtasks. All evaluations that reported on subtasks had some discussion on ORF and comprehension.

ORF gains were usually significant, but the comprehension gains were more mixed, sometimes with almost no practical effects. Three issues with comprehension measurements were: 1) a maximum of five or six comprehension questions (with one reading passage) on most of the EGRAs, which resulted in a lack of reliability and high measurement error; 2) near universal measurement of comprehension with oral reading but not with silent reading; and 3) focus on ORF in reporting rather than an equal balance between ORF and comprehension.

Besides ORF and comprehension, other subtasks showed mixed results in terms of gains. For instance, with Kyrgyz Republic QRP, gains were the highest in listening comprehension and familiar words. With Philippines Basa, letter sounds showed the greatest gains, though familiar and unfamiliar words also had positive effects. Correlations between these subtasks and ORF scores were variable, which made it difficult to draw conclusions on skills that helped the most in terms of improving reading fluency. With the Kyrgyz Republic QRP, pre-reading skills such as initial letter sounds and oral vocabulary had low correlations with ORF while other skills such as nonsense words and familiar words were highly correlated. The gains for different subtasks and the relations between those subtasks and increases in ORF and comprehension need further study.

2. Support for Classroom Instruction

Guiding Questions

What kinds of instructional methods are leading to improved reading levels? How is the distribution and use of materials related to improved reading? What types of materials? What is the impact of materials and instruction in languages with which children are familiar? How is fidelity of implementation of instructional interventions related to improved reading? Do supplementary materials make a difference in learning gains? What kinds of materials?

Results Summary

The review was inconclusive about the most effective instructional method for improving reading skills. The evaluations reviewed listed a variety of practices, including scripted lessons and active learning, but it was difficult to determine what methods were working better than others. A few of the activities increased time on task for reading and in some cases to reduce teacher absenteeism. Materials were distributed on a widespread basis by most activities, though sometimes not in a timely manner. Many materials were produced in local or national languages; but any gains associated with those materials were difficult to discern.

- *Four main subtopics were identified for support for classroom instruction: time on task, instructional guides, quality materials, and language appropriateness.*

A. Time on Task

- *Increased time-on-task was linked to positive impacts on student learning.*
- *A variety of strategies were used to increase time-on-task.*

Several evaluations reviewed related time-on-task improvements with positive impacts on learning achievement, though the approaches to increasing time-on-task were diverse. Malawi EGRA extended the school day to add time for more reading. Mozambique ApaL reappropriated the school day to spend more time on reading lessons or practice. Nigeria RARA reduced teacher absenteeism, thus increasing

classes meetings, which included reading time. Philippines Basa and Malawi EGRA supported extracurricular reading activities such as homework clubs or reading fairs. Malawi EGRA also promoted student use of reading materials outside of class, thus expanding reading time at home.

While there was some anecdotal evidence in the evaluations reviewed that these activities were helpful in improving reading levels, the evaluations were unable to specifically relate these interventions to reading gains through quantitative evidence. An exception was that increasing time spent reading outside of class did seem to have a positive impact on reading scores. Malawi EGRA and Philippines Basa calculated correlations between these activities and ORF. In Malawi, the evaluators found that students who took books home from school had positive gains in ORF; they also determined that student scores increased with every reading fair that was held at the schools.

Other issues involving time devoted to reading were also noted by some of the evaluations reviewed. On Philippines Basa, the evaluation showed that lower performing schools had fewer hours in school. On Zambia TTL, more time was devoted to literacy skills but positive effects were compromised by overcrowded classrooms.

B. Instructional Guides

- *Evaluations reviewed were not designed to isolate effects associated with instructional guides.*
- *Some limited evidence showed that guides help mitigate challenges from cascade training models.*

Many of the evaluations reviewed, e.g., Jamaica ETP, Malawi EGRA, and Philippines Basa, reported some anecdotal success in combining teacher guides, in-service training, and mentoring. As noted above with Jamaica ETP, evaluators found that the cascade model of training and support broke down before it reached teachers, which left the instructional guides as the only successful aspect of implementation. The significant gains in student learning outcomes in Philippines Basa were likely due to the combinations of interventions. While it is difficult to draw conclusions from a mixed set of information, it appears that training combined with materials provision has a good chance of succeeding since most of the activities that had these two interventions showed student learning gains. An issue to examine more fully is whether the interventions are implemented with a high degree of fidelity, including follow-up on the distribution and use of the instructional guides, which were not discussed in the evaluation report.

C. Materials Distribution²⁷

- *Materials development and distribution interventions were components in most reading activities.*
- *Materials included textbooks, resource books, library materials, pamphlets, and teacher guides.*
- *Timely distribution and replacement of worn materials were problematic for many activities*

New reading materials were features of most of the activities, with classroom use sometimes hampered by production delays and lack of full distribution to all schools.

²⁷ For more information about evaluations that referenced book production, procurement, and distribution, see: Thomaz Alvares de Azevedo and Sean Kelly, *Supplemental Topics from the Synthesis of USAID-Funded: Evaluations: Education sector, 2013–2016* (USAID, January 2018).

Activities generally were successful at producing and delivering reading textbooks—and sometimes supplementary materials—to the schools as part of their reading programs. However, there were instances of delays in producing materials, including for those activities that needed materials in multiple languages, and in promptly distributing materials to all supported schools, such as with Ethiopia READ II (Reading for Ethiopia's Achievement Developed II).

Nearly all evaluations reviewed included materials development and distribution as positive interventions, particularly for improving student learning. Materials included grade-appropriate textbooks, resource books, library books, pamphlets, and guides for teachers. While most of the evaluations reviewed reported on activities that developed new materials, some reported on activities that increased the appropriateness of existing materials. Djibouti Projet AIDE revised existing materials to make them more appropriate to the national context. Malawi EGRA revised materials so that they would be more inclusive of student gender sensitivities and students with disabilities.

As mentioned above, the distribution of the materials in a timely manner as well as replacement of worn materials was noted as problematic in many of the evaluations reviewed. However, on the positive side, the fact that the materials were used—including in activities that allowed students to bring materials home such as Malawi EGRA, Mozambique ApaL, and Rwanda L3—was a positive finding. Activities such as Mozambique ApaL and Rwanda L3 also noted that both classroom materials and teacher guides have a limited durability, and additional distributions of books were a necessity after some years of active use. As such, evaluation recommendations suggested replenishment of teacher and student materials in the schools be included in activity planning and budgets. There was no evidence that materials replacement was implemented on activities.

D. Language Appropriateness

- *Evidence linking language of learning to reading outcomes was lacking in most evaluations reviewed.*
- *Some evaluations reviewed showed positive correlations between use of local languages and improved learning.*

Evidence of language appropriateness linked to improved student learning was lacking in most of the evaluations reviewed. In addition, any generalizations or lessons learned with individual language of instruction effects should be taken with caution since the evaluations reviewed did not provide comparative evidence on instruction in local languages.

Some of the evaluations reviewed provided correlations between the use of local languages and improved learning outcomes. For example, the Malawi EGRA evaluation determined that speaking that local language (Chichewa) at home or with friends was positively correlated with learning outcomes. Similarly, the Nigeria Northern Education Initiative (NEI) evaluation cited survey evidence showing that reading outcomes were better for children who reported Hausa as the language they most commonly speak at home than for peers who report not speaking Hausa at home. Positive trends were also noted in the Philippines Basa evaluation on improvements in learning outcomes due to having some instruction in the local language.

3. Teacher Training

Guiding Questions

Which kinds of teacher training activities are leading to improved reading? How is increased duration and dosage of teacher training related to reading gains? Are there coaching models that appear to be particularly effective in improving reading? Are teacher's guides related to student learning gains? What kinds of guides? Is training to teach in languages that students speak and understand leading to improved reading?

Results Summary

The review found that in-service programs were the most popular and appeared to be the most successful. Pre-service training was often not linked with in-service training and was frequently isolated from the schools. When linkages between different types of teacher training were made, this sometimes led to significant improvements in reading scores. Mentoring and coaching models were not used on many activities, but the evaluations reviewed on those interventions showed some promise. A combination of interventions that were implemented with high fidelity seemed to have the best effects, though the combinations that produced those effects could not be discerned.

- *Almost all evaluations reviewed (18 of 21 activities) included some analysis of a teacher training component, making it the most popular type of intervention. In-service training workshops served as the centerpiece of most activities (also 18 of 21), complemented by in-school mentoring (13 of 21) and district-level coaching (6 of 21). Pre-service training was an intervention in fewer than one-third of activities (6 of 21).*
- *In-school mentoring, district-level coaching, and new teacher induction programs were often cited as follow-up activities to training to ensure skills gained from in-service training workshops and pre-service training were implemented in the classroom. Several evaluations—including Ethiopia READ II, Indonesia PRIORITAS, Jamaica ETP, Malawi EGRA, Mozambique ApaL, and Zambia TTL—reported that beneficiaries considered these follow-up practices critical to behavior change.*
- *Generally, activities that channeled a portion of resources into two or three complementary training components seemed to have a greater chance of success than those that focus exclusively on a single type of training intervention.*

A. Pre-service

- *Student learning was difficult to associate with pre-service training or certification requirements.*
- *Pre-service training when combined with in-service support led to improvements in EGRA scores.*

Evaluations reviewed found that changes to pre-service training and certification requirements were often challenging to judge in relation to student learning outcomes for a variety of reasons, including lack of coordination between teacher training institutes and primary education offices, different curricula and materials used for teacher training and classroom instruction, and low percentages of teacher training graduates entering the teaching force. Those activities, however, that linked pre-service with widespread in-service training—such as Ethiopia IQPEP, Indonesia PRIORITAS, and Rwanda L3—found evidence that the linkages, while often far from ideal, led to improvements in EGRA scores.

Five of the six evaluated activities with pre-service training components also included in-service training components, and the only activity excluding in-service training—Guatemala Reform in the Classroom Project (REAULA)—took place while other development partners were implementing in-service trainings targeting the same population. As mentioned above, the Ethiopia IQPEP activity combined in-service training for current teachers and pre-service reforms for entering teachers to align abilities with expectations through the entire teaching workforce. Joint pre-service and in-service reforms are possible when making specific changes to teacher training, such as new techniques to improve comprehension, as in Ethiopia READ II and Guatemala REAULA, but not when the entire teacher training system needs broader reform. In Liberia LTTP, the evaluation found that the pre-service training component was over-funded and inefficient, with the authors even labeling them as “white elephants.” The Liberia LTTP evaluators called the lack of follow-up monitoring and of oversight of training institutions to be a “fundamental failing” of the teacher training institutions.

B. In-service

- *Most activities implemented in-service teacher training but relationship to learning was unclear.*
- *Evaluations reviewed did not provide sufficient detail on training duration, materials, or follow-up support.*
- *Cascade training models tended to result in lower training quality and limited impact on learning.*
- *Attrition among trained teachers was a common issue.*

Teacher in-service training was the most popular intervention in the sample of activities. Almost all activities organized periodic in-service workshops. Most of these workshops received positive reviews by the teachers, such as with Indonesia PRIORITAS, though there were some problems with unavailability of materials at the time of the training. **Most measurements of teacher behavior change, using tools such as through well-designed classroom observations, were either not conducted (e.g., Ethiopia READ II) or not linked to student learning outcomes (e.g., Indonesia PRIORITAS).**

While almost all evaluations reviewed included reviews of some form of in-service training, sufficient details were not provided in these evaluation documents to adequately synthesize the results across the activities to identify which types of in-service training worked better than others. Descriptive information showed that the 16 activities that provided in-service training also provided some materials distribution, and 14 included in-school mentoring as follow-up support. While this compounding of interventions provided more comprehensive models, it also made it more difficult to determine which interventions contributed most to learning gains. Conversely, it was possible to gain some information from evaluations in which the services were not delivered as planned. For example, in Ethiopia READ II, evaluators were positive about the teacher training interventions but said their effect was limited because materials meant to supplement the training were not distributed to the schools, implying that without materials training impact is limited.

A key issue with in-service training related to the time and structure of the training. **While it would be useful to determine whether intensive teacher trainings with many hours of work are significantly more effective than shorter trainings, the evaluations reviewed did not provide sufficient detail to determine the optimal number of days, hours, or activity types.** This lack of detail was the case in all the evaluations reviewed pertaining to in-service teacher training, most likely due to a lack of a theory of change and subsequent evaluation design. Duration and frequency of in-service training, including the levels of follow-on support provided, should be included in all future

evaluations of training programs so inputs could be prescribed according to optimal impact on classroom practice.²⁸

Several evaluations noted common pitfalls in implementing in-service training. Usually trainings are disseminated through a cascade model, starting with a core group of experts, and continuing through district or regional staff, to headmasters, and then to teachers. As demonstrated by the Jamaica ETP evaluation, if there are too many levels in the cascade, or if a level is under-resourced, the chain breaks down resulting in no or low-quality training and limited impact. In Zambia TTL, trainers at the school levels could not replicate the original training-of-trainers (TOT) because the initial trainings used technology and equipment that were not available at the school level, specifically digital projectors. In addition, the trainers said that critical training materials were not available at the final level of the training cascade.

A final issue common to activities that include training components—whether pre-service or in-service—is retaining teachers for assignments in target schools. For example, in Mozambique ApaL, 30 percent of teachers trained by the activity left their positions within two years of training. Similar issues were reported in Kyrgyz Republic QRP. In Indonesia PRIORITAS, movement among schools was common, resulting in trained teachers relocating out of USAID target regions. **The evaluations reviewed found that training makes teachers more marketable, and teaching in private schools, teaching in higher grades, or taking other positions is advantageous, causing schools to suffer significant leakage.**

C. In-school Mentoring and District-level Coaching

- *Teacher coaching and mentoring models showed promise in teacher behavior change.*
- *Effects on student learning outcomes were not evident, perhaps due to a small body of evidence.*
- *Most evaluations reviewed relating to mentoring and coaching activities revealed logistical and other challenges.*

Some teacher coaching and mentoring models showed promise, but most had implementation difficulties due to local staffing, training, and logistical issues. There were challenges with designing and implementing the teacher coaching models to allow for sufficient support in the schools by the district specialists, such as in Dominican Republic ESP (Effective Schools Program), Guatemala REAULA (Reform in the Classroom Project), Malawi EGRA, and Zambia TTL (Time To Learn). Mentoring models were easier to implement, such as with Jamaica BEP (Basic Education Project). Unfortunately, the coaching and mentoring interventions were usually not evaluated except through anecdotal evidence, as in Liberia LTTP II (Teacher Training Program II). An exception was Kyrgyz Republic QRP, which evaluated its coaching model through ratings of teachers in relation to best teaching practices and then associated those ratings with student scores; however, it found no significant improvements in learning outcomes from teacher coaching.

In general, for the evaluations reviewed, in-school mentoring, which was sometimes called pedagogical leadership, referred to training at the school level by school personnel. District-level coaching referred to outside personnel providing support in the schools during periodic visits. Of the thirteen activities that included a component related to mentoring or coaching, all had in-school mentoring, with six including both mentoring and coaching. Mentoring and coaching were generally used as one among several interventions, and often played a role in conjunction with in-service training workshops, such as in

²⁸ Note that extensive work has been conducted in the U.S. on this topic. In particular, see Garet, Birman, and Yoon, “What makes professional development effective: Results from a national sample of teachers?”

Ethiopia IQPEP, Philippines Basa, and Philippines WSRP. Many qualitative findings in the evaluations suggested that well-implemented mentoring or coaching was successful, but also that implementation was often challenging. **The evaluations reviewed did not show that a particular model—i.e., mentoring, coaching, or induction—was necessarily better than another; factors such as teacher workloads and transportation infrastructure may guide which is more appropriate in different contexts.**

In-school mentoring was considered an important, and usually necessary, follow-on support to training interventions, with examples provided in evaluations of Dominican Republic ESP and Kyrgyz Republic QRP. **However, one important lesson from the reviewed evaluation reports is the need for sufficient time and resources to be given to mentors or coaches as they expand their roles. Activities that simply add this new responsibility onto current tasks for the mentors or coaches were less successful.** For example, in Dominican Republic ESP, mentors were teachers elected to the role within each school, which led to positive reports about the help that the mentors provided, but the reports also stated that the mentors were heavily constrained by their own responsibilities as full-time teachers. The time constraint posed by existing tasks hindered district-level coaches as well as in-school mentors.

District-level coaching appeared to be particularly challenging to implement successfully. The evaluations of Dominican Republic ESP, Guatemala REAULA, and Malawi EGRA found significant struggles in carrying out the field visits by district pedagogical advisors. Only Jamaica ETP, Liberia LTTP II, and Nigeria RARA highlighted coaching as a source of activity success.²⁹ Even where there was a reasonable level of implementation fidelity, results were mixed. Kyrgyz Republic QRP found no significant improvements in learning outcomes from teacher coaching. In contrast, Liberia LTTP II, which had other issues referred to in the sections above, called coaching a “reliable source of instructional supervision” in its schools and a “working model [the] MOE [Ministry of Education] could emulate.” Coaching may be more attractive than in-school mentoring because coaches may have greater expertise; however, constraints mentioned in the evaluations included training/supervision for the coaches themselves and transportation to reach the schools. For example, the evaluation of Zambia TTL noted that the greatest challenge with district-level coaching was the time and logistical complexity of travel. Evaluations in Malawi EGRA and Liberia LTTP II reported similar challenges for school visits.

D. Pedagogy

- *Most activities included teacher training components and provided classroom materials.*
- *Lack of implementation fidelity data was common.*
- *Teachers cited a lack of on-going support and delay in material distribution as primary constraints.*

The evaluations reviewed listed a variety of pedagogical practices—including scripted or structured lesson plans, teacher guides, and active learning methods—that the activities implemented to improve reading. **Most of the pedagogical methods included some sort of training or mentoring component and well as provision of associated classroom materials. However, the evaluations reviewed did not describe these practices in enough detail to distinguish and group the practices, making it difficult to discern which methods are working better than others.** Despite a consensus on the part of the evaluations reviewed that the pedagogical methods used by the activities were successful in improving learning, the evaluations themselves did not point to aspects of the instructional method that improved reading levels. There were few exceptions, such as the

²⁹ The remaining activities did not clearly specify whether coaching was a source of success or failure.

Nigerian RARA activity, where evaluators showed a link between the “I do, we do, you do” methodology and a reduction in ORF zero scores.

Despite overall reporting that implementation of pedagogical practices was successful, there were also instances of a lack of full implementation fidelity. This often revolved around two related pedagogical components: insufficient ongoing training and delayed/insufficient delivery of materials. Examples of this were in Jamaica ETP, Malawi EGRA, Mozambique ApaL, and Nigeria RARA. While teachers may have received initial training, anecdotal and observational data in the evaluations reviewed indicated inadequate follow-up support, therefore, a lack of long-term change in teacher behavior due to insufficient fidelity to promoted practices. Several evaluations reported on classroom observations—either conducted by the evaluators or by the activities—though there were few explanations of how these observations related to the pedagogical practices promoted by the activities. More common were observations that factors such as large class size or lack of classroom materials were barriers to implementing new pedagogical methods, as was the case with Ethiopia IQPEP, Kyrgyz Republic QRP, Malawi EGRA, Mozambique ApaL, and Nigeria RARA.

There were findings in a few of the evaluations stating that teachers appeared to embrace some pedagogical changes more readily than others. The evaluations for Jamaica ETP and Kyrgyz Republic QRP found that teachers changed their pedagogical approach to teaching phonics but not reading comprehension. However, as stated above, the lack of detail in the evaluations reviewed made it difficult to discern patterns across activities in terms of pedagogical practices for this sample of evaluations.

4. Community Engagement

Guiding Questions

What models are being implemented to engage parents and communities? Which models are most effective? Are reading data shared with parents and communities? If so, how do they use this information? Are community engagement activities related to learning gains? What kinds of activities?

Results Summary

The review found that there were many different models, and it was difficult to determine the success of the various approaches due to a lack of data. PTAs and SMCs had some success in providing additional support for reading interventions and monitoring activities at the school level. A few of the interventions attempted extracurricular activities, but the evaluations of these interventions were more descriptive than analytical. Several evaluations discussed sharing information on student academic performance with parents, but they provided little information about how the information was used and whether there were any effects on student learning.

- *Community engagement activities were included in more than two-thirds of the activities evaluated. The most popular activities were tied to parent teacher associations (PTAs, sometimes called parent teacher councils, or PTCs) or school management committees (SMCs). Other interventions, including extracurricular activities and sharing information with parents, were only found in a small number of the evaluation reports reviewed.*

A. PTAs/SMCs and Parental Engagement

- *Community engagement models focused on support to PTAs and SMCs.*
- *These models showed only minimal effects on student learning.*
- *Positive impacts on schools resulted from community initiatives.*

Support to parent-teacher associations (PTAs) and school management committees (SMCs) focused on two areas: 1) school improvement and 2) student learning. The PTAs were more successful in school improvement. Indonesia PRIORITAS, Nigeria NEI, and Rwanda L3 activities included small grant components for infrastructure, while Djibouti Projet AIDE targeted girls' attendance and retention.

While not related directly to a learning goal, several PTA/SMC groups—e.g., Nigeria RARA, Rwanda L3, and Ghana PAGE—were lauded for their monitoring of both student and teacher attendance, which directly impacted time-on-task and, presumably, had a positive effect on learning. Evaluations noted that when community/parents were involved in monitoring activities, teacher absenteeism was reduced and parents were more likely to send their students regularly. Guatemala REAULA worked on community level homework support. Nigeria RARA expanded on this idea, noting that when parents were exposed to classrooms, then they were willing to provide more support for their students, including books and materials like pencils, in addition to ensuring attendance. **Effects of community initiatives on learning outcomes were not examined except through anecdotal information.**

Despite the lack of direct evidence linking community engagement to learning outcomes, most reviewed evaluations recommended more training for parents or community members. Many of the evaluation reports stated that goals of community capacity development were to provide additional reading support outside of the classroom instruction, manage grants and small activities at the schools, monitor attendance, and promote general quality at the school level. Examples were in Djibouti Projet AIDE, Ghana PAGE, Indonesia PRIORITAS, Nigeria NEI, and Rwanda L3. Such training for parents was recommended in response to qualitative data from teachers and principals, who often placed blame for low student outcomes on the parents by noting a lack of parental support and a need for parents to reinforce learning at home.

B. Extracurricular Activities

- *Promotion of reading activities outside the school day showed positive effects on learning.*
- *Limited evidence was found showing positive effects of extracurricular reading on ORF.*

Parents were encouraged to support reading activities outside of the classroom, including reading as homework, on several activities. On a small number of activities, such as Kyrgyz Republic QRP, Philippines Basa, and Malawi EGRA, evaluations showed positive effects of community activities in reading and better student learning outcomes. However, on other activities such as Ethiopia IQPEP (Improving Quality of Primary Education Program) and Guatemala REAULA, evaluators did not examine these interventions except through some anecdotal evidence.

Less than one-fourth of the evaluations reviewed cited interventions involving extracurricular activities (5 of 21). **While potentially useful extracurricular activities were implemented, the evaluations of those activities lacked details, particularly in terms of effects related to student learning outcomes in reading.**

With Kyrgyz Republic QRP, innovative community activities included parents and communities, with activity support, implementing 21-day long summer reading camps and “book days” for students; unfortunately, neither of these interventions was evaluated for effectiveness. With Jamaica ETP, parents were invited into the classrooms to support reading instruction, but again, the effectiveness of the activity was not evaluated, including in relation to learning outcomes. Similarly, Guatemala REAULA provided “schools for parents” with specific training to help parents better support their students' homework, but this was only described and not formally evaluated.

Only one evaluation reviewed made correlations between learning gains and extracurricular activities. For Malawi EGRA, as mentioned above, the evaluation report showed a positive correlation between the number of learning fairs and student performance in reading fluency.

C. Information for Parents

- *Few evaluations reviewed discussed the sharing of information with parents and communities.*
- *When information was shared, evaluations reviewed lacked detail on how that information was used.*

Ghana PAGE used local radio broadcasts to inform parents, advertise after school activities, and to encourage parents to monitor their children’s homework; however, the impact of this intervention was not discussed in the evaluation report. In Liberia LTTP II, a low technology solution to information sharing was to install a dedicated bulletin board at the school where PTA information was posted for the public; this information seemed to be limited to the PTA meeting agenda and minutes—not information related to the school, teachers, or students—but it served as a first step for sharing information between the schools and the communities.

The evaluations reviewed provided little information about how any shared information was actually used. The intention of the school report cards in Nigeria NEI, for example, was to allow parents to compare their school results with results from other schools. However, the school report cards were not yet underway at the time of the evaluation, so the impact on parents’ involvement with their children’s learning was unavailable.

5. Systems and Policy Strengthening

Guiding Questions

What systems have been developed and implemented for monitoring, evaluation, and accountability? Are reading data shared and used to inform policy analysis and management decisions? What kinds of capacity building has taken place and has it been successful?

Results Summary

The reviewed evaluation reports often mentioned improvements in monitoring and evaluation (M&E) by ministries of education, but the review found that the implementation and sustainability of these systems were lacking. Similarly, some activities had interventions related to improved policy analysis involving data, though the uses of the data were generally ineffective, with one or two exceptions. Most reviewed evaluations reported on capacity building, especially with ministry officials, though actually implementing capacity building activities and sustaining interventions were usually unsuccessful.

- *For systems and policy, three categories of results were identified: M&E systems, data sharing and use, and capacity-building.*
- *The general category of capacity building was the most widespread intervention, with other activities—M&E systems and policy and data—implemented in fewer than half of the activities.*

A. M&E Systems

- *Activity support for government monitoring and evaluation systems was generally unsuccessful.*
- *Data were rarely used in efforts to reform policy around reading improvement.*

The evaluations reviewed found that building monitoring and evaluation (M&E) systems in support of reading initiatives with Ministries of Education was rarely done. In Liberia LTTP II, systems were established to monitor reading materials at the school level, but the monitoring largely failed since the results were not reported. As with the district coaching models described above, some activities found that ministry officials were not able to travel to the field to collect data, such as in the case of Malawi EGRA. In addition to a lack of data, when data were available, sharing those data for policy reform and decision-making was generally not done. For instance, in Mozambique ApaL and Rwanda L3, extensive data were collected but the Ministries were unable to apply the data to their policy analysis and planning.

M&E activities were mentioned in several evaluations reviewed, but building M&E systems with ministries of education was rarely done, particularly in support of reading initiatives.

For instance, under Nigeria NEI, communications and monitoring systems were established between the state level and schools, which the evaluation noted was successfully helping to increase enrollment and retention rates; however, there was no information on whether this improved reading levels. Liberia LTTP II supported improved education management information systems (EMIS), and efforts were made at monitoring materials at the school level through the Principal's Monthly Report, but it largely failed; in fact, the evaluation stated that monitoring educational quality at any level was largely "impossible," and the activity data only related to access to education and not reading improvements.

Guatemala REAULA established an M&E system by training decentralized officials and deploying computers to the departments (districts). However, the activity lacked links back to central ministry offices and the evaluation voiced concerns over a lack of ongoing management support to sustain the decentralized system. With Mozambique ApaL, the final evaluation made a recommendation to strengthen M&E systems since nothing had been done to help the Ministry in collecting their own data on reading activities. Malawi EGRA supported the design of the ministry's M&E systems, but the evaluation stated that the lack of roads and transportation prevented district coordinators from actually visiting the schools to collect information, and some education officers even had difficulty attending the M&E training the activity provided on data collection and use. The findings in Malawi EGRA reflected issues cited above involving transportation and logistical challenges faced by ministry officials in supporting various activity interventions, such as M&E, but also in-service teacher training and district-level coaching.

B. Policy and Data

- *Evaluations reviewed found minimal evidence of data sharing.*
- *Data use for policy reform and decision-making was limited even when data were available.*

According to the evaluation report, Djibouti Projet AIDE was the main exception in that they had considerable success with strategic planning using information. The EMIS was improved at the central level and user-friendly EMIS software was installed and used by regional focal points. The report stated that the regional officials found the system difficult to implement at the beginning but they were eventually able to gain the necessary skills, along with an appreciation of the national data which were aggregated and then communicated back to the regions. With Djibouti Projet AIDE, the report stated that the EMIS data were used in conjunction with the national EGRA through the preparation of sampling

frames and entering school-level reading results into the EMIS database. At the same time, the evaluators stated that other EGRA-related information, such as from teacher and head teacher surveys, was not used effectively. Another exception was Ghana PAGE, where data on teachers, head teachers, and classrooms were collected from all 46 activity districts, and head teachers said that these data were used to generate more discussions with decision-makers about policy and reforms.

The Nigeria NEI evaluation stated that the institutional environment was generally not supportive of education, and the use of data was minimal at best. In Liberia LTTP II, several research studies were conducted to support policy reform and improve programming to boost reading; however, even though plans were developed in these areas, there was little progress due to weak management systems, budgetary constraints, inattention to reforms (including to decentralization), and high personnel turnover in the Ministry of Education. In fact, in interviews conducted for the Liberia LTTP II evaluation, few MOE officials had even heard of the new Education Reform Act (ERA), which was intended as the blueprint for education reform in the country. Similarly, with Malawi EGRA, there was a strong focus on the new National Reading Strategy (NRS) but the evaluation stated that few schools took the ministry policy seriously and there was limited implementation of action plans.

With Guatemala REAULA, the policy environment was cited in the evaluation report as “weak.” The evaluation noted that the ministry did not have the experience, expertise, or political will to carry out a strategic planning process. With Rwanda L3, there were extensive data collected by the activity on issues such as attendance, repetition, distances to school, and class size; however, there was little evidence that these data and a subsequent descriptive analysis informed policy analysis or decision-making. Similarly, with Mozambique ApaL, data were collected on issues such as class size, repetition, achievement, and absenteeism, and there was a statistical report on the findings, but no evidence that the data impacted policy decisions.

With Mozambique ApaL, there was an intervention to produce data through the School Management Assessment (SMA) but the evaluation stated that the data were not used effectively. Very little feedback was provided from the schools through the national EMIS on the implementation of the reading strategy. For instance, there was a policy to increase the amount of time for reading in the classrooms, but the evaluation did not have information on whether the policy was implemented.

C. Capacity Building

- *Most activities had interventions related to building capacity in the MOE.*
- *Evaluations reviewed showed little evidence of government ability to take over reading interventions.*

Nigeria NEI had a clear objective of capacity strengthening, but the evaluation stated that institutionalization and sustainability of education activities was still at an incipient stage at the time of the evaluation. A recommendation was made to develop an institutional sustainability plan identifying a government team to operate interventions, but this was not implemented by activity close-out. Mozambique ApaL made the recommendation of planning for sustainability, but it was not clear from the evaluation report whether such planning ever took place. There was some focus on sustainability at the school level but the evaluation stated that the interventions were only minimally supported with on-going assistance and that the reading interventions were not maintained.

As noted, an exception was Djibouti Projet AIDE, where capacity building was an explicit part of the activity. Nearly 200 people in the ministry were trained on strategic planning through an initiative called Fundamental Quality and Equity Levels (FQEL). High turnover hampered progress but the ministry maintained a core group of officials to analyze and use information. This included staff working on EMIS,

which was developed in collaboration with the ministry. There was also substantial training for decentralized education planning officials in the regions, including training for school directors to plan, budget, and monitor school performance. The evaluation report stated that the activity successfully built the capacity of the MOE to manage the FQEL reforms.

6. Crosscutting Themes

Researched Themes	Results Summary
<ul style="list-style-type: none"> • Gender • Disability • ICT • Innovative finance • Scaling and sustainability 	<p>The review found that interventions related to gender issues were the most successful, particularly for improving gender equality in the classroom. Many evaluations reviewed reported reducing gaps in learning between boys and girls. Other crosscutting interventions were not as widely implemented, such as those targeting students with disabilities, innovative financing, or information technology. Scaling up and sustainability were frequently cited in the reports, though most evaluations reviewed emphasized that more work should be done in these areas, particularly by focusing on sustainability early in activity cycles and planning for scaling up of effective interventions.</p>

- *Crosscutting issues were found inconsistently in the reviewed evaluation reports. Over half of the activities reported activities in gender and scaling/sustainability but disability, innovating financing, and information technology activities were only found on fewer than one-third of the reports. For more information, see Table 8 in Annex 3.*

A. Gender³⁰

- *Evaluations reviewed generally showed positive trends toward reducing gender gaps in reading scores.*
- *Gender awareness activities included training focused in girls' education, gender-balanced materials, and positive gender modelling.*

Although most activities implement some interventions related to girls' education, across the evaluation reports reviewed, reporting of gender-related issues was inconsistent and most evaluations mentioned limitations to attribution due to other donor activities also focusing on girls' education, often in collaboration with USAID implementing partners. **Teacher training with gender awareness was the most common gender-related activity, followed by modification of teaching and learning materials to ensure a gender balance and positive modelling for girls.** Overall, most activities showed positive results, specifically in terms of trends toward reducing gender gaps in reading scores, yet over half of the evaluations reviewed recommended improving gender awareness implementation, particularly in relation to girls' education.

³⁰ Guiding questions included: Are evaluations adequately examining gender dynamics? To what degree do evaluations find differential results for male versus female beneficiaries, both in terms of activity access and outcomes? What contextual and implementation factors help explain gendered outcomes? Has school-related gender-based violence been examined? What effects does school-related gender based violence have on reading outcomes?

Several activities included creative gender-related approaches that were evaluated as successful in the reports. In Djibouti Projet AIDE, interventions included the development of teacher training modules and pamphlets promoting gender awareness, with the provision of school kits especially designed for girls, scholarships for girls to stay in school and continue their education, and an associated girls' mentorship program. Ethiopia IQPEP supported girls' education advisory committees, girls' clubs, and dedicated spaces for girls in schools. Ghana PAGE, for instance, conducted a complete gender analysis, but it was only cited in the evaluation report as an annex, with no discussion of findings in the main body of the report.

The Indonesia PRIORITAS evaluation identified a potential for reverse discrimination, given that participation rates and academic performance for boys lagged behind those of girls. The evaluators suggested this may have been a result of teachers' increased gender awareness and their resulting emphasis on increasing girls' participation in class at the expense of boys. In Philippines Basa and WSRP, as well as Kyrgyz Republic QRP, a reverse gender divide—with boys behind girls in reading scores—was identified. In Nigeria RARA, gender results in learning outcomes were mixed, with girls' academic performance higher in one target province with boys scoring higher in the other target province. In all other activities, the boys generally outperformed the girls, though interventions were identified as having helped to narrow achievement gaps, particularly in the early primary grades.

As mentioned in the learning outcomes section, many of the student scores were disaggregated by gender. Given the transparency of these data which were shared within the education structures and school communities of some activities, understanding of gender issues will continue to develop. As this happens, more detailed correlations between activities and gender improvements may become possible.

B. Disability³¹

- *Few activities included interventions targeting students with disability.*
- *Teachers reported greater awareness of disability issues but requested additional assistance in addressing student needs in classrooms.*

Only about one-fourth of the evaluations reviewed (5 of 21) mentioned disability in the reports. Of the activity approaches involving disability, there were some activities, such as conducting classroom assessments for both gender and disability issues at the same time and then revising materials with the Ministry to increase gender sensitivity and improve services to students with disabilities.

About one-fourth of the evaluations reviewed discussed activities involving students with disabilities (often referred to as special needs students). Dominican Republic ESP worked toward reducing violence in school-related settings, which included violence and bullying toward students with disabilities. Results included improvements in students' self-esteem, better behavior and communications skills, a greater acceptance of children with disabilities, and a reduction in bullying. Teachers reported having a greater awareness of disability issues, but requested additional assistance in addressing student needs in the classroom.

In Indonesia PRIORITAS, teachers noted the presence of students with disabilities in their classrooms but said that there was a lack of technical expertise to provide adequate education solutions for these students. At the same time, the Indonesia OVC activity, which had a specific focus on the acquisition of reading skills by students with disabilities, provided training to teachers who were subsequently “more

³¹ Guiding questions included: To what degree do activities address disabled populations and their needs? How is disability treated in evaluations, including in the evaluation questions? Is disability included as a sampling stratum? Is there a separate data collection module? Is disability discussed separately in findings/conclusions/recommendations?

patient” and reported (in anecdotal evidence from interviews) that their students with disabilities had made progress in acquiring reading and other academic skills.

Perhaps the most innovative approach to supporting students with disabilities was recorded by Malawi EGRA, where, after an assessment of available teaching and learning materials, a Braille version of the textbooks was developed. Furthermore, a review of existing classroom materials included an assessment of issues around both gender and disability. Following this review, some of the materials were revised in collaboration with the Ministry of Education and teachers reported increased awareness and knowledge on how to serve students with disabilities.

All activities that targeted more inclusive education—including Dominican Republic ESP, Indonesia OVC, Indonesia PRIORITAS, Jamaica ETP, and Malawi EGRA—identified progress in terms of increased awareness, but all evaluations of these activities noted there remained room for improvement.

C. Information and Communications Technology³²

- *Classroom use of Information and Communications Technology (ICT) was evaluated in only three activities.*
- *Three other activities used IT in relation to EMIS systems for scaling and sustainability.*
- *Maintenance was cited as an ongoing issue with computers but cellphone technology is promising.*

Information technology (IT) was a focus of the activities in less than one-third of the evaluations reviewed. Three activities—Djibouti Projet AIDE, Liberia LTTP II, and Nigeria NEI—emphasized EMIS systems, and three other evaluations—Ethiopia IQPEP, Rwanda L3, and Tanzania Bridge IT—described technology use at the school level. All reported that much of the equipment stopped working due to lack of maintenance. Conversely, use of cell phones by teachers to access instructional materials (Rwanda L3) and use of tablets by district supervisors to collect reading data (Malawi EGRA) were shown to have strong possibilities, particularly as the apps became more user-friendly.

The use of EMIS data was reported as it related to implementation, scaling, and sustainability. As mentioned above, the Djibouti Projet AIDE evaluation cited several EMIS-related successes on the activity. Nigeria NEI provided a success story in which high quality information was processed and shared in a transparent manner. In Liberia LTTP II, conversely, EMIS was cited as lacking quality standards and thus limiting the potential use of available data.

With Ethiopia IQPEP, evaluators found minimal IT infrastructure in the schools, even though there had been efforts at improving IT at the school level. Computer rooms were described as seriously lacking in working computers, much of which was the product of poor maintenance. Even when available and functioning, observations at the Ethiopian schools reported little evidence that students or faculty had made significant use of computers at the school level. Conversely, the Rwanda L3 initiative provided cell phones, speakers, and data cards to teachers, most of whom reported that they used the technology multiple times per week in their classrooms, providing a case study of successful use of appropriate technology for schools.

³² Guiding questions included: To what degree do activities utilize ICT in intervention modalities, and what are the main uses? Is there any evidence of the effectiveness of ICT-based approaches versus face-to-face training and interventions?

D. Innovative Finance³³

- *The evaluation reports reviewed did not include meaningful information on innovative finance.*

There were some explorations of innovative financing on activities but there were no activities that actually implemented such activities. In fact, financing was only analyzed on one activity, and that was a non-innovative analysis of unit costs associated with reading improvements.

Only one evaluation contained information regarding a financing-related education activity. Malawi EGRA analyzed unit cost of incremental improvements in reading outcomes. According to the evaluation, however, this cost estimate should be interpreted with caution since the activity was in early intervention stages and total unit costs included activity costs not directly related to reading improvements.

E. Scaling and Sustainability³⁴

- *Capacity-building and scaling up were features of most activities.*
- *Resulting ability of countries to sustain/expand successful interventions was often lacking.*

Nigeria NEI had a clear objective of capacity strengthening, but the evaluation stated that institutionalization and sustainability were still at an incipient stage during the activity. Similarly, Mozambique ApaL developed a plan for sustainability but the evaluation provided no evidence that sustainability activities, e.g., transfer of responsibilities during the activity, took place. An exception was Djibouti Projet AIDE, which explicitly included capacity building at the central and decentralized levels with the goal of having the Ministry implement many of the activities before the end of the activity.

Over three-quarters of the activities evaluated contained design elements that promoted scaling or sustainability to ensure that successful activities would continue beyond the period of performance. These activities usually included interventions targeting central ministry or decentralized government offices and schools, with training to strengthen capacity in areas such as instruction, production, training, and information systems. A small number of activities, such as Djibouti Projet AIDE and Rwanda L3, had personnel who were embedded within the ministries and **activity staff collaborated with ministry colleagues daily to build capacity and increase sustainability.**

Most evaluations reviewed noted that more work should be done on improving the sustainability of activity-funded interventions, i.e., build on activities involving sustainability that were taking place. Examples of this were Ethiopia READ II, Indonesia PRIORITAS, Malawi EGRA, Mozambique ApaL, Philippines Basa, Tanzania Bridge-IT, and Zambia TTL. Financing constraints were usually cited as a limiting factor for sustainability, as were political shifts or transfer of personnel that tended to undermine the development of ownership of activity initiatives, such as with Djibouti Projet AIDE. One reason for a lack of sustainability may have been that the implementation of most sustainability-related interventions was weighted toward the final period of the activity so that the interventions did not receive necessary attention.

³³ Guiding questions included: To what degree have the use of private versus public funds been included as part of project design? What have been any innovative financing schemes, if any, and what have the results been?

³⁴ Guiding questions included: How many of the evaluated activities have been scaled? What have been the results of the scaling processes? With which kinds of partners? How effective have they been at scale? How sustainable? Do evaluations of activities that are promising but have not been scaled make recommendations concerning scaling? What patterns exist if any in these recommendations on future scaling? To what extent do the projects focus on sustainability? How do they define it and measure it? To what extent is sustainability a focus of evaluations? What are the findings and recommendations relating to sustainability?

OBJECTIVE 2: SYNTHESIS OF FINDINGS AND LESSONS LEARNED RELATED TO USAID EDUCATION STRATEGY GOAL 2 (WFD AND HIGHER EDUCATION)

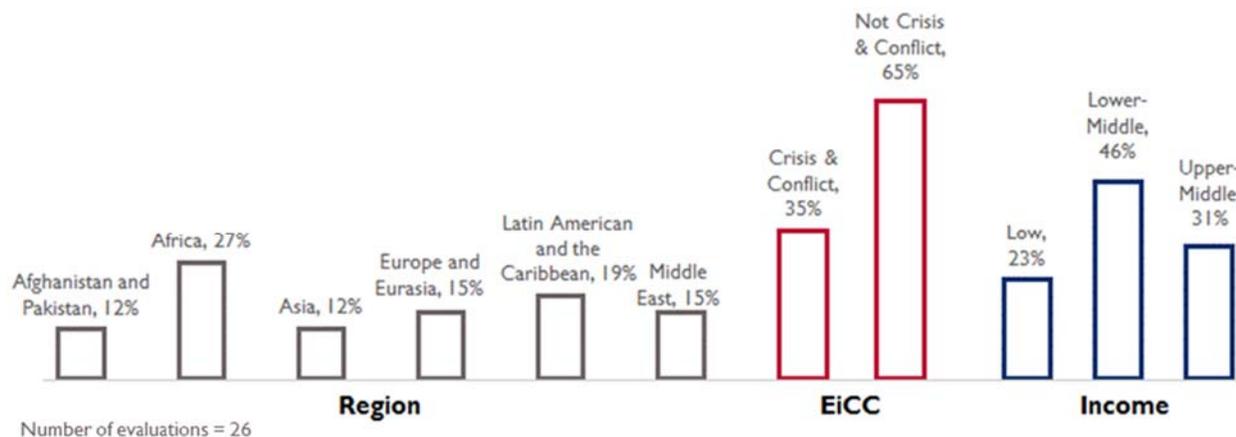
Goal 2 programming addresses the high-level Education Strategy goal of “improved ability of *tertiary and WfD* programs to generate workforce skills relevant to a country’s development goals.” This review synthesizes the results for 26 evaluations covering 30 activities related to Goal 2. Overall, this constitutes a medium-sized body of evidence.

Evaluation Sample, Context and Limitations

All Goal 2 related evaluations reviewed in this study were funded by USAID and met the Office of Education criteria for evaluation quality. These included performance and impact evaluations published between 2013 and 2016.

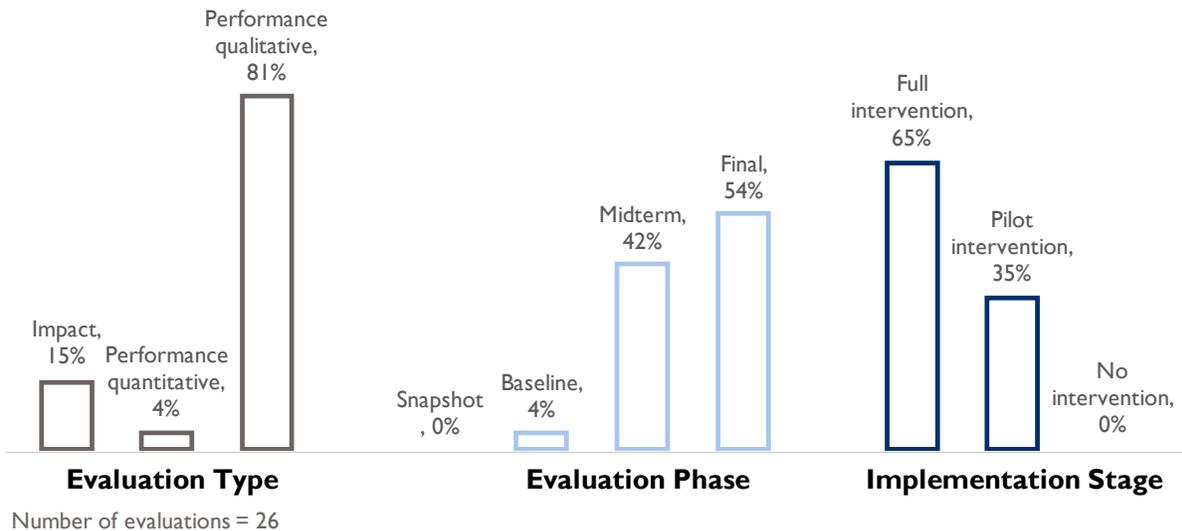
The Goal 2 review drew on global evidence with a mix of regions that are roughly reflective of USAID investment (recognizing that the availability of evaluations is not necessarily identical to the size of USAID investment in programs). Most of the body of evidence comes from environments considered as not in crisis or conflict (65 percent) and from lower-middle or low-income countries (69 percent).

FIGURE 7: PERCENTAGE OF EVALUATIONS REVIEWED UNDER GOAL 2 BY REGION, CRISIS AND CONFLICT STATUS, AND COUNTRY INCOME



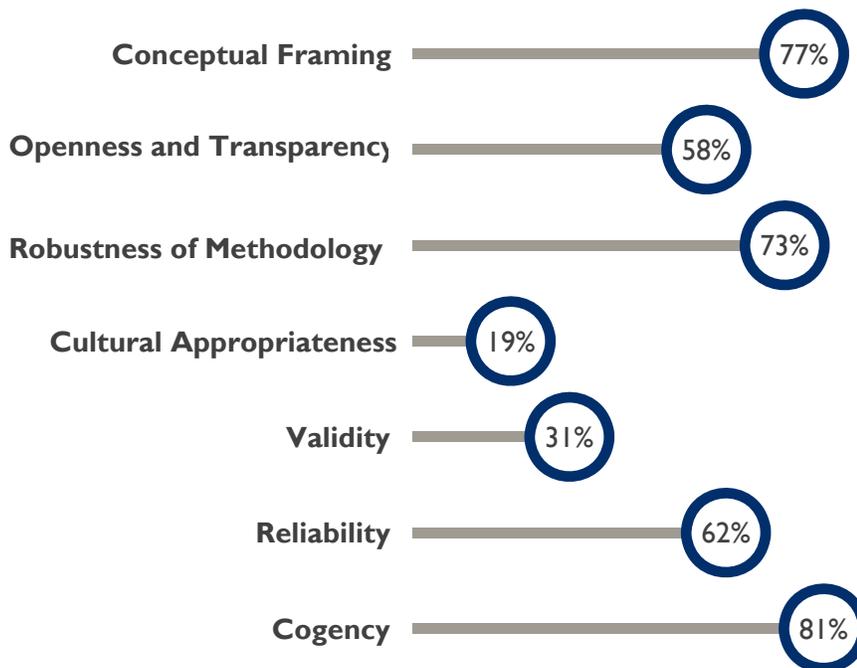
The body of the evidence for the Goal 2 review comes mostly from qualitative performance evaluations (81 percent). However, only about half are final evaluations (54 percent) and about a third are pilots (35 percent).

FIGURE 8: PERCENTAGE OF EVALUATIONS REVIEWED UNDER GOAL 2 BY EVALUATION TYPE AND EVALUATION/IMPLEMENTATION PHASE



Based on the expert judgments of the seven principles of quality, cogency (81 percent) was rated as “adequate” most frequently, while cultural appropriateness (19 percent) was rated as “adequate” the least. Results on Figure 9 indicates that despite the Office of Education’s minimum quality standards for inclusion in the synthesis, the general quality of the body of evidence is only moderate.

FIGURE 9: PERCENTAGE OF EVALUATIONS REVIEWED UNDER GOAL 2 BY PRINCIPLES OF QUALITY



This review provides a landscape perspective on the range of activities conducted under Goal 2, though not necessarily proportional to all USAID investment in Goal 2 programming. This is because many activities have not been evaluated, evaluations are not yet available, or evaluations did not clear the Phase I quality review. However, drawing on available evaluations, this synthesis does provide a snapshot perspective on trends in outcomes achieved, both short-falls and accomplishments. In some cases, a clear view of results was compromised by the lack of data available, usually because activities were not structured to collect the types of data needed to make the strongest judgment about their effectiveness. For example, a workforce readiness program that trains adolescent youth before they reach the labor market is typically not designed to track the youth years from program completion when their success in the labor market becomes most apparent. Furthermore, many areas of Goal 2 programming are notoriously difficult to measure in a rigorous manner, such as soft skills development in young people, institutional reform in higher education entities, policy reform, and the quality of public-private partnerships.

Also, Goal 2 activities are extremely diverse in their design, target beneficiary, and outcomes—in part due to diverse funding sources—and as a result, measurement indicators are inconsistent across activities. This lack of consistency in outcome measurement makes generalization difficult.

Finally, there are very few experimental or quasi-experimental evaluations of Goal 2 activities designed to isolate the effects of specific interventions or combinations of interventions, as well as the effects of various levels of treatment duration and intensity. Therefore, at this stage it is not possible to render a judgment about the effectiveness of specific Goal 2 interventions that could be reliably generalized to other contexts.

On the other hand, USAID is currently investing in the development of more robust youth soft skills assessment tools which should increase and improve the measurement of these critical intermediate outcomes. While they may first be piloted in workforce development activities, they are unquestionably relevant to higher education activities seeking to build a more flexible and adaptable country workforce.

Intervention Modalities

Goal 2 programming relied on three activity types: workforce development (WfD) for economic growth and employment/self-employment, WfD for violence prevention and countering violent extremism (CVE),³⁵ and higher education (HE). The Goal 2 evaluations in this review reflect what is true for WfD and higher education programming across the Agency: that is often cross-sectoral and funded beyond what are strictly Education Sector funds. For example, WfD activities for violence prevention and CVE may be funded with security-oriented funds, or WfD activities in agriculture may be funded by Feed the Future Initiative funds. The implication of diverse funding sources is that activity theories of change and, with them, outcome monitoring and evaluation tend to be diverse. This diversity makes the sector's ability to aggregate evidence around the effectiveness of specific interventions more challenging. The purpose of this review is to show where there are commonalities in evaluating Goal 2 outcomes and where there are gaps or inconsistencies.

³⁵ According to the USAID Violent Extremism and Insurgency Policy (2011) and the USAID Crime and Violence Prevention Field Guide (2016), violence prevention and countering violent extremism are two distinct outcomes though there may be commonalities in programming. CVE refers to programs that address social, economic, governance, and other grievances that can fuel violent extremism or radicalization of individuals and communities. Violence prevention consists of actions and interventions that seek to decrease or eliminate underlying risk factors that lead to violent and/or criminal behavior, and that promote citizen security. Violence can be political, gang-related, and/or inter-personal (including domestic).

A. Workforce Development Activities

These interventions usually comprise a mix of skills training, employment services³⁶, and entrepreneurship promotion³⁷ aimed at enhancing the livelihoods of individuals, the vast majority of whom are young people aged 15-35 years old.³⁸ Every WfD activity in the evaluations reviewed had some skills training components, though they varied widely in scope and intensity, and typically these were complemented by employment services, and less commonly by entrepreneurship promotion. Skills training includes technical skills,³⁹ soft skills,⁴⁰ life skills,⁴¹ workforce readiness,⁴² vocational skills,⁴³ and basic skills.⁴⁴ See Tables 9-12 in Annex 4 for the range and frequency of types of WfD activities including skills training, employment services, and entrepreneurship promotion.

In addition to individual services, all WfD interventions included some element of institutional change. WfD activities strive to create systemic and institutional change through public-private partnerships with businesses or community-based organizations,⁴⁵ cost sharing with the private sector,⁴⁶ demand-driven curriculums,⁴⁷ industry associations,⁴⁸ labor market assessments,⁴⁹ WfD policy authorities,⁵⁰ labor market information systems,⁵¹ policy reforms,⁵² technical and vocational education and training (TVET) capacity

³⁶ Employment services encompass activities such as internships, mentoring, apprenticeships, job placement, career services and guidance, job fairs, and networking. Within employment services, most of the intervention components are self-explanatory. “Career services” involves counseling and guidance for choosing and pursuing professional opportunities, and “networking” concerns structured opportunities for youth or young professionals to meet and speak with experienced professionals in particular sectors, or, in the case of youth in village livelihoods, to link to new markets for selling goods.

³⁷ Entrepreneurship promotion consists of entrepreneurship skills development, business coaching, and access to finance. Entrepreneurship skills development includes training on operating a small business such as basic accounting and marketing. Business coaching involves seasoned business owners’ offering their advice and guidance to nascent entrepreneurs. Access to finance refers to an activity’s directly providing loans or grants or facilitating access to loans or grants through a third party.

³⁸ This categorization follows Kluge et al., *Interventions to Improve the Labour Market Outcomes of Youth: A Systematic Review of Training, Entrepreneurship Promotion, Employment Services, and Subsidized Employment Interventions*. None of the USAID evaluations reviewed for this synthesis included subsidized employment interventions, so that category is not included.

³⁹ Within skills training, technical skills refer to job skills applicable to a wide array of professional jobs such as using computer hardware and software.

⁴⁰ Soft skills refer transferable skills, attitudes and behaviors that enable people to be successful on the job (such as critical thinking, teamwork, communication, self-control, positive self-concept, etc.).

⁴¹ Life skills are sometimes used synonymously with soft skills though they often refer to practical coping skills such as personal hygiene, health, and household finances.

⁴² Workforce readiness may also include soft skills but adds skills needed to land a job such as job-searching, CV-writing, and interview skills.

⁴³ Vocational skills refer to skills development toward a specific occupation, trade, or craft.

⁴⁴ Basic skills are literacy (reading and writing) and numeracy.

⁴⁵ Public-private partnerships are initiatives in which government agencies form formal or informal relationships with the private sector or NGOs to deliver WfD services.

⁴⁶ Cost share occurs when the private sector or NGOs contribute their own institutional funding toward activity implementation.

⁴⁷ Demand-driven curriculum refers to skills training curriculum that are developed to train youth and young professionals in skills that the private sector asserts are in high demand.

⁴⁸ Industry associations are initiatives by activities to utilize existing private sector associations or create new ones in order to advance activity objectives.

⁴⁹ Labor market assessments include formal and informal efforts to gauge current and future skills needs by employers to inform activity services.

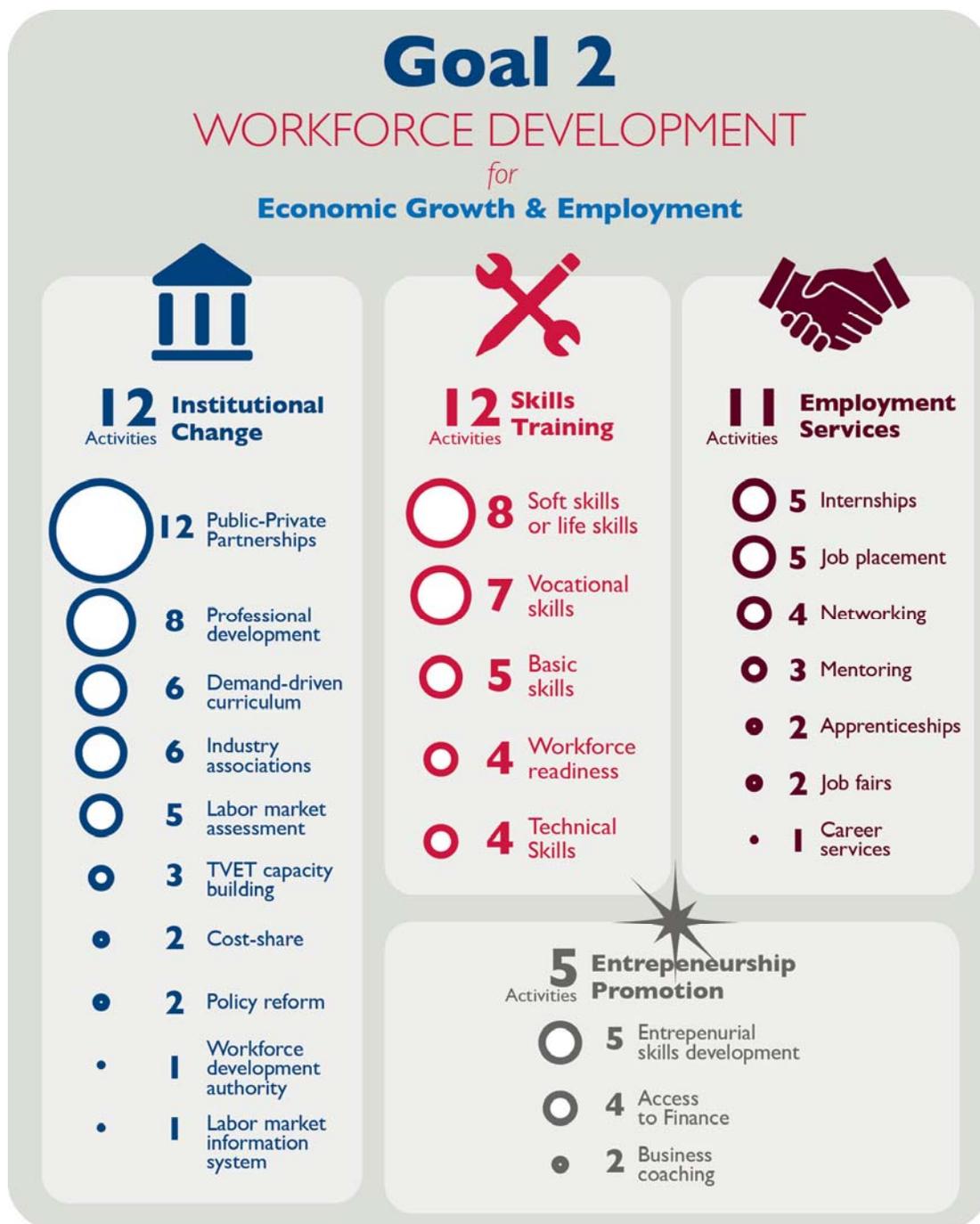
⁵⁰ WfD authority means the explicit inclusion of a national WfD authority in some element of activity delivery.

⁵¹ Labor market information system (LMIS) designates the creation a new LMIS in a country.

⁵² Policy reform includes efforts by an activity to change a country’s existing regulations, laws, or institutional approach to WfD issues.

building,⁵³ and professional development.⁵⁴ See Table 13 in Annex 4 for the range and frequency of types of WfD systems and institutional change.

FIGURE 10: GOAL 2 INTERVENTION MODALITIES RELATED TO ECONOMIC GROWTH AND EMPLOYMENT



⁵³ TVET capacity building are targeted efforts to improve the quality of vocational training provide by government institutions, private firms, or NGOs.

⁵⁴ Professional development occurs when an activity improves the skills or capacity of staff within government, NGOs, or private sector entities participating in the WfD activity.

B. Workforce Development for Violence Prevention and Countering Violent Extremism

The WfD activities for violence prevention and CVE aimed to reduce violence or extremism through improvements to youth employability and employment or livelihoods. WfD activities are combined with other peace-building interventions such as juvenile justice reform, community and police engagement, and youth civic engagement in local decision-making. The main outcomes of interest in these activities pertain to crime, violence, and youth engagement. Consistent with WfD activities for economic growth, WfD activities for violence prevention and CVE were largely focused on skills training. However, in contrast to WfD activities for economic growth, WfD activities for violence prevention and CVE more frequently paired entrepreneurship promotion with skills training. The WfD components of violence prevention and CVE activities did not typically feature institutional change as a component. See Tables 14-17 in Annex 4 for the range and frequency of types of WfD services for violence prevention and CVE services in skills training, employment services, entrepreneurship promotion, and other peace-building activities.

FIGURE 11: GOAL 2 INTERVENTION MODALITIES RELATED TO VIOLENCE PREVENTION AND CVE



C. Higher Education Activities

The components of higher education activities in the evaluations reviewed differed from WfD interventions. Higher education activities focused primarily on faculty training and institutional capacity building. These efforts involved supporting the creation of new academic programs or the strengthening of existing courses of study to match international norms or the skills demanded in the private sector.

Higher education activities also built partnerships with private sector firms and other organizations within their countries, provided external training in agriculture and education to participants outside of the university, and improved university research capacity. See Table 18 in Annex 4 for the range and frequency of higher education activities.

**FIGURE 12: GOAL 2
INTERVENTION MODALITIES
FOR HIGHER EDUCATION**



Findings

I. WfD: Responsiveness to Labor Market Demands

Guiding Questions

To what extent are WfD activities responsive to local, national, and international labor market demand as demonstrated by private sector engagement in activities and strong job placement for graduates? What have been the most successful modalities for private sector engagement including cost-share? What are the most common activity modalities for demand-driven skills development? Given that soft skills have been in demand globally, to what extent is soft skills development integrated into education and training? Is there any evidence for the difference in outcomes that soft skills training provides when combined with other intervention components?

- *Skills development activities often integrate many types of skills, imparting them to youth through experiential methods to meet labor market demand. However, because demand is weak, youth are often stigmatized, and labor market information systems are weak, job placement remains challenging.*
- *Nearly all activities focus on linking youth trainees to employment through active engagement with the private sector through diverse modalities.*

Results Summary

The review found that activities often conducted labor market assessments, consulted with local businesses, linked trainees to internships and jobs, and established experiential learning for secondary and university students.

A. Investing in a Mix of Skills

- *USAID-funded activities fostered a rich mix of skills in youth including soft skills, technical skills, basic literacy and work-readiness skills, depending on the needs of both youth and the labor market demand.*

Goal 2 activities are investing in a comprehensive mix of skills, based on both the needs of targeted beneficiaries and (mostly) local labor markets. USAID activities are providing the holistic mix of skills that global research indicates is needed to ensure youth success in rapidly changing labor markets, including technical skills, soft skills, basic literacy, and work-readiness skills. Soft skills, life skills, basic skills, and work-readiness skills—all of which are cross-functional—are provided as frequently as technical or vocational skills, which are more occupation-focused. Approximately 80 percent of WfD activities that provide skills training deliver two or more skill types; for WfD activities providing skills development, four focus on one type of skill, while 15 focus on between two and four skill types, for WfD for violence prevention and CVE, five focus on two skill types while the remaining two focus on four skills types.

For example, Nepal EIG (Education for Income Generation) combined basic literacy and life skills training with vocational training for employment and entrepreneurship, and technical skills for raising agricultural productivity. A crosscutting theme in all trainings was peace and reconciliation education to promote attitudes to reduce conflict and promote human rights.

B. Mainstreaming Private Sector Engagement and Job Placement

- *Skills training was virtually always linked to employment services and/or entrepreneurship development to link youth to labor market opportunities.*
- *Work-learning opportunities including internships, apprenticeships, mentoring, and networking were increased and improved in terms of quality.*
- *Private sector businesses and business associations were engaged by education and training providers to improve training quality, identify job openings, offer mentoring and develop internships and mentor.*

Activities are mainstreaming private sector engagement and linking training with placement. **All evaluated WfD activities included in this synthesis pair skills training with either employment services (e.g., job placement, internships, career networking) or entrepreneurship support.** Furthermore, engagement of the private sector has been an expected part of all activities, and they do this in diverse ways. This represents major progress from earlier generations of workforce activities that provided training without engaging employers and without follow-up services to ensure that trainees obtain jobs. **Job placement rates were not always consistently tracked, but virtually every evaluation report reviewed for Goal 2 activities included in this synthesis includes design elements that actively connect young people to the labor market.**

There is evidence that activities are conducting consultations with local businesses to design or revise skills training curriculum, although it was impossible to determine from the evaluations alone how many activities conducted labor market assessments as part of activity design (or during the early stages of implementation). Interventions engaging the private sector were nonetheless common. For example, the Afghanistan AWDP (Afghanistan Workforce Development Program) required its training partners to consult with local businesses and use their knowledge of the local labor market to develop TVET training that meets market demand. As a result, the program exceeded its target for job placement for its 8,200

trainees by 9 percent. The program was, however, more successful in securing raises for those currently employed (86 percent) in comparison to placing new job seekers in jobs (only 26 percent).

Internships, apprenticeships, and entrepreneurship development were important experiential approaches used to train youth in areas that meet market demand, be it for employment or self-employment. The Azerbaijan YBLP (Youth Business Leadership Program) developed partnerships with British Petroleum Azerbaijan for program funding, and with the American Chamber of Commerce to support internships for over 100 university students in 32 companies. When surveyed, their mentors believed the internships would make the students competitive for entry-level jobs after graduation.⁵⁵ Junior Achievement Armenia provided applied economics and business training as extracurricular activities coupled with student-run businesses. The students founded and ran 100 businesses that sold their products at trade fairs organized by private sector partners. Notwithstanding the positive outputs, the ultimate employment outcomes for the YBLP and Junior Achievement Armenia are not known, as these would require a tracer study.

While there is evidence that WfD activities are striving to gauge and respond to labor market demand, results in terms of job placements are not straightforward given the challenges of developing country labor markets. For example, Jordan's Youth for the Future aligned its skills training content to the National Employment Strategy 2011-2020.⁵⁶ Evaluators found the youth who underwent the training to be "in a better position now" to find jobs in the hospitality, retail sales, and heating, ventilation and air conditioning (HVAC) sectors. However, the employment opportunities for these jobs were limited outside of urban areas, and HVAC positions were deemed culturally unacceptable for women. Youth for the Future surveyed 352 participants and found that 47 percent were employed and 27 percent were studying in an educational program. Surprisingly, 29 percent of the employed cited the program as the reason for their job, while 71 percent credited other means.⁵⁷

There were other challenges in aligning skills training with employment and increased earnings. Liberia's Advancing Youth Project identified livelihoods opportunities through a youth-led labor market assessment and provided short-term skills training in soap making, beekeeping, food preservation, candle making, and baking. Unfortunately, livelihoods data was not collected so it is difficult to know how useful these trainings were.

The Iraq Foras Project conducted a labor market assessment that identified a need for employability skills in entry-level jobs. The activity provided online skills training, including training courses focused on human resources, retail, sales, security, and hospitality. While training participants gave positive reviews of the courses, a survey identified 150 participants from these courses and found that only 13 percent had secured jobs as a result of the courses.⁵⁸

Nepal EIG built formal partnerships with 146 employers and business associations which appeared to bear fruit in terms of beneficiary earnings. Although no comparative data is available for non-participants, the activity reported mean individual income increases from 1,836 to 27,724 rupees and mean household income increases from 107,187 in 2010 to 170,936 rupees (across all activity components). The activity estimated the employment rate for all participants at 76.3 percent.⁵⁹

Through close collaboration with government partners, Djibouti AIDE identified businesses in construction, hospitality, IT, and shipping that ultimately provided internships for 240 participants as part

⁵⁵ Statman and Babayev, *Final Performance Evaluation of the Youth Business Leadership Project (YBLP) in Azerbaijan*, iii-iv.

⁵⁶ Roberts et al., *Mid-Term Performance Evaluation Afghanistan Workforce Development Program*, 12-14.

⁵⁷ Moubayed et al., *Final Performance Evaluation of the USAID/Jordan Youth for the Future (Y4F) Project*, 19-21, 29.

⁵⁸ QED Group, *USAID/Iraq Foras Project Final Performance Evaluation Report*, 19-20.

⁵⁹ Population, Health, and Development Group, *Final Evaluation Report Education for Income Generation*, 8-10, 21.

of a pilot project. The firms generally viewed the training as effective in producing skilled, hard-working employees.⁶⁰

Honduras METAS forged formal partnerships with businesses so that METAS graduates could obtain jobs and internships. In particular, the internships strengthened the job search skills of METAS youth, who experienced a 21.8 percent increase in job search skills, a statistically significant increase in comparison with a control group. Morocco's FORSATY (Favorable Opportunities to Reinforce Self-Advancement for Today's Youth) activity had one particularly fruitful partnership with a textile firm in which 98 percent of youth believed the internship was a positive experience and 57 percent received employment after the internship. Company officials attributed the youth's maturity and preparedness for the work environment to the activity's life skills component provided by local out-of-school youth-serving organizations.⁶¹

2. WfD: Entrepreneurship Integration

Guiding Questions

To what degree are activities incorporating entrepreneurship training and support into technical training or general education? What are the most common modalities for entrepreneurship skills development, by participant and institutional type? What have been the results?

Results Summary

The review found that activities often promoted initiatives to train youth to write business plans and run a business, and a few offered coaching to start-ups, training in agricultural inputs, counseling on the use of credit and leasing arrangements, and financing.

- *Entrepreneurship training and support has been provided to beneficiaries with diverse characteristics: from university students to out-of-school youth in conflict-affected settings. Numbers of businesses established has been tracked, but rarely are their success rates tracked.*
- *In-school youth benefit from entrepreneurship learning experiences designed to promote critical thinking and creativity, without the immediate expectation of establishing viable businesses.*
- *Youth livelihoods development outside of formal educational institutions often occurs in networks of youth-led associations, but earning results are mixed.*

Activities are integrating entrepreneurship into skills training, general education, and youth engagement interventions. USAID activities are responding to global research and country analysis of labor market conditions that point to the need for a greater focus on entrepreneurship. About 50 percent of WfD activities (9 of 21)⁶² and over 75 percent of VP/CVE-WfD activities include entrepreneurship promotion. Employer demand for more flexible thinking and behavior in the workforce, coupled with the paucity of formal sector jobs in many developing countries, has prompted a stronger emphasis on entrepreneurship promotion. **USAID interventions include support for development of entrepreneurial mindsets and skills, business coaching, linkages to markets for improved livelihoods, and access to finance.**

⁶⁰ McClellan and Elabeh, *Projet AIDE Performance Evaluation 2009-2013*, 29.

⁶¹ Boyle, Bouasla, and Abderebbi, *Mid-Term Evaluation Favorable Opportunities to Reinforce Self-Advancement for Today's Youth*, 16, 31.

⁶² A Ganar in Honduras, Akazi Kanoze in Rwanda, Education for Income Generation (EIG) Program in Nepal, Empowering Jamaica's Youth, Georgia Economic Prosperity Initiative (EIP), Iraq Foras Project, Junior Achievement Armenia, Somali Youth Leadership Program (YLP) in Somalia, and Yes Youth Can in Kenya.

These are often provided in combination with general education and vocational skills training.

Entrepreneurship programming is very diverse and usually involved forming productive linkages in support of youth businesses. Junior Achievement activities in Armenia and Jamaica cultivated private sector partnerships to create practical opportunities for youth to market and sell their products. Similarly, Iraq Foras employed start-up weekends and business competitions for emerging entrepreneurs, but also attempted to facilitate access to microloans for winners, albeit with limited success. The Georgia EPI (Economic Prosperity Initiative) connected small- and medium enterprises (SMEs) and farmers with financial advisors so that the businesses could receive sound financial advice and successfully apply for small business loans from banks. Nepal EIG linked its farmers to markets and agriculture support services from other businesses. Finally, Yes Youth Can in Kenya required youth associations (called *bunges*) interested in additional financial support to form savings and credit associations (SACCOs) and deposit savings into accounts that the activity would match.

The results from these 10 evaluations fell into 4 categories: no results reported, output results, performance outcomes (lacking valid comparison groups or total attribution), and rigorous outcomes. Iraq Foras did not track or report results specific to entrepreneurship. Three activities had output results: Junior Achievement Armenia produced 100 student-run businesses;⁶³ Georgia EPI assisted three firms to access leasing and spurred \$170,000 in loans to farmers and SMEs; and Empowering Jamaica's Youth trained 40,000 students in 200 schools in applied economics and entrepreneurship.⁶⁴ Performance outcomes came from Somalia SYLP and Nepal EIG. SYLP reported the share of youth claiming no income dropped from 88.7 percent to 14.6 percent and the mean income of youth increased from \$12.20 to \$146.40.⁶⁵ It is unclear how much of this achievement resulted from the entrepreneurship training. EIG reported that participants experienced an increase in mean individual income—R1,386 at baseline to R21,518 at endline—and mean household income of R126,189 in 2010 to R191,220 in 2012. The activity estimated an employment rate of 85.7 percent for participants in the agriculture and enterprise component.⁶⁶

Finally, Akazi Kanoze and Yes Youth Can reported rigorous outcomes based on valid comparison groups. For Akazi Kanoze, there was a statistically significant difference-in-difference of 15 percentage points between treatment and control groups for employment, but an insignificant difference-in-difference of 4 percentage points between treatment and control for self-employment.⁶⁷ Yes Youth Can analyzed household income and an asset index, but it did not distinguish self-employment in the results even though 60 percent of surveyed participants claimed to have participated in income generating activities with their *bunge*. For household income, there was a statistically insignificant difference between treatment and control. The asset index showed a statistically significant difference at the 5 percent level between treatment and control. On average, youth who participated had experienced an increase in their assets relative to youth who had not participated in Yes Youth Can, though the effect size was very small.⁶⁸

⁶³ Turner et al., *Mid-Term Performance Evaluation for Junior Achievement of Armenia*, 25.

⁶⁴ Social Impact, *Assessment of the Caribbean Basin Security Initiative*, 71.

⁶⁵ Swedberg and Reisman, *Mid-Term Evaluation of Three Countering Violent Extremism Projects*, 54.

⁶⁶ Population, Health, and Development Group, *Final Evaluation Report Education for Income Generation*, 8-10, 21.

⁶⁷ Alcid, *A Randomized Control Trial of Akazi Kanoze Youth in Rwanda*, 23, 29.

⁶⁸ NORC at the University of Chicago, *Yes Youth Can! Impact Evaluation Final Report*, 15, 20-12

3. WfD: Access for Marginalized Groups

Guiding Questions

To what degree do activities reach marginalized communities and nontraditional learners? How do activities tailor their interventions to ensure learner success? What aspects of activities contributed to increased learner success, including graduation rates? What kinds of financial assistance activities (including scholarships) have been provided and what have been the results of these?

- *Second-chance education, training, and employment services target marginalized youth who face a range of risks, including gang membership, recruitment by violent networks, and extreme poverty.*
- *Non-formal WfD activities address the stigmatization of out-of-school youth by mediating between the youth and potential employers, and by building positive self-confident through life skills workshops.*

Results Summary

The review found that at-risk youth were most commonly targeted, with some activities focusing on girls/women, conflict-affected youth and internally displaced persons (IDPs).

The majority of workforce development activities focus on marginalized youth. Nearly two-thirds of WfD activities, and virtually all WfD activities for violence prevention and CVE, targeted marginalized youth: either young people who have dropped out of school and are unemployed, or in-school youth who are at-risk of engagement with gangs or violent networks.⁶⁹ Youth living in regions affected by conflict and youth from stigmatized groups such as low-caste Indians or ethnic minority groups have been targeted by activities.

Nineteen of the 30 evaluations were of activities focusing on marginalized or nontraditional learners. By far the largest group is disadvantaged or at-risk youth, which includes out-of-school youth, unemployed youth, and youth failing in school; these youth are targeted by 13 activities.⁷⁰ Some challenges faced by youth include stigmatization due to their neighborhood/region of residence, race, or ethnicity, fewer networks for access to higher education or jobs, geographic distance from educational opportunities and jobs, lack of assets to pay for education fees or acquire loans, and trauma due to stressful environments and poverty. Internally displaced people (IDPs) and conflict-affected populations are served by the Advancing Youth Project in Liberia, EIG in Nepal, and Iraq Foras. EIG also reached 421 Dalit (low-caste) youth, half of whom, as a result of the activity, were enrolled in higher education with scholarships, and another 20 percent were employed.

Activities for disadvantaged youth were generally not significantly different from activities for other youth in that they involve combinations of employment or entrepreneurship training, life skills training, mentoring, scholarships, job placement, remedial education, and internships or apprenticeships. Many evaluations reviewed provided anecdotal evidence of the challenges of meeting the needs of marginalized youth, but there were only a few cases of systemic testing of strategies designed to enable these youth to remain in learning or training environments. One is Liberia's Advancing Youth Program which provided an extensive adult

⁶⁹ When one considers both marginalized and mainstream youth, aged 12-35 years, all reviewed evaluations of WfD activities focused on youth, though two activities were not framed as youth activities per se (in Afghanistan and Iraq.)

⁷⁰ Alerta Joven in the Dominican Republic, A Ganar in Honduras, Advancing Youth Program in Liberia, Education for Success in Nicaragua, FORSATY in Morocco, Garissa Youth in Kenya, KTI-Eastleigh in Kenya, Proyecto METAS in Honduras, Skills and Knowledge for Youth Employment in Guyana, Somali Youth Leadership Program in Somalia, University Scholarship Program in Lebanon, Yes Youth Can in Kenya, and Youth for the Future in Jordan

basic education curriculum with work-readiness training to young adults who had missed education due to the conflict. While enrollment numbers are high, there is little evidence that many youth have been able to complete the program, given the pressure for earning a living and caring for family. This non-formal education program, conducted in collaboration with the Ministry of Education, was an important avenue for continuing education for older youth, though the livelihood effects were unclear. A second example of a tailored activity comes from Iraq. Given the difficult security situation in Iraq, Foras delivered technical skills training to IDPs in northern Iraq and ensured its job platform is compatible with Facebook and mobile web browsers.

4. WfD: Policy and Systems Strengthening

Guiding Questions

To what degree do activities include systems strengthening as part of the activity, including establishment or strengthening of coordinating bodies; administrative or organizational capacity building; curricular and pedagogical improvement; policy reform; social messaging; advocacy; and partnership development? What are the major modalities for capacity-building of formal and non-formal TVET providers, including HICD (human and institutional capacity development) approaches? What have been the results of these efforts and how have they measured?

- *Reform of both formal and non-formal technical-vocational education and training (TVET) provision involved facilitating linkages between TVET entities and the local private sector for curriculum modernization and internship development.*
- *Few activities involved efforts to improve macro-level labor market failure or workforce development policy.*

Results Summary

The review found that while nearly all activities included at least some institutional capacity-building and private-sector partnership, there were few examples of diagnoses and well-defined theories of change for systemic inefficiencies, such as bottlenecks, information gaps, lack of coordination, or disincentives for better functioning workforce systems.

While institutional capacity building and partnerships are common, there is a lack of strategic focus on systems strengthening. WfD activities typically focus on delivering training and/or employment services to young people, while at the same time working to build the effectiveness and reach of the larger WfD ecosystems; nearly all activities included institutional capacity building and private-sector partnerships alongside service delivery. However, **there were few examples of well-defined theories of change related to the needs of country or subnational workforce systems.** Typically, activities were designed on the general assumption that more demand-driven training was needed and that services for disadvantaged youth needed to be expanded. Neither of these assumptions are incorrect: all education systems are becoming more responsive to labor market demand as global competitiveness increases, and nearly all developing countries are experiencing a demographic youth bulge such that investment in youth is needed. However, **there was little evidence of assessment and diagnosis of systemic inefficiencies, such as bottlenecks, information gaps, lack of coordination, or disincentives for a better functioning system, that might have been conducted prior to activity design.**⁷¹ As a result, systems-strengthening activities tended to

⁷¹ An example of a comprehensive workforce development systems analysis is the World Bank's "[Workforce Development SABER](#)."

be diffuse and without well-crafted monitoring and evaluation. Few evaluations reviewed described in much detail the specific interventions utilized to improve systems.

There were some exceptions to the diffuseness of systems-strengthening. One exception is Iraq Foras, which introduced a web portal for job matching because a significant gap in labor market information was identified. Another exception among the WfD for violence prevention and CVE activities is the Eastern Caribbean CARSI-funded activity that pilots an innovative public health model. Rather than looking at crime as only a security issue, the model considers primary, secondary and tertiary levels of risk factors facing youth, and introduces an ecological model that brings together the individual, family, and community.

WfD activities often provided professional development and institutional capacity building for workforce training providers and/or relevant local or national governmental agencies. Capacity-building efforts were usually supported by small grants to local technical and vocational education and training (TVET) providers. Institutional capacity was described by a panoply of qualitative indicators that include demand-driven curricula, inter-ministry coordination, partnership development and management capacity. While findings are generally positive, there was little evidence of rigorous measurement of progress. For example, a typical approach for measuring progress in partnership development is a catalog of partners involved in the activity's implementation, from which a number of partnerships is inferred. The nature, productivity, and sustainability of these partnerships was not measured.

The AWPDP initiative in Afghanistan provided grants for 10-15 TVETs in which grantees were required to show that trainers and curriculum developers were of high caliber, the training facilities were upgraded, and courses were based on modern, demand-driven curricula and related to current employer demand. A series of incremental benchmarks and learning activities such as coaching and workshops were provided to support TVET grantees' development. Not unlike TVET in most developing countries, AWPDP grantees faced the challenge of transforming instruction from theory-based to active learning and competency-based pedagogies. The introduction of private sector consultation, employment services, job placement, and post-activity tracking of trainees in the labor market represented large shifts for hitherto supply-sided activities.

The Advancing Youth Project sought to improve instructional quality by training teachers, but it only monitored the number of Ministry of Education teachers and teachers trained at rural teacher training institutes without assessing how the training affected their classroom instruction. Iraq Foras was the only WfD activity to develop systems, specifically an online job portal, which was judged a positive result due to its creation and large number of registered users. Projet AIDE assisted the Ministry of Education to link its efforts with the Ministry of Labor's Employment Office, and to expand its efforts to increase the employment opportunities of Djiboutian women.

In terms of WfD for violence prevention and CVE, five activities had capacity building results: USAID CVE activities and Yes Youth Can. The USAID CVE activities had indicators for institutional capacity, partnerships, and sustainability, while Yes Youth Can had one indicator for institutional capacity related to *bunges*. Institutional capacity for USAID CVE activities was conceptualized as management capacity, and both G-Youth (Garissa Youth Project) and SYLP reported positive results. Yes Youth Can judged institutional capacity result as negative because implementing partners and *bunges* struggled to engage and coordinate effectively. The USAID CVE activities assessed partnerships qualitatively by describing engagement with government officials and private sector partners.

Very few of the evaluations reviewed included efforts at targeted policy reforms. One exception was the efforts of FADCANIC (Foundation for the Autonomy and Development of the Atlantic Coast) in Nicaragua which focused on developing an effective model for at-risk youth development for replication across the southern Caribbean coast by diverse partners. Across the evaluations, there was no evidence

of support for the creation of national-level advisory bodies in WfD. Social messaging and advocacy appeared only as an intervention for peace building, and not (as it has been used elsewhere) as a vehicle for reducing stigma associated with TVET.

5. WfD: Responsiveness to Needs and Dynamics in Conflict-Affected Regions

Guiding Questions

Conflict-affected situations include post-conflict recovery (e.g., Liberia), violent extremism (e.g., Pakistan), political violence (e.g., Kenya), and gang-related violence (e.g., Honduras). Have activities been implemented differently in conflict-affected regions than in stable countries, such as inclusion of specific services (such as psycho-social support) and/or adaptive implementation approaches (such as the use of risk assessments)? Have evaluations of these activities gone beyond capturing traditional WfD outcomes to include conflict-related results, such as reduction of risky behavior and perceptions of citizen security? How have conflict-related dynamics affected workforce results?

Results Summary

The review found that WfD interventions appear as significant components of activities focusing on violence prevention, post-conflict recovery, and countering violent extremism. They seek to prevent youth participation in violent activities as well as to mitigate some of the negative labor market and community conditions affecting young people as the result of conflict and crime. Although sustainable livelihoods and workforce participation are thought to be promotive factors for stability and peace, few activity designs made explicit linkages between WfD components and violence prevention and CVE outcomes.

- *The relationship between employment, and violence prevention and countering violence extremism is complex and unique to each situation, and may in some instances be counter-intuitive. As a result, there is evidence that nuanced youth assessments were conducted before activity designs and that they did inform the choice of interventions.*
- *Some of the soft skills associated with conflict-related protective factors for youth, such as self-efficacy and civic engagement, were also important for youth labor market success. But these skills alone did not ensure employment or higher earnings.*

The links between youth employment and violence prevention/countering violent extremism are not well understood. WfD activities for violence prevention and CVE were grounded on theories of change that assume that higher levels of youth, family and community engagement can reduce youth vulnerability to joining gangs or networks espousing violent extremism. Although sustainable livelihoods and workforce participation are thought to be promotive factors for stability and peace, few activity designs made explicit linkages between WfD components and violence prevention outcomes, and only one evaluation—though a significant one—was structured to elucidate this link. Livelihood and youth employment outcomes are typically reported in parallel with other outcomes predictive of violence (such as attitudes and risky behavior), but not in concert with them to understand the interaction between them.

- The Kenya Yes Youth Can evaluation found that youth remained active in youth associations (*bunges*) even though they did not experience improvement in their employment or earnings. They did, however, believe that as a result of their experience with their *bunges* their political views were taken more seriously by those around them and that trust had improved between themselves and their communities. While the findings are discouraging from a livelihood point of view—interventions were admittedly very light in this area, the findings are encouraging for social cohesion and trust which are critical for building stability in Kenya.

- The East Africa CVE evaluation shows positive effects for civic engagement, and to a lesser degree, identity and the value of youth associations. Here, “identity” includes youths’ perceptions of their preparedness for the labor market and belief in education and training as more important than family connection for finding a job. All three activities showed some significant positive effects for youth in the full treatment group versus the comparison group.

The discussion of the links between youth employment and violence prevention/countering violent extremism in the body of evidence reviewed suggests that this topic is not yet well-understood. WfD activities have been implemented in conflict-affected countries to try both to prevent youth participation in violent activities and to mitigate some of the negative labor market and community conditions affecting young people as the result of conflict and crime. Three types of conflict-related contexts were represented in the evaluations reviewed: post-conflict recovery (e.g., Liberia), violent extremism (e.g., Pakistan, Afghanistan), and gang-related violence (e.g., Honduras). All three types of conflict or violence are different, and every country context differs in terms of conflict drivers, even within the same broad type of conflict or violence. Therefore, this review examined to what degree the theory of change for those conflict or violence prevention activities that included WfD interventions explicated *how* the WfD components contributed to the conflict/security-related results.

The Caribbean Basin Security Initiative (CBSI) report is not strictly an activity evaluation but rather an assessment of 23 activities across 10 Caribbean countries designed to reduce illicit trafficking, increase public safety, and promote social justice. At least four of these are youth WfD activities (Alerta Joven in the Dominican Republic, SKYE [Skills and Knowledge for Youth Employment] in Guyana, A Ganar in the eastern Caribbean, and Empowering Jamaica’s Youth).⁷² While the report does not explicate a theory of change linking WfD to violence prevention, the Eastern Caribbean Mission is implementing an innovative public health model that considers primary, secondary and tertiary levels of targeting risk factors, and introduces an ecological model that brings together the individual, family, community and broader society.⁷³

The A Ganar activity in Honduras is a seven- to nine-month, four-phase integrated job training activity that combines sports-based field and classroom activities, vocational training, internships/apprenticeships, service training, mentoring, entrepreneurship workshops, and various follow-on activities. The A Ganar theory of change involves mutually reinforcing increases in employability, life skills, technical skills, work experience, and relationships, which lead, in turn, to the primary outcomes of increased employment, re-entry into the formal school system, and entrepreneurship, as well as the secondary outcomes of changed gender roles and perceptions and increased security (i.e., reduced risky behavior). The activity hypothesizes that sport can act as an incentivizing force, engaging and retaining participants who may not have otherwise applied or stayed with a traditional WfD program.⁷⁴ Retention of marginalized young people in pro-social activities that are engaging to them is one important element of WfD activities that simultaneously seek to prevent youth from participating in risky or violent behavior.

Three CVE activities⁷⁵ in East Africa (Kenya and Somalia) are reviewed together under the USAID CVE policy framework of “push” and “pull” factors that draw youth into networks espousing violence.⁷⁶ Workforce and livelihoods development and civic engagement are thought to be intertwined with these push factors (grievance, marginalization, frustration) and pull factors (social status and material reward for joining violent groups). The evaluation framework examines five factors or drivers of violence extremism: civic engagement, perception of efficacy, value of youth associations, identity, and rejection of the

⁷² Social Impact, *Assessment of the Caribbean Basin Security Initiative*, 65-71.

⁷³ *Ibid.*, 4.

⁷⁴ Duthie, Pucilowski, and Murphy-Graham, *A Ganar Alliance Impact Evaluation Summative Baseline Report*, 11-12.

⁷⁵ The Kenya Transition Initiative – Eastleigh, the Somalia Youth Livelihoods Program, and the Garissa Youth Project.

⁷⁶ Swedberg and Reisman, *Mid-Term Evaluation of Three Countering Violent Extremism Projects*, 15.

violence in the name of Islam. Rigorous evaluation methods were used to make comparisons across three groups of youth: full beneficiaries, partial beneficiaries, and nonparticipants. Full beneficiaries in the activities demonstrated significantly higher levels of engagement with local authorities and moderate levels of advantage in efficacy, identity, and the power of youth association in comparison to the comparison group.⁷⁷ However, higher rates of engagement with low levels of efficacy (belief on one's own power to effect change) may indicate that more emphasis needs to be placed on working with authorities to be more responsive to youth priorities and on opening channels of communication.⁷⁸ Traditional WfD outcomes were reported in the course of the evaluation, and they were largely positive, though they were not correlated with the CVE measures. This may be a topic of analysis for future studies.

Both the CVE and gang prevention WfD activities identify engagement of local stakeholders—local governments, NGOs, civic associations, religious and clan leaders, and local businesses—as vital to activity success. The CVE evaluation notes that the lack of engagement of family members represents a critical gap in the evaluated activities,⁷⁹ considering the key role families play in identity formation in youth, as well as their proximity to youth for detecting changes in youth dispositions. Both the KYI-E and the G-Youth activities faced challenges with local capacity building of NGOs and youth associations receiving grants, as well as struggles over whether to provide stipends for youth attendance at trainings. These challenges are very critical to anticipate and resolve in a context where there is a pervasive and heightened sense of grievance that may erupt in violence.

Evaluators of Kenya's Yes Youth Can activity determined self-efficacy based on the Generalized Self-Efficacy Scale, a four-point scale with a strong academic literature base. While they could not characterize the difference as large or small, the results meant that participants in Yes Youth Can had a greater sense of self-efficacy, a fundamental soft skill highly correlated with many positive adult outcomes. However, while overall impacts were roughly similar for male and female youth participants, women showed smaller gains in self-efficacy.

6. WfD: Measurement Issues

Guiding Questions	Results Summary
<p>What are the overall results in terms of beneficiary outcomes (e.g. employment; self-employment; training; program completion/certification)? What are the measures (indicators and assessments) used in evaluations to track outcome results?</p>	<p>The review found that outcome results for WfD include employment and self-employment and were reported in most reviewed evaluations of WfD activities; however, there was little use of comparative data needed to evaluate effectiveness. Furthermore, results for earning were seldom reported, and all but three WfD activities had inconsistent reporting of enrollment and program completion figures.</p>
<ul style="list-style-type: none"> ➤ <i>Diverse measures of workforce success have hindered the aggregation of results to allow comparability across activities.</i> ➤ <i>The lack of experimental and quasi-experimental evaluation designs has hindered our ability to judge the effectiveness of specific intervention components.</i> 	

⁷⁷ Ibid, 5.

⁷⁸ Ibid, 7.

⁷⁹ Ibid, 7.

A. Weak Outcome Measurement of Employment and Workforce Readiness

- *Employment and earnings outcome data was rarely presented with comparable benchmarks so judgment of the ultimate labor market success of participants is difficult.*
- *Workforce readiness or employability is often treated as an intermediate outcome especially where youth are too young for employment. Few if any rigorous measures exist for this area of skill development, so anecdotal evidence is often very ample.*

Weak outcome measurement of employment and workforce readiness makes it difficult, if not impossible, to evaluate the full effectiveness of workforce development activities. Employment outcomes were reported by 9 out of 12 activities, but they are presented as employment rates or number of people employed without comparison groups or baselines. For the vast majority of evaluations reviewed, it was impossible to render an objective judgment as to whether the activities' employment results were positive or negative, because, absent the activity, young people will experience some change in employment status or earnings over time. The question is whether the activity improved their prospects beyond the opportunity cost of participating in the activity. In some cases, activities did not collect employment data at all.⁸⁰ On the other hand, reviewed evaluations do typically report positive qualitative data about youth satisfaction with employability and life skills components—though this should not be accepted as the full evaluation of activity effectiveness.

- Of 12 activities, 3 had objectives in employability that did not necessitate collection of employment outcome data. Of the remaining nine evaluations, seven used poor employment data, one used high-quality data, and one was of moderate quality.⁸¹
- Liberia Advancing Youth Project served some 22,256 young adults in a non-formal education program, coupled with WfD training and livelihood support; however, output data show that only 3,103 completed the program and there is virtually no employment/earnings outcome data. The evaluation recommended a post-program tracer study, as well as more attention in the future to rigorous outcomes data monitoring.
- The Jordan Youth for the Future evaluation found, through phone interviews with 352 participants, that 47 percent had found jobs, but only 29 percent of these attributed this to help from the activity. It is difficult to determine whether the activity was successful without a comparison with the employment rates for similar youth.

Employment outcomes were reported by 9 of the 12 WfD activities. (The Georgia EPI did not track employment outcomes, and employment was not a relevant outcome for Junior Achievement Armenia or the Youth Business Leadership Project.) The indicators used to monitor employment were the number of people employed, employment rate, and income. Akazi Kanoze and Proyecto METAS presented rigorous, positive results using experimental research designs with treatment and control groups. Nicaragua's Education for Success produced the only negative result, a rigorous finding that the vulnerable youth the activity targeted struggled to find wage paying employment relative to the comparison group. Afghanistan AWDP and Nepal EIG had positive results based on comparisons of pre-activity baselines with post-activity endlines. Seven evaluations presented employment rates or number of people employed without comparison groups or baselines. In those cases, it was impossible to render a judgment as to whether the activity's results were positive or negative. Nonetheless, some activities appeared to perform well, such as the Nepal EIG, which cited 76 percent of youth employed, while Iraq

⁸⁰ For some activities, collection of employment data is not relevant because participants are too young or in-school so employment would not be expected by activity completion.

⁸¹ Here, quality means the activity data is comparative with either a control or comparison group.

Foras claimed that 13 percent of trainees gained employment or improved their economic status. The employment indicators were typically disaggregated by gender.

Seven activities focused on self-employment, but only four evaluations detailed self-employment outcomes. The dominant indicator for self-employment was an employment rate, but Nepal EIG reported income for the self-employed as a comparison of baseline and endline. There were three rigorous results for self-employment: Akazi Kanoze, Education for Success, and Proyecto METAS. All self-employment outcomes were disaggregated by gender and some differentiated between geographic locations. It's important to note that Akazi Kanoze was found to provide a protective factor: both intervention and control groups dropped in employment, but the Akazi Kanoze group dropped less. Without a control group, it could have been judged as a failure.

Only three of the nine violence prevention and CVE activities reported employment or livelihood outcomes: FORSATY, SYLP, and Yes Youth Can. FORSATY reported the number of young people employed, but without a baseline comparison. The SYLP reported a positive comparison based on a tracer study: the share of participants earning no income declined 74.1 percent and mean income increased from \$12.20 to \$146.40. Yes Youth Can reported income and asset ownership indicators for employment. In the case of income, there was no statistically significant difference between treatment and control groups. However, the asset ownership demonstrated a slightly positive and significant result between groups. While the engagement aspects of the activity were considered positive, the livelihoods dimensions were not.

None of the WfD for violence prevention and CVE activity evaluations reviewed presented self-employment outcome data. However, the data from FORSATY suggest that some participants were self-employed and data from Yes Youth Can indicate that a substantial share of participants derived their income from agriculture or running small businesses.

Intermediate results for work-readiness largely derive from measures of satisfaction of students, instructors, and employers, but there are instances of evidence from experimental and quasi-experimental methods. Focus group discussions for the Advancing Youth Project showed that students felt work readiness training enabled them to access job opportunities. The course instructors highlighted a positive change in the students' ability to lead, communicate, and work as a team.⁸² A quasi-experimental work readiness assessment showed activity participants had a three-fold gain in their mean score versus a comparison group.

A survey of 181 FORSATY participants found that 47 percent had training on preparing CVs, job searching, and interviewing. Among the 47 percent, more than three-quarters agreed that work readiness training was useful. One human resources officer in Tangiers asserted that the work readiness and life skills training made FORSATY participants higher-caliber applicants for vocational jobs in the free trade zone.⁸³ Data collected from students, instructors, and private sector mentors of the YBLP show that the work readiness training was highly valued.⁸⁴ Surveys of Youth for the Future participants show that they overwhelmingly believe the life skills and work readiness training were beneficial and useful for self-confidence, sense of responsibility, motivation, and communication. The evaluation concluded that the training in life skills and work readiness combined with vocational training and internships had improved the position of participants as they transitioned to the workforce.⁸⁵ In general, youth who studied the basic skills curriculum had greater work readiness skills than a comparison group. However, reliability tests for the work readiness assessment produced low scores for internal consistency, which makes

⁸² Auten, Dabla, and Bassie, *Advancing Youth Project Performance Evaluation*, 35.

⁸³ Boyle, Bouasla, Abderebbi, *Mid-Term Evaluation Favorable Opportunities to Reinforce Self-Advancement for Today's Youth*, 16-17.

⁸⁴ Statman and Babyev, *Final Performance Evaluation of the Youth Business Leadership Project (YBLP) in Azerbaijan*, 8-9.

⁸⁵ Moubayed et al., *Final Performance Evaluation of the USAID/Jordan Youth for the Future (Y4F) Project*, 21-23.

definitive judgment about the effectiveness of the interventions difficult.⁸⁶ Akazi Kanoze's randomized control trial produced rigorous results. There were statistically significant gains for the treatment group in comparison to the control in three areas: proficiency in applying for a job, knowledge of business development plan components, and understanding of marketing strategy and tactics.⁸⁷ Overall, the evidence for work readiness represents a spectrum of rigor. Importantly, most activities that endeavored to inculcate work readiness skills attempted to substantiate outcomes.

B. Input-Output Measurement

- *Activities often neglected to collect data on both participant enrollment and completion, including a full definition of the standard for completion. Where data set is missing, it is difficult to judge the suitability of interventions for particular beneficiaries.*
- *Scaled activities (reaching many tens-of-thousands or even hundreds of thousands) often did not result in satisfactory completion rates, especially in conflict-affected environments where tracking is more challenging.*

Weak input/output measurement makes it difficult to evaluate the effectiveness of activities, especially those recruiting out-of-school youth in challenging environments. While not all evaluations reviewed were designed to answer the question of the overall outcome of the intervention, there is nonetheless a marked weakness in the collection and analysis of program enrollment data against data on program completion. This is particularly important for non-formal workforce training and livelihood activities with disadvantaged, out-of-school youth because dropout (attrition) can be very significant and costly. Without this data, USAID cannot determine whether activities are targeting the right beneficiaries for the planned intervention, or whether the intervention truly reflects the perceived needs and capacities of the targeted beneficiary.

Ten of the 12 WfD activities reported program enrollment or completion figures (Junior Achievement Armenia and Georgia EPI did not report monitor individual enrollment or completion figures). However, of these 10 activities, only 3 reported both. This is significant because without both enrollment and completion data (including a clear description of what completion means for that program), it is impossible to understand whether USAID activities are targeting individuals who can successfully complete, and presumably benefit from, the investment. Both data points are also needed to understand attrition and retention, significant factors in education and training activities. For the three activities, results are diverse: Proyecto META enrolled 35,000 youth and 8,140 youth completed the certificate training. Iraq Foras enrolled 175,000 youth in its online jobs platform and 16,500 were trained, though the training on use of the platform was very brief. The Advancing Youth Project enrolled 22,256, and 3,103 completed its training course. Although three data points are too few to draw firm conclusions, it appears that activities can reach large numbers of youth but that they fail to maintain anywhere near these numbers in terms of full completion of the program. This preliminary finding warrants further study.

The range of enrollments and completions were very divergent: from 105 students (48 percent male, 52 percent female) in Azerbaijan's Youth Business Leadership Project to 74,917 Dalit youth (78 percent male, 22 percent female) for the EIG Program in Nepal. Of the 10 programs reporting data, all but one disaggregated by gender. For more information, see Table 19 in Annex 4.

Two of the activities reported outcome results for additional education based on enrollment rates. Proyecto METAS had a rigorous approach, but a negative result, while Youth for the Future in Jordan

⁸⁶ Auten, Dabla, and Bassie, *Advancing Youth Project Performance Evaluation*, 46.

⁸⁷ Alcid, *A Randomized Control Trial of Akazi Kanoze Youth in Rural Rwanda*, 34.

provided an enrollment rate without any context or comparison. The results for additional education were disaggregated by gender.

C. Few Impact Evaluations

- *Too few impact evaluations are currently available to make reasonable judgments about the effectiveness of interventions.*

There are few impact evaluations available, especially for workforce development activities. Of 21 WfD activities for economic growth and for violence prevention and CVE, 7 had impact evaluations in 5 reports, of which only 4 had findings currently available.⁸⁸ WfD for violence prevention and CVE have more impact evaluations than WfD activities for economic growth, which had only two. Considering the large scale of many of the WfD activities, some involving tens of thousands of youth and activity budgets over \$20 million (for example, in Liberia, Afghanistan, Nepal), it is surprising that there are not more impact evaluations. **Without more impact evaluations, it will be difficult to understand whether activities are truly having an effect, and which interventions work best for which populations in which contexts.**

7. HE: University Extension Services

Guiding Questions

To what extent are USAID Higher Education activities supporting HEIs [higher education institutions] in developing countries to engage with their local communities, similar to “extension” type services offered by U.S. land grant institutions?

Results Summary

The review found that higher education extension services included efforts to contribute to agriculture by developing tools for farmers and to education by contributing to the development of education management information systems. There were also efforts to foster leadership and interest in agriculture among women, including outreach to families and communities to support their transition from the secondary school level.

- *Pilots to foster the extension of higher education expertise to solve country development challenges have been promising and creative, though there has been little focus on deepening their sustainability or scaling what has worked.*

Extension services expand the reach of universities to contribute to national development. At least four higher education activities were structured to provide relevant expert research and development for critical country objectives through extension services. Although activity monitoring did not always provide a clear picture of the results of services from the end-user perspective, this is a promising growth area for USAID high education programming.

In Georgia, a partnership between the Ministry of Education and Science and Georgia’s Ilia State University was supported to create an effective education management system and the country’s first master’s degree in education to train educators for careers in education management and administration. Activities included training 2,298 school principals and 280 Education Resource Center (ERC) officers in financial management and administration. Both efforts were aimed at improving autonomy and accountability. In general, principals found the financial management training useful and staff at the ERCs

⁸⁸ Impact evaluations are available for Rwanda Akazi Kanoze; Honduras METAS; East Africa CVE activities; Kenya Yes Youth Can; and Honduras A Ganar (later, findings available in near future). The CVE evaluation does not include workforce development features.

also received financial training to support the schools. Furthermore, survey results of principals indicated that the principals believed they could better forecast school financial needs, leading the evaluators to attribute declines in schools receiving additional funds to the principals' new budgeting skills. Finally, interactions between principals and ERC staff have increased, though both groups can now more easily resolve issues related to financial and academic reporting.⁸⁹

The Indonesia University Partnerships initiative included a collaboration between Columbia University and Institut Pertanian Bogor on climate change that had an extension component targeting 100 farmers with index insurance, a crop calendar, and fire risk management. Although after the farmer workshops, it was discovered that there was little interest from farmers in the index insurance, there was generally increased familiarity with concepts like crop calendars and crop failure. Ultimately, the activity monitoring did not capture whether these extension services improved the farmers' knowledge of the topics or led to better yields.⁹⁰

The Paraguay Women's Leadership Program used several modalities, with varying degrees of success, to reach out to farming families to promote young women's leadership in agriculture and entry into higher institutions of agricultural sciences. The activity reached secondary school students in communities outside of the National University of Asuncion's School of Agricultural Sciences. It provided training for young women in agriculture and decision-making. Participants in the training demonstrated leadership and enhanced self-esteem as well as knowledge of gender and human rights. Through an extension component, the activity also promoted workshops as well as internships in local communities, which targeted female participants and matched them with local female farmers. Through the National University of Asuncion, the activity sought to mentor female secondary school students and facilitate their enrollment in agriculture sciences; however, the mentorship component failed to produce tangible results. Nonetheless, the Probationary Admission Course for students from disadvantaged backgrounds yielded the enrollment of 14 students (9 women and 5 men).⁹¹

The Rwanda Women's Leadership Program deployed extension services in agriculture and education. The agriculture extension attempted to spread the expertise of the University of Rwanda's College of Agriculture, Animal Husbandry and Veterinary Medicine to local communities. In practice, only an advisory board of government and private sector representatives had been formed, leaving the ultimate outcome of the extension services unclear.⁹² In contrast, the education extension services from the University of Rwanda's College of Education strived to improve the access, retention, and completion of female students from underrepresented schools. The activity conducted adolescent health and reproductive health workshops in secondary schools. Moreover, presentations in local communities stressed the importance of girls' education and secured commitments from community leaders to encourage girls' education. Meanwhile, staff at the Kigali Institute of Education received training for a mentorship program with students, but the program failed to materialize after the training when it was deemed duplicative of other efforts at the university.⁹³

⁸⁹ Sedere, Shatirishvili, and Gorgadze, *Final Performance Evaluation of Georgia Education Management Project*, 3, 16-17, 22.

⁹⁰ Boardman et al., *Evaluation of the Indonesia University Partnerships Program*, 11-12.

⁹¹ Ramos-Mattoussi and Caballero, *External Evaluation of the Women's Leadership Program in Paraguay*, 36-39, 41.

⁹² Lancaster and Mirembe, *Evaluation of HED/USAID Women's Leadership Program-Rwanda*, 39.

⁹³ *Ibid*, 22-23.

8. HE: Entrepreneurship Promotion

Guiding Questions

To what degree are HE Business Schools promoting entrepreneurship and small business development across all the colleges within HE institutions? To what degree are HE institutions reaching out to TVET schools, community colleges and even secondary schools to teach and enhance entrepreneurial skills development?

Results Summary

The review found no evidence of business schools promoting entrepreneurship or small business development within their higher education institutions or beyond to TVET schools, community colleges, or secondary schools. However, some USAID activities included entrepreneurship training for students as part of other training.

- *Entrepreneurship development within higher education is a nascent area for USAID programming, though there are early signs that entrepreneurship support can be applied to both student preparation for the labor market and innovation in faculty research.*

Entrepreneurship promotion is included in some higher education activities as specialized training for students and, to a lesser degree, as a focus for faculty research. Specialized training in leadership and entrepreneurship were provided to higher education students in Lebanon and Paraguay to give them an edge in the labor market, and respectively, to better manage their own agro-business cooperatives. In Pakistan, one round of faculty research grants included entrepreneurship as a theme. While there appears to be some understanding that entrepreneurship skills are increasingly important in rapidly changing marketplaces, there is room for further development of this awareness and of specific interventions tailored to higher education programs.

9. HE: Access for Marginalized Youth

Guiding Questions

To what degree do activities reach marginalized communities and nontraditional learners? How do activities tailor their interventions to ensure learner success? What are aspects of activities contributed to increased learner success, including graduation rates? What kinds of financial assistance activities (including scholarships) have been provided and what have been the results of these?

Results Summary

The review found that although several of the reviewed HE activities focused on girls and women, only one scholarship program focused on disadvantaged young people.

- *Efforts to improve access for young women and marginalized populations to high-priority higher education programs have been mixed. Scholarship programs are critical for their financial support, but students also require a variety of other types of support to be successful, such as mentoring, counseling, and specialized trainings.*

Most higher education interventions of the reviewed evaluations focused on professionalization of faculty, as opposed to scholarships for marginalized youth. There were only two scholarship programs among the evaluations reviewed: the Pakistan Fulbright Student Program and the Lebanon University Scholarship Program. High-achieving but financially disadvantaged graduates of Lebanese public secondary schools received scholarships to study at Lebanese private universities. The evaluation reported some tension and challenges with the integration of the scholarship students into the mainstream campus, but these issues were eventually resolved. The latter explicitly targeted disadvantaged youth. The Women's Leadership Program in Paraguay strengthened the Center for

Leadership at the National University of Asuncion to develop innovative activities to promote gender equality including a focus on helping vulnerable populations access the means to attend the university.

10. HE: Policy and Systems Strengthening

Guiding Questions

To what degree do activities include systems strengthening, including establishment or strengthening of coordinating bodies; administrative or organizational capacity building; curricular and pedagogical improvement; policy reform; social messaging; advocacy; and partnership development? What are the major modalities for capacity-building of HE entities, including HICD (human and institutional capacity development) approaches? What have been the results of these efforts and how have they measured?

- *Higher education activities have been primarily designed as unique pilots and partnerships with U.S. institutions. Results are often positive; however, there is little evidence of follow-on investment to sustain and scale up reform across institutions and systems.*
- *Partnership with private sector industry councils and international businesses is important for lifting the quality of higher education training to international standards. The globalization of labor makes this effort an imperative.*

Systems strengthening for higher education activities focused on improving the quality of instruction and in promoting gender equality at all educational levels, as opposed to larger-scale institutional or policy reform. None of the evaluations reviewed included efforts at targeted policy reforms or creation or coordination of national-level advisory bodies. For ensuring the continuation of improved, active learning pedagogical skills acquired by faculty as the result of a USAID activity, the challenge is to shift from individual-level change to institution-wide change in the teaching culture. In Vietnam, HEEAP trained a small percentage of faculty at the five universities and three vocational colleges. However, change is needed at the other estimated 130 Vietnamese institutions teaching engineering. A solid strategy within and across institutions to disseminate the results of active teaching to those not trained by HEEAP was needed for a significant, sustainable difference in the teaching methods. Even at the individual level, faculty participants feared that the one-year stipend for a teaching assistant to relieve their workload to enable them to implement the new pedagogies would not be enough, as they would revert to the heavy workload the following year.

Results Summary

The review found indications that activities may be interested in expanding the reach, and achieving sustainability, of quality improvement efforts; however, there have been few efforts to assess sustainability.

11. HE: Responsiveness to Needs and Dynamics in Conflict-Affected Regions

Guiding Questions

Have activities been implemented differently in conflict-affected regions than in stable countries, such as inclusion of specific services (such as psycho-social support) and/or adaptive implementation approaches (such as the use of risk assessments)? Have evaluations of these activities gone beyond capturing traditional HE learning outcomes to include conflict-related results, such as reduction of risky behavior and perceptions of citizen security? How have conflict-related dynamics affected higher education results?

- *Higher education activities rarely, if ever, use Agency frameworks for programming in conflict-affected environments or for security-related outcomes. They do, however, function frequently in countries experiencing conflict.*

Higher education interventions are less focused on conflict prevention and stabilization.

Higher education activities did take place in countries experiencing violent extremism, such as Pakistan, Lebanon, and Kenya; however, the activity theories of change and designs did not focus on stabilization or violence prevention. Some activities, such as the Fulbright Program in Pakistan, reported promotion of goodwill and cooperation with the U.S., which could be a resource for peace building in foreign policy. A few higher education activities, such as the scholarship activity in Lebanon, focus on access to higher education for disadvantaged youth. Alleviation of feelings of marginalization among youth may have a stabilizing effect on countries prone to conflict, though this may be a longer-term, indirect effect.

Results Summary

The review found that while higher education activities are being implemented in conflict-affected environments, mitigating conflict-related issues was not a prominent component of these activities.

12. HE: Measurement Issues

Guiding Questions

What are the overall results in terms of beneficiary outcomes (e.g. placement in further education and research outputs)? What are the measures (indicators and assessments) used in evaluations to track outcome results?

- *There was no evidence of efforts to directly measure teaching quality in higher education activities; rather, student degree completion or employment rates were typically used as proxies.*
- *Satisfaction and perception surveys were typically used to measure the effectiveness of higher education partnerships with US universities. There is no evidence of the use of standardized objective measures and frameworks to capture results in HE institutional reform.*

Higher education activities are focusing on quality improvement, though few objective measures are used to assess progress over time. While a few higher education activities focus on

Results Summary

The review found that while higher education activities are focusing on quality improvement, few objective measures are used to assess progress over time.

supporting student access to and persistence in higher education programs, most USAID activities focus on bringing higher education institutions closer to international quality standards through faculty professional development, institutional capacity building, and research partnerships. They provide leadership opportunities for faculty to conduct and disseminate world-class research, train them in active learning pedagogy, and take measures to improve the educational environment for women to succeed.

Nonetheless, there was little or no evidence of efforts to objectively measure educational or research quality beyond beneficiary self-report or satisfaction surveys. Nor was there evidence of the use of labor market surveys to increase the knowledge within higher education institutions to better respond to country skills gaps. (One exception to the absence of labor market assessment is the Higher Engineering Education Program in Vietnam which arose from collaboration with Intel focusing on the skills gaps for potential hires, particularly around soft skills.) These activities focus on improving the leadership roles of higher education institutions to develop high-quality research and improve the quality of teaching based on international standards. Such activities were implemented in Indonesia, Vietnam, Pakistan, Rwanda, Georgia, and Paraguay.

In Rwanda, through partnerships with U.S. universities, Rwandan faculty participated in workshop-based training to develop gender-sensitive curriculum, pedagogy, and research agendas in the fields of education and agribusiness. A similar activity in Paraguay supported a Paraguayan university (National University of Asunción's School of Agricultural Sciences [UNA/FCA]) through a partnership with the University of Florida to revise its curriculum in agricultural sciences to support gender-sensitive leadership. As students received mentoring and leadership training, the university's Center for Leadership was strengthened.

Faculty development in Pakistan improved the quality and global dissemination of Pakistani research through research grants, exchange visits with U.S. universities, and training on new scientific equipment. Evaluation results were positive, suggesting that Pakistani researchers' critical soft skills and technical skills were improved, though further work is needed to incorporate the priorities of end-users of the research into the grant-making process. A similar finding is echoed in the Indonesia research partnership evaluation: that there should be more "documentation/dissemination of results for the local consumers of science and technology as well as "more work on 'technology transfer' to the field" through prototyping. In Georgia, a partnership between the Ministry of Education and Science and Georgia's Ilia State University was supported to create an effective education management system and the country's first master's degree in education to train educators for careers in education management and administration. The program is being expanded to a second Georgia university.

Institutional capacity indicators measure a multitude of results, none of which are the same across higher education activities, though the results are usually positive. Similar to WfD activities, higher education activities usually measure partnerships by counting the number of partnerships and considering any partnership as positive. Some higher education activities developed unique partnership indicators such as secondary school outreach, the enrollment of young women in an agriculture extension program, and the management of an agriculture public-private partnership.

One measure of activity quality is the preparedness of higher education graduates for employment, though data is sparse for this analysis. Lebanon's University Scholarship Program indicated that 22 percent of graduates found employment within three months after graduation. The graduates claimed they felt prepared to enter the workforce, though 90 percent intend to pursue graduate degrees, roughly half within six months.⁹⁴ None of the higher education activity evaluations reviewed provided outcome results for self-employment.

⁹⁴ Miller et al., *USAID Lebanon University Scholarship Program Mid-Term Evaluation*, 24.

Two examples of improving instructional quality through private sector engagement were found. Under the Indonesia University Partnerships, the University of Southern California (USC) formed a partnership with the Institut Teknologi Bandung (ITB) to strengthen a master’s program in geothermal energy, building on a previous partnership between Star Energy and ITB. ITB sought to provide its graduate students with higher quality education relevant to the geothermal sector, while Star Energy wanted better trained professionals. Results included a doubling the number of students enrolled in the geothermal degree, from 17 to 40, and with a greater share of women.

In the context of Higher Engineering Education Alliance Program (HEEAP), Intel engaged in long-term effort to shift engineering education in Vietnam from theoretical to practical, thereby producing graduates with skills ready to work in the sector. Intel designated the higher education institutions in Vietnam from which faculty in mechanical engineering, electrical engineering, and electronics could apply to attend six-week training courses on active learning and curriculum design at Arizona State University (ASU). Surveys of Vietnamese faculty that had attended the training at ASU showed that the faculty were applying the active learning methods in their courses. However, the evaluation acknowledged that larger institutional change across engineering education in Vietnam is unlikely given the small number of faculty trained and the missing follow-up support required to maintain the application of active learning methods. Nonetheless, Intel continues to support HEEAP as a component of its effort to shift the instructional emphasis of engineering in Vietnam.⁹⁵

Research outputs were the most common higher education outcome for five of the activities. The most common indicator was published research, though the Indonesia University Partnerships tracked the number of dissertations and theses, the HEEAP in Vietnam monitored faculty activities, and the US-Pakistan Science and Technology Cooperation collected data on new products and academic presentations. The Fulbright Student Program’s published research indicator was positive, comparing the number of graduates with academic publications before and after completing their Fulbright. In addition, the faculty activities in Vietnam had a negative result because most of the faculty failed to undertake or complete one despite its being a program requirement.

13. Crosscutting Themes

Research Themes	Results Summary
<ul style="list-style-type: none"> • Gender • Disability • ICT • Innovative finance • Scaling and sustainability 	<p>The review found that less than half of the activities took the minimal step to collect gender-specific data. Most activities that accomplished significant results for girls or women were not designed with gender as a consideration, but a select few incorporated gender-specific components or mainstreamed the concept throughout the design. Only one activity considered disability, and developed a curriculum for the deaf. While a few activities had private sector cost-share, none of the activities had innovative financing mechanisms. Information and communications technology (ICT) is most often employed for instructional purposes, though activities have created online job platforms, education management information systems, a website for professional networking, and agriculture analysis tools that advanced academic research.</p>

⁹⁵ Gilboy et al., *Mid-Term Evaluation of the Higher Engineering Education Alliance Program*, vii.

A. Gender⁹⁶

- *Gender is not a prominent organizing principle in the vast majority of the Goal 2 activities included in this synthesis.*
- *Some activities noted significant gains for young women in labor markets that are not usually very permissive for women. With a few exceptions, this may be due to women's tenacity in participating in the activity more than intentional, gender-sensitive design features.*

Overall, there is a lack of attention to the complex dynamics of gender in Goal 2 activities.

Gender-sensitive programming assumes that gender is a dynamic process and that both male and female perspectives and experiences need to be considered. Although there was attention to leveling the playing field for females, there was no indication that either activities or evaluations have considered male experiences from a gender perspective or opportunities to engage men in unique ways in the pursuit of gender equality.

For workforce development, there have been some impressive gains for women in specific activities, though little strategic focus on gender. Most WfD activities did not highlight gender-based targets or interventions; however, 10 of the 12 activities disaggregated data by gender. Four WfD for violence prevention and CVE activities disaggregated by gender, and it was unclear whether an additional four did, though their experimental evaluations did include gender in their sampling strategies. Still, six activities achieved significant results for girls or women without a theory of change or results framework focusing on gender.

- Djibouti AIDE placed 240 participants, of whom 64 percent were women, in jobs after receiving vocational training.
- Georgia EPI made a concerted effort to train women in vocational skills for the garment industry. While the activity did not track individual outcomes, anecdotal evidence suggests many women are largely employed and there is high demand for skilled garment workers.⁹⁷
- Nepal EIG had a dominant share of female participants for literacy training (95 percent) and agriculture and enterprise training (81 percent). Female-headed households had higher incomes than male-headed households, with the gap growing from roughly Rupees 10,000 in 2010 to more than Rupees 30,000 in 2012. Women who had completed the literacy training and agriculture and enterprise training contributed the most to increasing the household income gap.⁹⁸
- The Akazi Kanoze initiative in Rwanda found that its workforce readiness training benefitted women more than men; women had statistically higher gains in knowledge related to applying for jobs and searching for jobs relative to men. The gains closed a gender gap in knowledge between men and women.

Few activities intentionally set gender-based targets or training priorities. An exception is Afghanistan AWDP, which achieved its target of 25 percent women as TVET beneficiaries. To enable that to happen,

⁹⁶ Guiding questions included: Are evaluations adequately examining gender dynamics? To what degree do evaluations find differential results for male versus female beneficiaries, both in terms of activity access and outcomes? What are we learning from evaluations about the relationships between gender and livelihood and advancement opportunities? What contextual and implementation factors help explain gendered outcomes?

⁹⁷ Menendez England and Associates, *Mid-Term Performance Evaluation of the Georgia Economic Prosperity Initiative*, 39.

⁹⁸ Population, Health, and Development Group, *Final Evaluation Report Education for Income Generation*, 7, 10-11.

it required training institutions to hire more female instructors. Female job seekers were more likely to find a job than men, though women already in jobs were less likely to receive raises than men.

The evaluation of the East Africa CVE activities found lack of attention to gender to be a salient gap, and suggested that future activities should include a gender assessment. It noted that “in patriarchal societies [such as Somalia], women and girls can have a pacifistic influence on family members should they choose to exercise it.” The evaluation also found that young women are generally less engaged in the communities than young men. This was attributed to traditional Somali social structures that exclude women from decision-making. The study recommended that CVE interventions focus more intentionally on women since CVE activities tend to emphasize the role of young men in extremism, without analyzing underlying gender dynamics.

The Yes Youth Can evaluation found that overall impacts were broadly similar for women and men, but with some notable differences. Both economic outcomes were weaker for women, suggesting that the limited economic impacts of Yes Youth Can tended to reach male participants to a greater extent than female participants.

For higher education, activities reveal mixed attention to gender. Only five higher education activities disaggregated data by gender and/or had a substantial focus on gender. Inclusion of women as students, faculty and administrators in higher education systems is critical for a wide array of development outcomes, including more inclusive workforces and research and technology transfer that improves gender equality in all development sectors, from access to water to climate change. Therefore, it may be surprising that only four of nine higher education activities disaggregated results by gender. Nonetheless, two activities had extensive gender-focused objectives.

- The Women’s Leadership Program in Paraguay delivered workshops on leadership to girls and awarded scholarships for young women to study agriculture. However, the scholarship program at the National University of Asuncion only led to nine women studying agriculture.
- The Women’s Leadership Program in Rwanda worked extensively to design a gender-sensitive curriculum and implement gender-sensitive policies at the University of Rwanda, College of Education. The program also led outreach in local communities on girls’ education and attempted to augment the access of female students through e-learning platforms. Efforts to enhance the gender research capacity of faculty produced six research projects in progress at the time of the evaluation.
- The Vietnam Higher Engineering Education Alliance increased female participation in its faculty cohorts from 17 percent in the first year, to 33 percent and 30 percent in the second and third years. The program hailed the increased number of women as significant because women compose less than 30 percent of university-level engineering students in Vietnam.
- The Lebanon University Scholarship Program had nearly 500 graduates, of whom 60 percent were women. This share of female graduates mirrors the gender distribution in Lebanese public secondary schools. However, focus group discussions with female students revealed that they felt more empowered and gained a greater sense of equality due to their stipend. For these young women, the stipend was the first time in their lives they had personal money to spend at their own discretion.
- The Teacher Education and Professional Development initiative in Kenya designed a gender-sensitive curriculum with manuals for primary schools. Unfortunately, the Ministry of Education, Science, and Technology (MoEST) was slow to approve the curriculum and supporting materials.

B. Disability⁹⁹

- *There is little evidence for the inclusion of disabled learners in Goal 2 programming.*

Overall, there is little evidence for the inclusion of disabled learners in Goal 2 programming. Only one evaluation referenced disability. Empowering Jamaica's Youth revised its curriculum so that deaf students could study in Jamaican Sign Language, and provided a one-week training on personal development for students with disabilities which led to apprenticeships for the top performers.¹⁰⁰

C. Information and Communications Technology¹⁰¹

- *The use of ICT in Goal 2 programming was diverse, ranging from instructional support to labor market intermediation to faculty exchange.*
- *Technical difficulties hindered the full utilization of the potential of the ICT-based interventions.*

Overall, there were significant and, at times, impressive efforts to utilize information and communications technology, though results often fell short. Nine activities substantively deployed information and communications technology (ICT) for instructional purposes, employment services, education management, professional networking, and agricultural research. However, there were a wide range of technical and design difficulties in getting the systems to work as intended. While few would argue with the need to continue to mainstream ICT solutions into Goal 2 programming, more attention is needed both to design issues around the usefulness of the selected technologies and equally to measuring the impact of these investments.

- In Indonesia, faculty from the University of Southern California (USC) designed an introductory geothermal course for delivery over the internet to students at Institut Teknologi Bandung. Unfortunately, USC could not overcome technical issues with communication and voice quality.
- Iraq Foras contained an e-learning platform for online courses, but users had trouble accessing them because the vast majority of courses required English and slow internet speeds made it difficult to use the platform. Only 173 people completed at least one course out of nearly 3,000 that viewed course materials. While Foras registered an impressive 175,000 job seekers for the portal and 3,000 firms, only 20 percent of users actively used it. Furthermore, the activity did not track individual employment outcomes, so it is impossible to determine the platform's impact.
- Similarly, the Somalia Youth Livelihoods Program (SYLP) designed InfoMatch, a job platform primarily for mobile phones. Roughly 10,000 Somali youth registered for InfoMatch through which they uploaded resumes and applied for jobs. However, SYLP did not monitor individual employment outcomes for InfoMatch users, so the study team cannot ascertain the added value of this component.
- The Teacher Education and Professional Development initiative in Kenya trained teachers on integrating technology into their classroom instruction using pre-service and in-service training. Teachers who received in-service training reported that their students had become more motivated and performed better in school.

⁹⁹ Guiding questions included: To what degree do activities address disabled populations and their needs? How is disability treated in evaluations, included in the evaluation questions? Included as a sampling stratum? Is there a separate data collection module? Discussed separately in findings/conclusions/recommendations?

¹⁰⁰ Social Impact, Assessment of the Caribbean Basin Security Initiative, 26.

¹⁰¹ Guiding questions included: To what degree are activities utilizing technology for instructional purposes? What are other uses of technology have been utilized and what have been the results?

- The Vietnam Higher Engineering Education Alliance developed a website through which Vietnamese faculty who attended training at Arizona State University (ASU) could communicate with each other and their mentors at ASU. The evaluation determined that the website existed, but evaluators could not access its content, and a survey of Vietnamese faculty revealed that they had used the website infrequently.¹⁰²

D. Innovative Finance¹⁰³

- *There is no evidence of the use of innovative finance in the reviewed evaluations.*

There is no evidence of experimentation with innovative financing, defined as social impact bonds, development impact bonds, and results-based financing, as well as special financing schemes to support students such as student loans or entrepreneurs such as venture capital. Workforce development and higher education are critical elements of a national education system, but for a variety of reasons tend to be more expensive than basic education and more likely to be privately funded by families, learners, and, sometimes, the private sector. For this reason, innovative financing could be used to expand the reach of workforce and higher education systems. Although evaluated activities boasted a wide variety of private-sector partnerships and some in-kind cost-sharing (which is now mainstream for WfD), none of the evaluations reviewed revealed programmatic efforts to test innovative finance.

E. Sustainability¹⁰⁴

- *The evaluations reviewed reveal an interest by USAID in achieving sustainability by expanding the reach of services and quality improvement achieved through USAID funding, however, there have been few systematic efforts to assess sustainability.*
- *Performance evaluations are weak instruments for evaluating sustainability. Sustainability studies would need to be structured to review the country operating environment quite a bit after USAID funding has ended, but most performance evaluations are conducted immediately upon activity completion before sustainability efforts would have had time to play out.*

Overall, evaluations reviewed reveal an interest in expanding the reach and achieving the sustainability of quality improvement, though there have been few efforts to assess sustainability. For workforce development activities, many evaluations noted the importance of private sector partnership as critical to sustainability with demand-driven training aligned with employees' skills needs having the potential to attract cost sharing from businesses. For scholarship programs, there is little expectation of sustainability without USAID funding. However, for other types of higher education activities, evaluations reveal a consistent concern in USAID for the sustainability of institutional capacity-building. Still, while evaluations have enumerated the discrete achievements of higher education partnerships, there has been little effort to measure or evaluate systems-level impact and sustainability of reforms. For example, while a small number of faculty or institutions may benefit directly from USAID

¹⁰² Gilboy et al., *Mid-Term Evaluation of the Higher Engineering Education Alliance Program*, 6, 14-16.

¹⁰³ Guiding questions included: To what degree have the use of private versus public monies been included as part of project design? What have been any innovative financing schemes, if any, and what have the results been?

¹⁰⁴ Guiding questions included: To what degree has scaling of activities been part of activity objectives? Where scaling is an objective, how have design features contributed? What have been the results? Do evaluations of projects that are promising but have not been scaled make recommendations concerning scaling? What patterns exist if any in these recommendations on future scaling? To what extent do the projects focus on sustainability? How do they define it and measure it? What are the findings and recommendations relating to sustainability?

support to upgrade student services, modernize pedagogy, or develop research agendas, it is unclear how these reforms would be spread across tertiary institutions and systems.

- Georgia’s EPI subsidized both students and manufacturers by supporting the training, though most students are required to pay part of their tuition.
- Junior Achievement Armenia engaged with businesses to mentor youth in entrepreneurship activities, and sustainability was built into the model from the beginning by keeping overhead and costs down.
- Several evaluations, such as those of research partnerships in Indonesia and Pakistan, recommended improved sustainability through more inclusive research processes and a call for more technology transfer from the university to the field through prototyping.
- The evaluation of the Lebanon scholarship program revealed that program managers believed that intervention components, such as workshops, community involvement, leadership training, and capstone projects, had substantial value for all students, and could strengthen the universities’ overall educational programs. Apart from the scholarships themselves, university staff recognized and appreciated the impact of program experiences on students: that they had become better students, stronger leaders, and more employable.

Sustainability was addressed by most, but not all, of the evaluations reviewed, and evaluators made efforts to put activity sustainability into context. For example, some activities operate in low-resource environments and are focused on populations without strong champions (such as at-risk youth or adult learners)—factors which impose limitations on scale and sustainability. Also, not all activity components are intended to be sustainable but have demonstration effects or humanitarian or short-term stabilization objectives. Furthermore, sustainability includes multiple dimensions including financial, organizational and technical—not all of which need to be in play for an activity to have positive effects over the longer term.

For **workforce development** activities, many evaluations reviewed noted the importance of private sector partnership as critical to sustainability (i.e., WfD activities reviewed focused squarely on partnering and responding to the needs of the private sector) Effective demand-driven training that is truly meeting employees’ skills needs is more likely to attract cost sharing from businesses, though it is not clear that activities are structured to facilitate this cost share. Georgia’s EPI “subsidized both students and manufacturers by supporting the training, though most students are required to pay part of their tuition. Whether the companies would pay for such courses themselves remains an open question but there are indications from the manufacturers that if the trained workers demonstrate increased productivity and the price for training is reasonable, they might consider paying for it.”¹⁰⁵

Even for activities that work with younger, in-school youth who may be several years from labor market entry, the private sector is a vital partner. Junior Achievement Armenia successfully engaged with businesses to mentor youth in entrepreneurship activities. Scalability and sustainability was built into the Junior Achievement Armenia model from the beginning by keeping overhead and costs down; this is the only evaluation that provided detailed costing data and analysis, as well as a global analysis that compared Junior Achievement Armenia costs with U.S. Junior Achievement chapters.¹⁰⁶

More difficult sustainability issues exist with activities serving out-of-school or at-risk youth populations. For example, the sustainability of the Advancing Youth Project in Liberia is in doubt due to shifting priorities within the Ministry of Education from access to quality. This may involve a greater focus on higher quality education for school students than on access to functional literacy for adult learners. It was also noted that the lack of a tracer study of the impact of adult basic education on livelihoods is a barrier

¹⁰⁵ Sedere, Shatirishvili, and Gorgadze, *Final Performance Evaluation of Georgia Education Management Project*, 40.

¹⁰⁶ Turner et al., *Mid-Term Performance Evaluation for Junior Achievement of Armenia*, 27-28.

to institutionalizing the program. On the other hand, as the result of the activity, the capacity of NGO partners was strengthened to use the Advancing Youth Project materials, and there are plans by other donors and international NGOs to expand the use of the curriculum and active learning pedagogies with both the adult learners and younger youth cohorts in other regions of Liberia.

For **higher education** activities, the evaluations reviewed suggest that in scholarship programs, such as Lebanon's, there is no expectation that the scholarships themselves are sustainable without USAID funding. However, program managers believed that components of the activity, such as workshops, community involvement, leadership training, and capstone projects, had substantial value for all students, and could strengthen the universities' overall educational programs. University staff recognized and appreciated the impact of the experiences on students, noting that they had become better students, with strong leadership skills, and thus, more employable.

For higher education research partnerships, the concern was about how to sustain the quality of the research for the "end-user," mostly public agencies (such as the national basic education system), international research networks, or the community members and businesses. The Indonesia university research partnerships evaluation recommended more interdisciplinary research and program planning, improved linkages with the private sector, and continued linkages with internationally-recognized research institutions. It was also noted that a minimum of three years of funding or more is needed for sustainable action research (more than was provided through the USAID activity), and that sustainability would require more inclusive participation, including, for example, in mental health research, and in partnerships among hospitals, subdistrict governments, and the communities.

OBJECTIVE 3: SYNTHESIS OF FINDINGS AND LESSONS LEARNED RELATED TO USAID EDUCATION STRATEGY GOAL 3 (EICC)

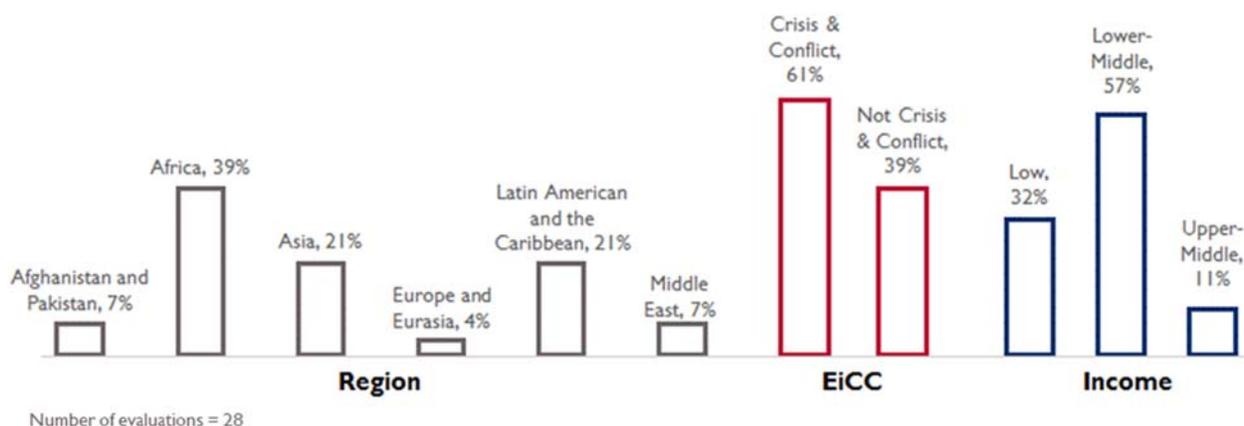
Goal 3 (EiCC) programming addresses the high-level Education Strategy goal to “increase equitable access to education in crisis and conflict environments for 15 million learners.” For this study, the Office of Education was interested in findings and lessons learned related to access to education from all evaluations that qualified for the syntheses phase. This review synthesizes the results for 28 evaluations covering 18 activities related to Goal 3. Overall, this constitutes a medium-sized body of evidence.

Evaluation Sample, Context and Limitations

All Goal 3 related evaluations reviewed in this study were funded by USAID and met the Office of Education criteria for evaluation quality. These included performance and impact evaluations published between 2013 and 2016. The Office of Education also requested the inclusion of a few research studies that did not evaluate a specific intervention.

As shown in Figure 13 below, there are evaluations from every region, but the largest share (39 percent) comes from Africa. Unlike Goals 1 and 2, which are dominated by evaluations from East Africa, the sub-Saharan African countries’ evaluations under Goal 3 are almost entirely from West Africa and Central Africa. The second most predominant region represented is Latin America and the Caribbean, primarily due to the number of evaluations from countries in Central America’s Northern Triangle affected by generalized gang violence. Given the hotspots of conflict and global instability, it is striking that Goal 3 does not have more evaluations from the Middle East and North Africa, regions that have experienced an incredible amount of turmoil and conflict since the Arab Spring in 2010.

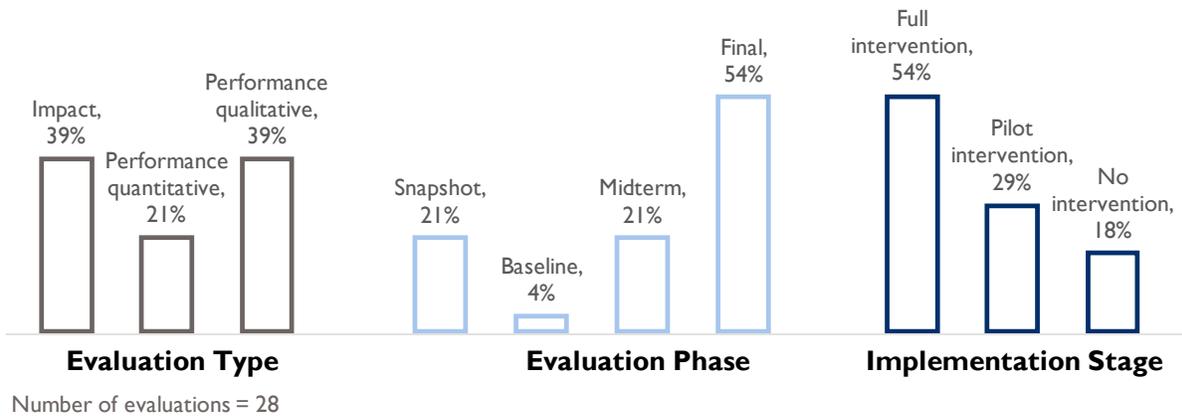
FIGURE 13: PERCENTAGE OF EVALUATIONS REVIEWED UNDER GOAL 3 BY REGION, CRISIS AND CONFLICT STATUS, AND COUNTRY INCOME



The body of the evidence for the Goal 3 review is well distributed across evaluation types; however, the substantial number of impact evaluations reviewed is partly due to the four evaluations done for the School Dropout Prevention Pilot Program (SDPP) in four countries. It should also be noted that only

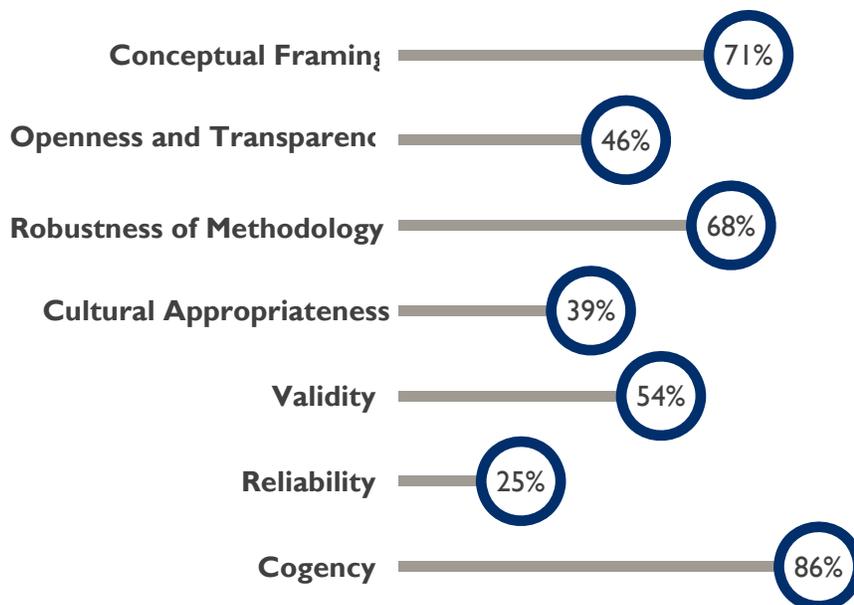
about half of the evaluations reviewed are final evaluations (54 percent) and only half evaluated full interventions (54 percent).

FIGURE 14: PERCENTAGE OF EVALUATIONS REVIEWED UNDER GOAL 3 BY EVALUATION TYPE AND EVALUATION/IMPLEMENTATION PHASE



Based on the expert judgments of the seven principles of quality, cogency (86 percent) was rated as “adequate” most frequently, while reliability (25 percent) was rated as “adequate” the least. Results on Figure 15 indicates that despite the Office of Education’s minimum quality standards for inclusion in the synthesis, the general quality of the body of evidence is only moderate.

FIGURE 15: PERCENTAGE OF EVALUATIONS REVIEWED UNDER GOAL 3 BY PRINCIPLES OF QUALITY



The findings in this report should not be considered generalizable across all Goal 3 activities for several reasons:

- This report focuses on a distinct set of priorities for USAID's Goal 3 team, such as CVE and violence prevention, safer learning environments, and social and emotional learning (SEL). These priorities are not situated within a global conceptual framework, though USAID's Goal 3 team has indicated this is under development.
- Further, as documented in the evaluation reports reviewed, each intervention was contextualized to the conflict setting in which it was implemented. The unique nature of each intervention made it difficult to establish trends in findings across the evidence base.
- The evaluations reviewed defined and measured outcomes in unique ways, which also made it difficult to establish consistency across studies.
- While there is a wealth of high-quality qualitative and mixed-methods assessments and studies of Goal 3 interventions, the evaluations reviewed were not able to disentangle the value added of specific components in many instances. This makes it difficult to assess the benefits or effectiveness of specific activities within complex interventions, further limiting the generalizability of findings in this report.
- Finally, the Goal 3 evidence base lacks rigorous experimental data such as randomized controlled trials (RCTs). While RCTs are costly, and often there are ethical considerations with implementing such a design in a conflict setting, the absence of larger quantitative evaluations limits the generalizability of findings in this report.

In 2016, the Education in Crisis and Conflict Network (ECCN) conducted an analysis of the indicators used in performance monitoring and evaluation plans from 25 USAID Education activities implemented in 16 countries between 2007 and 2018.¹⁰⁷ This analysis revealed a lengthy list of indicators used to measure several of the priority topics in this report: access to education (47 indicators), equity (21 indicators), school safety (31 indicators), and education quality (118 indicators).

Most indicators focused on outputs, without clear linkages to outcomes. Indicators related to retention, equity, school safety, and conflict-sensitive education are not standardized, making comparisons across activities difficult. The analysis also revealed that 103 indicators (28 percent of all indicators) covered categories not addressed by interventions' theories of change.

Clear recommendations have emerged from this analysis that align with the recommendations emerging from this synthesis. Specifically:

- USAID-funded activities should be encouraged to use a balance of standardized indicators, developed from commonly agreed upon definitions of key topics, and customized indicators.
- These indicators should align with activity results frameworks and theories of change.
- Activities should also ensure that theories of change are explicit about the unique value of specific activities toward individual outcomes so that such theories can be evaluated.

Intervention Modalities

Goal 3 programing focused on four broad programmatic types: improved access to basic education through strengthening institutional, environmental, and material capacity; violence prevention and CVE activities with youth empowerment and WfD components; situational assessment of the ways in which equitable access to quality education is achieved in emergency/conflict settings; and primary and

¹⁰⁷ ECCN, *Analysis of Indicators Used in USAID Education Activities in Crisis and Conflict Environments*.

secondary drop-out prevention programming. The range and frequency of these programmatic types are summarized in the infographic below.

FIGURE 16: GOAL 3 INTERVENTION MODALITIES



Findings

Given the variable nature of the findings, the tendency of each activity to be contextualized to the conflict context where its beneficiaries are situated, and the focus of the evaluations on unique sets of outcomes, it was challenging to describe results in broad strokes. Therefore, results are presented using a case study approach throughout much of the Goal 3 section of the report.

Synthesis results parallel extant evidence of education activities in crisis and conflict contexts. Reports such as [What Works to Promote Children’s Education Access, Quality of Learning and Wellbeing in Crisis](#)¹⁰⁸ and [Education in Emergencies and Protracted Crises: Toward a Strengthened Response](#)¹⁰⁹ were used to cross-reference findings from the current synthesis with existing reviews of education activities in crisis and conflict contexts. Consistencies were found with respect to equity, SEL, and gender. No relevant findings were found with respect to safer learning environments, improved programming, strengthened institutional capacity building, and activities targeting violence prevention or CVE.

¹⁰⁸ Burde et al., *What Works to Promote Children’s Educational Access, Quality of Learning, and Wellbeing in Crisis-Affected Contexts*.

¹⁰⁹ Nicolai, Hine, and Wales, *Education in emergencies and protracted crises: Toward a strengthened response*.

I. Strengthened Institutional Capacity to Provide Services

Guiding Questions

To what degree do EiCC activities strengthen local institutional capacity? What strategies are employed to strengthen capacities of local institutions in EiCC activities and how effective are they? Are there any notable innovations in measuring results in institutional capacity strengthening?

Results Summary

The review found that about a third of the reviewed evaluations mentioned improved access to basic education through strengthened institutional and/or environmental and material capacity. However, these evaluations also represented the greatest variation of targeted beneficiaries and methods, which makes it unclear which are the most effective approaches.

- *Strengthened institutional capacity at local and national levels led to improved access to basic education.*
- *Education activities employ a variety of models to improve institutional capacity. In some instances, the reviewed evaluation reports did not contain enough information to determine the effectiveness of particular education activities.*

According to the evaluations reviewed for the synthesis, strengthening institutional and environmental/material capacities was one of the most common approaches to ensuring access to basic education (10 of 28). Skills training to improve teachers' pedagogical approaches and education officials' school management abilities ranked highest among the range of approaches used to strengthen institutional capacity. Public-private partnerships were the least often cited and evaluated forms of institutional capacity building.

These activities also represented the greatest variation in beneficiaries and their corresponding methods to improve institutional capacity building (6 of 28), including:

- Community-level/NGOs;
- National-level ministerial officials;
- Different line ministries such as health, education, or social welfare;
- Teachers and school administrators; and
- Parents.

Table 23 in Annex 5 presents a list of all activities by target type to demonstrate the variety that exists across Goal 3 activities, also suggesting that, given the myriad contexts in which USAID operates, there is no one-size-fits-all approach to strengthening institutional capacities.

Evaluation results suggested that the most effective methodologies to enhance institutional capacity were:

- Skills training to improve pedagogical approaches;
- Classroom management techniques;
- School management capacity;
- Training workshops (continuous, intensive); and
- Curricular enhancements that respond to the needs, educational and vocational interests, and opportunities for learners.

These activities were most effective when they reflected the political, security, and cultural contexts in which learners live.

Interestingly, this synthesis found that public-private partnerships (i.e., Nigeria NEI and Senegal EBD) and civil society engagement were infrequently employed and studied (i.e., Nigeria NEI, Afghanistan ALSE, and Ghana Public Works Construction¹¹⁰). Yet the reviewed evaluation reports often referenced the strong potential of such partnerships to increase access to education for learners in conflict and crisis settings.

As discussed later in this report, both approaches led to greater efficiency and effectiveness in activity implementation and management. Public-private partnerships and civil society engagement should feature prominently in USAID's strategic approach to increasing equitable access to education in crisis and conflict environments for learners.

Strengthening institutional capacity is one of the greatest challenges USAID faces to fulfilling its development goals globally. Training, however, is only one piece of the organizational performance puzzle. USAID's Behavioral Engineering Model illustrates that, on an individual level, knowledge and skills, capacity, and motives need to be aligned with the work required. On an environmental level, information, resources and tools, and incentives need to be clearly present for key organizational capacity to improve.¹¹¹

The 10 activities that sought to improve institutional capacity also had the greatest variations in beneficiaries and methods to improve capacity to deliver equitable access to quality education.

A. Types of Capacity Building

- *The most effective methods were skills training to improve pedagogical approaches and school management. Least explored in the reviewed evaluations were public-private partnerships.*

The evaluations reviewed under Goal 3 that address capacity building are categorized into five specific types (partially aligned with USAID's Human and Institutional Capacity Framework).¹¹² The first, "**improved teaching methods**," are implemented to improve student instructional outcomes and include such activities as the Nigeria RARA intervention. Improved teaching methods are identified as a key approach in USAID's Behavior Engineering Model (i.e., individual knowledge and skills). This activity implemented an innovative instructional approach to improve Hausa early reading outcomes among primary 2 pupils in the Northern Nigerian states of Bauchi and Sokoto.

The second category is "**improved school resources**," illustrated by the Afghanistan Basic Education, Literacy, and Technical-Vocational Education and Training (BELT) activity. This activity addressed shortages of learning materials through emergency procurement of textbooks. USAID's framework to enhance individual and organizational change also specifies that making available necessary resources and tools within the environment is a critical ingredient to foster capacity strengthening.

The third category is "**improved learning environment**," illustrated by the majority of interventions under the broader capacity-building category. Environmental improvements are also specified in USAID's Behavior Engineering Model to strengthening individual and organizational capacities. An example is the

¹¹⁰ While Ghana is not categorized as a crisis or conflict-affected setting, the objectives of the project are toward increasing access to education through infrastructural improvements, priorities of the Goal 3 portfolio. Therefore, the Ghana evaluation was included in the Goal 3 synthesis.

¹¹¹ *Human and Institutional Capacity Development Handbook*, 6.

¹¹² This framework (https://pdf.usaid.gov/pdf_docs/Pnadt442.pdf) specifies that training is only one necessary ingredient to enhancing individual or organizational performance. According to the Updated Behavior Engineering Model (page 6), individuals need knowledge and skills, capacity, and motives, while at the environmental level, information, resources and tools, and incentives are required to engineer behavioral change.

Democratic Republic of Congo (DRC) Opportunities for Equitable Access to Quality Basic Education (OPEQ) activity that focused on improving SEL environments for boys and girls in conflict-affected areas.

The fourth subcategory is “**improved local governance and leadership in school management**,” also a repeated theme across a majority of the reviewed evaluations that fall under the category of capacity building. This is well-aligned with a number of individual-level factors specified in USAID’s Behavior Engineering Model—knowledge and skills, capacity, and motives. A notable example is the Jordan Learning Environment Technical Support (LETS) activity, which sought to assess the implementation of activities focused on improving the learning environment, such as results-based M&E systems and tools to build a range of capacities within the school/education system. The enhanced capacities aligned well with the strategic priorities of USAID’s Goal 3 team—specifically, increasing access to education involves a multilevel approach to capacity strengthening and skills enhancement (classroom, school, and policy environment).

For example, LETS activities included the “joint use of a) Individual School Profiles (ISP), and b) Results-Based Benchmarking (RBB). Together, these two instruments [composed] the learning environment improvement (LEI) assessment tool and were used periodically to guide adjustments to interventions.”¹¹³ LETS coaches also assisted schools to identify school and community-based activities through these ISPs. Capacity strengthening also took place at a national level through the Jordan LETS activity—specifically, the activity focused on enhancing the MOE’s capacity to sustain and institutionalize learning environment improvements and to prepare ASK for Capacity Development, a local implementing partner, to compete directly for funded work.

Nearly all activities assessed in the reviewed evaluations sought to improve the quality of learning environments in which learners were situated and to improve local governance or school leadership and management. The second most common objective was improved teaching/student outcomes. Physical infrastructure improvements (i.e., Ghana Public Works Construction and Yemen Community Livelihoods Project [CLP]) and improvements to school resources (Afghanistan BELT, Senegal EBD, and Yemen CLP) were the focus of a smaller number of education activities.

B. Beneficiaries of Capacity Building

- *Nearly all interventions reviewed by this evaluation synthesis targeted capacities of school administrators and management teams, community-level officials or NGOs and ministry-level officials. Line ministries and parents were the least targeted beneficiary groups for capacity-building activities.*

Beneficiaries of capacity building are part of each activity’s unique strategy or intervention approach to tackle challenging education issues in different contexts. The beneficiary groups of Goal 3 activities (included in this synthesis) ranged from community-level stakeholders to ministerial officials. The objectives of capacity building for each of these groups differed.

Nearly all interventions reviewed by this synthesis targeted school administrators and management teams in capacity-building activities, except for Yemen CLP. Nearly all interventions also focused on building the capacity of community-level officials or NGOs to implement education-focused interventions in conflict and crisis contexts, except for DRC-OPEQ. The next most frequently targeted beneficiary group for capacity-building activities was ministry-level officials. Interestingly, line ministries and parents were the least targeted beneficiary groups for capacity-building activities. This finding aligns well with the EiCC priority to strengthen institutional capacity—but also illustrates the hurdles that exist (e.g., developing

¹¹³ Moubayed et al., *Final Performance Evaluation USAID/Jordan Learning Environment Technical Support Program*, 5.

necessary cross-sectoral partnerships at a national level and strengthening parental capacity). There may be an assumption that parental capacities are targeted by community NGOs; this should be tested through future evaluative work.

Community-Level Capacity Building. The Afghanistan ALSE activity demonstrates one effective strategy for community-level capacity building through community-based education (CBE). Specifically, the activity focused on teacher recruitment, community capacity building, and mobilization efforts aimed at maximizing school learning and access. Overall, implementing CBE in remote villages increased education access and learning significantly. Additionally, it significantly improved children’s performance in school. Adults living in communities with CBE activities showed a substantial increase in trust and in the perceived legitimacy of education service providers, which, according to the evaluation, created a ripple effect of increased trust in overall government legitimacy. In terms of teacher recruitment, there were no real negative effects from hiring more qualified teachers who were not from the beneficiary community or who may have been unfamiliar with that community. In terms of the parent and community mobilization component, implementing an enhanced package did not create any statistically significant effect on access or learning outcomes. This may be because the enhanced package did not adequately address the reasons why some parents persistently do not send their children to school, or the activities were not implemented at the appropriate “dosage” to be effective.

In light of the subpar findings related to parental support, USAID should continue to explore creative and low-cost ways to engage parents, especially those who are illiterate, in supporting their children’s learning. CBE and efforts to enhance community capacities resulted in positive effects, such as increasing community trust in government to provide access to education and quality educational opportunities. Groups that include community members, such as SMCs, were also often the first to support infrastructure enhancements to schools, and were harnessed to monitor teacher attendance and performance. These positive effects suggest USAID’s investments in CBE and capacity-building activities targeting community members have contributed to increasing access to education for learners.

Ghana Public Works is an intervention that attempted to target both **national-level ministerial officials and line ministries** to support children’s access to education. The activity’s construction of 159 schools likely contributed to increased access to education for kindergarten-aged children. However, the aim of creating an environmentally friendly atmosphere for learning was only somewhat achieved due to the lack of running water and gender-segregated toilets, which remained barriers to proper use. Government officials’ involvement in the maintenance of newly constructed facilities consisted of incorporating them into the District Assembly’s list of facilities to maintain. However, no additional budget or support was provided to facilitate maintenance after construction by any of the line ministries or donors. The implementing mechanisms, supervision processes, and quality control activities were not effective enough to ensure activity outcomes in a timely and cost-effective manner. The lesson learned here is that national line ministries need to be adequately supported to develop and implement clear construction guidelines, monitor budget/financial protocols, and provide oversight of implementing partners, not only during construction activities, but also during ongoing maintenance. Such capacity strengthening activities will enhance the sustainability of educational structures.

Teachers are the third beneficiary group for capacity-building activities in Goal 3 education activities. The Nigeria RARA evaluation is illustrative of this because it focuses on improving instructional methods of teachers. Overall, the RARA-developed instructional approach had a significant positive impact on teachers’ reading instructional practices, children’s reading outcomes, and instructional leadership practices. Improved teachers’ reading instructional practices, in turn, resulted in significant improvement in pupils’ foundational reading skills in Hausa, including their ability to identify letter sounds, decode nonwords, and read a short narrative text. Pupils’ letter and word writing (spelling) skills also

improved.¹¹⁴ This may be due in large part to RARA's approach, which involved increasing the skills of school supervisors, who then served as reading coaches for teachers. This improved the supervisors' relationships with teachers, and built supervisors' capacity to use various RARA monitoring tools, such as classroom observations needed to better support teachers in improving reading instruction. It can be surmised that although teachers were the key target group here for capacity building, the work with teachers was done in tandem with building the capacity of their supervisors to better support them in their improvement.

The fourth beneficiary group of capacity building activities in Goal 3 education activities is **school administrators and school management teams**. An example of this is in the Jordan LETS activity, which worked directly on building the M&E capacity of school administrations to enhance their functional capacity. The LETS activity sought to institutionalize and sustain improved learning environments in schools through building the capacity of key stakeholders at the school level and at the central level in the MOE. The evaluation results show that 71 percent of sampled schools established all five components for monitoring and improving the learning environment. Nearly all schools had conducted the Results-Based Benchmark self-assessment at least once, a key component of M&E capacity-building activities. Almost all sampled schools reported a certain level of learning environment improvement, with more than half (65 percent) reporting a satisfactory level of learning environment improvement as a result of LETS, though perceptions of the level of improvement varied across schools according to gender and school level.

The evaluation results show that the LETS activity largely targeted teachers and principals, neglected to train parents and wider community members, and provided limited training to MOE staff and field directorates. Therefore, besides capacity building at the school level, the extent of LETS capacity building among local counterparts was limited.¹¹⁵ It may be concluded from this that an integrated, multilevel approach to capacity building is key to producing any key desired outcome, in this case improving the school administrative M&E capacity.

The last beneficiary group targeted by Goal 3 education activities is **parents**. The activity that best illustrates this is the Ethiopia School Community Partnerships Serving Orphaned and Vulnerable Children (SCOPSO), which focused on building the capacity of parents to better support vulnerable children to continue in school. According to the evaluation, at the household level, SCOPSO developed appropriate income-generating activities for OVCs and their guardians. As part of its economic strengthening component, SCOPSO provided training and seed capital at schools for income-generating activities such as school gardens, running microenterprises, and establishing linkages with microfinance institutions.¹¹⁶ The results indicate that even though the household-level economic strengthening support was carried out to ensure sustainability of care on the part of beneficiaries, less than 22 percent of OVC households were served by the income-generating activity support. Limited available budget, particularly for income-generating activity and food support, was the key barrier to effective implementation and outcome results.

To address this limitation and maximize support to households and communities, SCOPSO strengthened capacities of community leaders, health care providers, community-based organizations, religious leaders, and Women's Affairs officials to support OVCs and their guardians. Specifically, these stakeholders were empowered to plan an activity and manage implementation, including the M&E process. This example helps to illustrate the challenges that can occur for an education intervention and the real-time course

¹¹⁴ RTI International, *Nigeria Reading and Access Research Activity (RARA): Results of an Approach to Improve Early Grade Reading in Hausa in Bauchi and Sokoto States*, 3-7.

¹¹⁵ Moubayed et al., *Final Performance Evaluation USAID/Jordan Learning Environment Technical Support Program*, 6.

¹¹⁶ PRIN International Consultancy and Research Services PLC, *Final Performance Evaluation of the School-Community Partnership Serving Orphan and Vulnerable Children Affected by HIV/AIDS (SCOPSO) Project*, 6.

corrections that are needed to ensure that capacity strengthening outcomes are achieved in spite of budgetary, logistical, and other challenges.

C. Methodologies Employed to Improve Institutional Capacity

- *The Goal 3 evaluations reviewed demonstrate that education activities employ a variety of models to improve institutional capacity.*
- *In many instances, this synthesis infers the effectiveness of these models through reporting of positive, negative, or no effects on key outcomes, but in some instances (noted throughout this report) the reviewed evaluation reports did not contain enough information to determine the effectiveness of particular education activities.*

Seven methodology approaches are used to improve institutional capacity, including training workshops (e.g., Nigeria RARA), mentoring models (e.g., Ghana-Public Works), secondment of international NGO experts in ministry offices (e.g., Afghanistan BELT), curricular revisions (e.g., Jordan LETS), improved skills training (e.g., Ethiopia SCOPSO), civil society engagement (e.g., Afghanistan ALSE), and public-private partnerships (e.g., Senegal EBD).

The most popular methodology used by interventions to improve institutional capacity (specifically to improve teaching or school management capacity) is skills training, followed by training workshops and curricular revisions. The methodology cited as most effective was improving teaching in tandem with improved school management, as illustrated in the Nigeria RARA evaluation. The methodology least used is public-private partnerships, which was undertaken only by Nigeria NEI and Senegal EBD. Civil society engagement, of increasing political importance, was employed as a methodology to build capacity in only three activities included in this synthesis: Nigeria NEI, Afghanistan ALSE, and Ghana Public Works Construction.

2. Violence Prevention and Countering Violent Extremism

Guiding Questions

What models have EiCC activities developed to counter violent extremism? What distinct populations are targeted by these activities? How effective are these activities and under what conditions?

Results Summary

The review found that models included initiatives for youth empowerment, social and economic inclusion, media and messaging, improved local governance, reconciliation, and conflict mitigation. Most evaluations reviewed had a focus on social and economic inclusion, which usually took a WfD approach by offering people skills training, entrepreneurship promotion, and employment services. Another common approach was improving local governance. Other approaches used for conflict mitigation included youth empowerment, and media and messaging. The least common approach was reconciliation, which was part of only three activities.

- *Violence prevention and CVE activities are complex and uniquely tailored to diverse political, social, cultural, and economic contexts.*

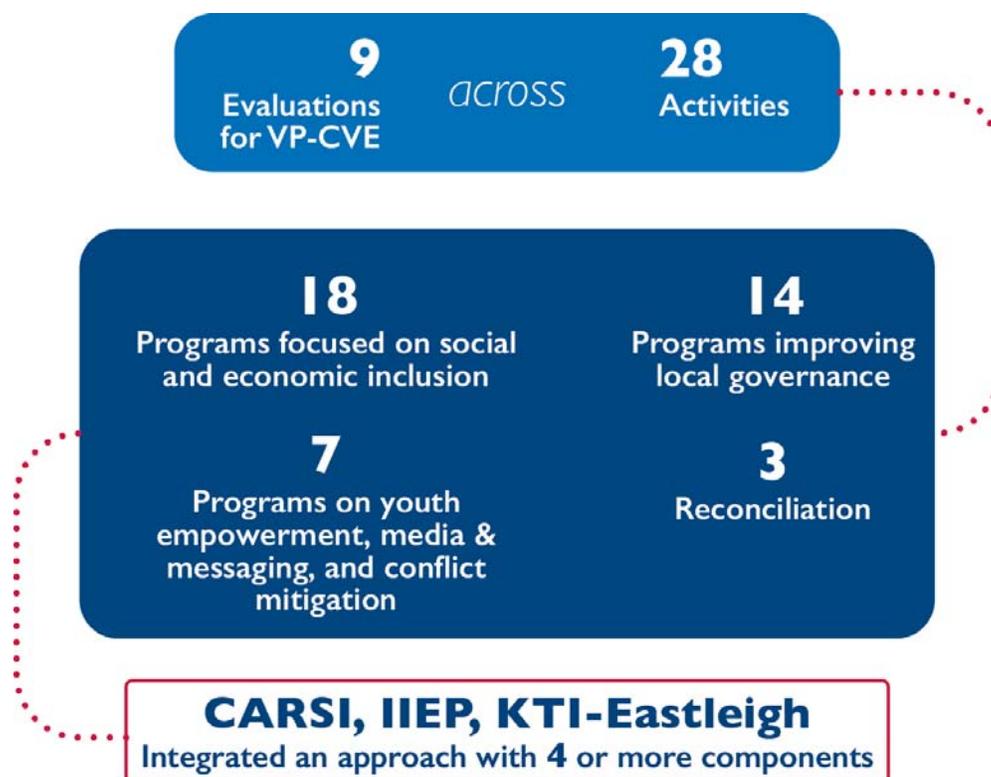
A. Definitions of Violence Prevention and CVE

- Violence prevention and CVE are two distinct outcomes, though there may be commonalities in programming.
- This analysis suggested the need for a multipronged, cross-sectoral approach to implementing effective and sustainable violence prevention and CVE activities.

“Violence prevention and CVE” refers to both violence prevention and countering violent extremism as program outcomes. They are two distinct outcomes, though there may be commonalities in programming. CVE refers to activities that address social, economic, governance, and other grievances that can fuel violent extremism or radicalization of individuals and communities. Violence prevention consists of actions and interventions that seek to decrease or eliminate underlying risk factors that lead to violent and/or criminal behavior, and that promote citizen security. Violence can be political, gang-related, and/or interpersonal (including domestic).¹¹⁷

A range of groups are targeted by violence prevention and CVE activities, including at-risk youth and other vulnerable populations (e.g., survivors of sexual and gender-based violence), government partners, and other agencies such as local law enforcement and judicial institutions. The interventions documented in the evaluations reviewed for this synthesis differed widely across the varying contexts in which they were implemented. It is thus challenging to pull out broad categorizations of what programmatic approaches are most or least effective. The analysis did suggest the need for a multipronged, cross-sectoral approach to implementing effective and sustainable violence prevention and CVE activities.

FIGURE 17: INTERVENTIONS BY FACTORS RELATED TO INDIVIDUAL RADICALIZATION AND RECRUITMENT



¹¹⁷ USAID, *Crime and Violence Prevention Field Guide*.

Five activities focused on violence prevention (CARSI El Salvador, Guatemala, and Honduras; the broader Caribbean CBSI; and Honduras A Ganar). Two activities focused on peace building and reconciliation processes as well as violence prevention (Midterm Performance Evaluation of USAID/Macedonia's Interethnic Integration in Education Program and Final Evaluation Youth Engagement to Promote Stability Timor-Leste). One evaluation from East Africa focused on CVE specifically (Midterm Evaluation of Three Countering Violent Extremism Projects in East Africa).

B. Beneficiaries of Violence Prevention and CVE Activities

- *Violence prevention and CVE activities are deeply contextualized, making it difficult to discern broad trends on the most reliable approaches to ensuring effective violence prevention or CVE activities.*
- *Such activities target a range of beneficiaries including vulnerable communities and populations, government institutions and partners.*

Violence prevention and CVE activities targeted at-risk youth, government partners and local law enforcement authorities, judicial institutions, students, vulnerable communities, and vulnerable populations (i.e., LGBTI individuals and survivors of trafficking and sexual and gender-based violence). Fifteen of the activities focused on at-risk youth, defined most often as unemployed and out of school (i.e., G-Youth, KTI-Eastleigh, and SYLP) or as juvenile offenders (i.e., CARSI). However, three evaluations reviewed focused on activities that targeted at-risk children in school (i.e., Interethnic Integration in Education Program [IIEP] in Macedonia, Improved Reading in Jamaica, and CLP in Yemen). Only one activity addressed LGBTI individuals and survivors of trafficking and sexual and gender-based violence (i.e., Fi Wi in Jamaica).

The violence prevention and CVE activities took varied approaches under unique conditions. Given the different contexts, intervention approaches, and beneficiaries, the outcomes measured were not the same across the reviewed evaluations. Thus, it is difficult to discern broad trends on the most reliable or successful approaches to ensuring effective violence prevention or CVE programming. Nonetheless, it is possible to specify the programmatic elements that seemed most effective in particular contexts. Examples are provided for the CARSI activities, CBSI activities, IIEP in Macedonia, USAID CVE activities in East Africa, and Youth Engagement to Promote Stability (YEPS) in Timor-Leste.

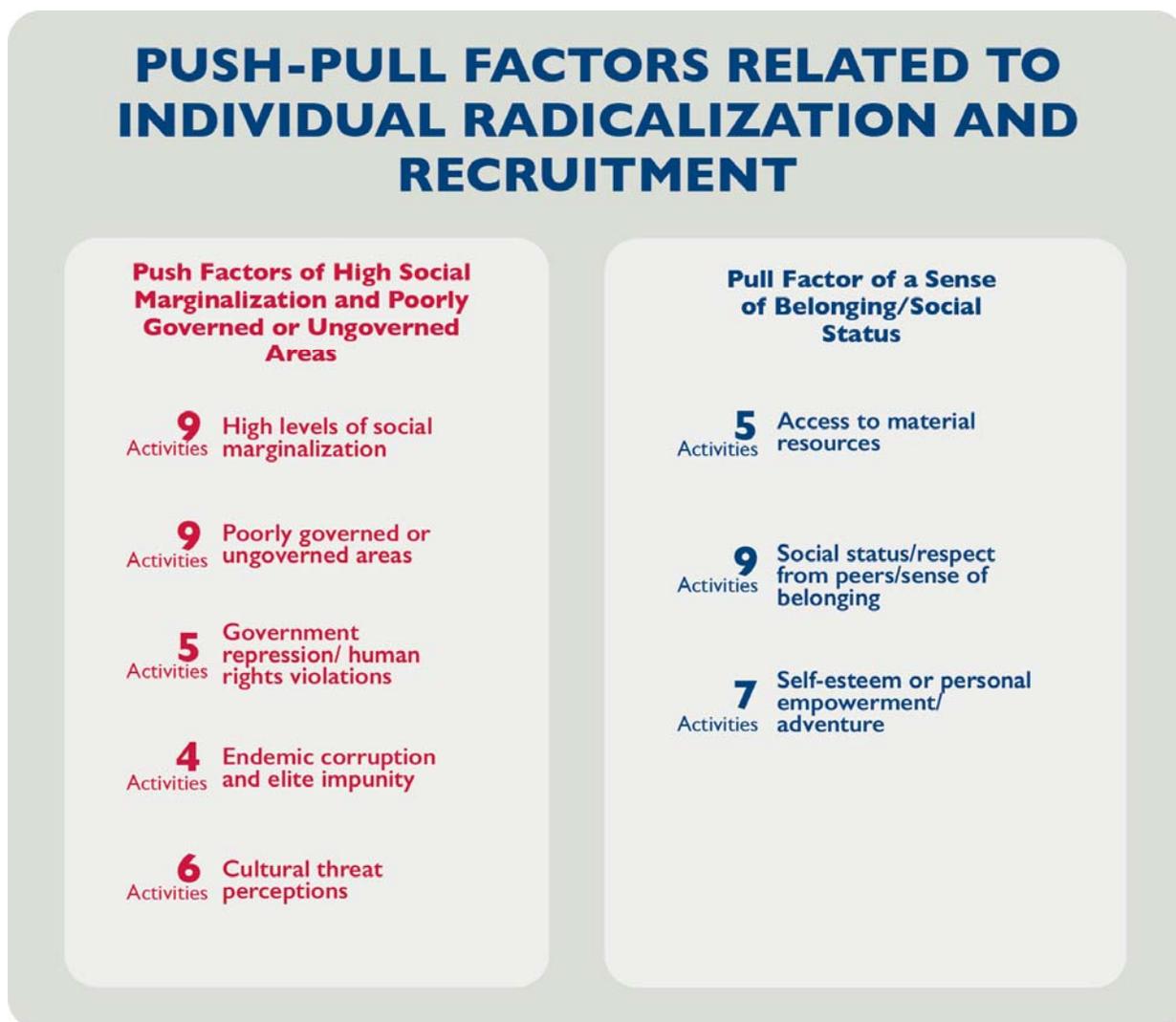
C. “Push-Pull” Factors Related to Individual Radicalization and Recruitment

- *The most effective intervention approaches to increase learners' access to education while also mitigating causes of violence and rehabilitate communities were: engaging youth and increasing employment opportunities for them; promoting community participation; strengthening local governance capacity (e.g., justice reform measures, assistance to police); and improving access to quality services in education and health.*

The policy framework outlined in USAID's Development Response to Violent Extremism and Insurgency: Putting Principles into Practice (USAID 2011) puts forth a series of principles to improve the effectiveness of its development tools in responding to violent extremism and insurgency and building capacity of partners in these challenging environments. The policy identifies the push and pull factors that lead to violent extremism or insurgency as well as those that can influence radicalization of individuals. The “push” factors analyzed from this cluster of evaluations include: high levels of social marginalization and fragmentation; poorly governed or ungoverned areas; government repression and human rights violations; endemic corruption and elite impunity; and cultural threat perceptions. The “pull” factors measured include access to material resources; social status and respect from peers and a sense of

belonging; adventure; and self-esteem or personal empowerment, particularly for individuals who feel victimized and marginalized.

FIGURE 18: PUSH-PULL FACTORS RELATED TO INDIVIDUAL RADICALIZATION AND RECRUITMENT



See Table 20 in Annex 5 for the range and frequency of push and pull factors related to individual radicalization and recruitment.

The seven violence prevention and CVE activities included in this synthesis sought to increase access to primary education and youth empowerment while also including WfD components to reduce “risky behaviors.” Such programmatic components were seen as vehicles primarily aimed at mitigating the causes of violence and instability to rehabilitate conflict-affected or high-crime-zone communities.

The most effective intervention approaches to increase learners’ access to education while also mitigating these causes of violence and rehabilitate communities were:

- Engaging youth and increasing employment opportunities for them;
- Promoting community participation;

- Strengthening local governance capacity (e.g., justice reform measures, assistance to police); and
- Improving access to quality services in education and health.

All seven of the evaluations classified under violence prevention or CVE addressed push factors of high social marginalization and poorly governed or ungoverned areas. A majority addressed cultural threat perceptions, except for A Ganar in Honduras. The CARSI evaluations in El Salvador, Honduras, and Guatemala addressed all eight push and pull factors described above. Macedonia IIEP was the only intervention focused on building capacity for interethnic integration. All interventions addressed the pull factor of a sense of belonging/social status. Yemen CLP is the only activity implemented in a direct conflict setting. The rest of the interventions focused on gang violence and crime prevention. For how each of the activities addressed factors most often related to individual-level radicalization and recruitment (see also Table 20 in Annex 5).

In some instances, activities addressed multiple factors in a regional setting. For example, CBSI covered a constellation of interventions across Antigua and Barbuda, Barbados, Dominica, Dominican Republic, Grenada, Guyana, Jamaica, St. Lucia, St. Kitts and Nevis, St. Vincent and the Grenadines, Suriname, and Trinidad and Tobago, all of which focused on reforming the juvenile justice system. CARSI, a multicountry intervention in El Salvador, Guatemala, and Honduras, used schools as a key mechanism for crime prevention and conflict meditation through their “School Harmony” (*Convivencia Escolar*) activity. Further, this activity sought to mainstream a variety of student leadership groups, along with training teachers and students in mediation and alternative conflict resolution, to improve learning environments for youth at risk of being radicalized in school.

D. Core Programmatic Principles for Activities Targeting Violent Extremism

- *USAID’s Goal 3 team has highlighted the need for education interventions in crisis and conflict contexts to demonstrate (a) flexibility, agility, and procurement speed; (b) intensive program management; (c) innovation, evaluation, and learning; and (d) informed risk-taking.*

During the programming cycle, a core body of program principles guide USAID’s design and implementation of development programming targeted at violent extremism and insurgency.¹¹⁸ These include: (a) focus on drivers of violent extremism and insurgency, (b) promote inclusive country ownership, (c) exercise selectivity, (d) take a coordinated and integrated approach, (e) tailor and coordinate communication, (f) think locally and bring an entrepreneurial approach, and (g) consider transnational strategies.

USAID’s Goal 3 team has highlighted the need for education interventions in crisis and conflict contexts to demonstrate (a) flexibility, agility, and procurement speed; (b) intensive program management; (c) innovation, evaluation, and learning; and (d) informed risk-taking, among other variables.

¹¹⁸ USAID, *Development Response to Violent Extremism and Insurgency: Putting Principles into Practice*.

FIGURE 19: CORE PROGRAMMATIC PRINCIPLES FOR ACTIVITIES TARGETING VIOLENT EXTREMISM



See Tables 21 and 22 in Annex 5 for the range and frequency of core programmatic principles for activities targeting violent extremism.

E. Case Studies Demonstrating Effectiveness of Varying Approaches

The CARSI activities in El Salvador, Guatemala, and Honduras all sought to counter crime and violence. These countries have each struggled with intense gang violence and crime. CARSI attempted to reduce crime and violence by promoting a holistic set of interrelated approaches to enhancing community policing and citizen security. Specifically, the activities sought to strengthen community capacity to combat crime by creating educational and employment opportunities (described below) for at-risk youth in participating countries in Central America. The evaluation demonstrated that the community-driven policing (versus more traditional law enforcement approaches) was effective as it attempts to address the root causes of crime, instead of addressing crime after it has become a severe problem.

The impact evaluations illustrate that CARSI reduced crime and violence (i.e., robberies, illegal drug sales, extortion, and murder), decreased perceptions of neighborhood insecurity, lowered disorderly behavior

such as loitering and gang fights, increased communities' sense of control in their neighborhoods and sense of interpersonal trust among neighbors, and improved community satisfaction with police—an interesting finding given that communities were now empowered to monitor and ensure their own safety and well-being.

The evaluation, through qualitative inquiry, also suggested the critical positive role that families, churches, and schools play in ensuring at-risk children and youth stay in school, engage in socially positive activities, and decrease their likelihood of participating in criminal activity, such as joining a gang, by reducing out-of-school suspensions (so students are not left at home unsupervised and at greater risk of joining gangs or engaging in criminal activity). The results demonstrated the critical role that schools play in crime prevention by creating better, safer environments for at-risk youth with innovative *Convivencia Escolar* (School Harmony) activities, institutionalization of student leadership groups within schools, and the training of teachers and students in mediation and alternative conflict resolution. Taken together, these findings point to the need for a multipronged, cross-sectoral approach to implementing effective and sustainable violence prevention activities.

Due to its broad coverage of 12 Caribbean countries and the design/nature of a performance evaluation, there are limited examples of **CBSI's** effectiveness in violence prevention. In general, the CBSI initiative tackled common issues among the Caribbean countries such as the illicit trade of drugs and small arms, crime and violence, government corruption, the judicial system's treatment of juvenile offenders, and the lack of economic and social opportunities for at-risk youth (i.e., unemployed and out of school).

Evaluations of the Caribbean Youth Empowerment Program in the eastern Caribbean and Empowering Jamaica's Youth demonstrated that the activities contributed toward improving economic opportunities for at-risk youth by strengthening life skills and increasing graduation rates and employment of youth. Intervention activities such as establishing a marketplace in Kingston, Jamaica, for youth to apply skills and knowledge learned in the Junior Achievement's entrepreneurial curriculum (Empowering Jamaica's Youth) were found to be innovative and uniquely effective for participating youth.

In contrast to the CARSII evaluation, which demonstrated the effectiveness of the community policing approach, neither the Community Empowerment and Transformation Activity II (COMET II) evaluation in Jamaica nor the evaluation of Police Reform in the Dominican Republic demonstrated successful community and law enforcement cooperation. This may suggest that the activities need to further strengthen their outreach to schools, communities, churches, universities, and youth groups, especially at-risk youth, to create ownership and buy-in for the community policing approach.

Through education activities, **IIEP** seeks to mitigate ethnic violence, primarily between Macedonians and Albanians, as the Macedonian government resolves interethnic tension by recognizing rights of Albanian minorities. As described earlier, the program succeeded in gaining the support from local schools more than from district education officials and parents, though the evaluation did not specify reasons for this.

Media and messaging campaigns typically received attention from local media, but less from national media, perhaps due to localized outreach efforts. Such campaigns were effective locally as they provided students with a chance to voice their experiences with IIEP and their increasing feeling of safety in their schools. Teacher training aimed at interethnic collaboration led to implementation of some integration topics, but some teachers still experienced resistance from students and parents (perhaps as this required coverage of very sensitive content). Teachers from demonstration schools supported by IIEP reported satisfaction with their training on ethnic integration, and students even reported increased contact and dialogue with members of other ethnic groups.

Involving parents remained a significant challenge for the demonstration schools. Given the critical role parents and families have been found to play in ensuring a host of positive academic, behavioral, and

social-emotional outcomes for learners more broadly, future activities to prevent violence in the form of CVE must take care to strengthen parent engagement.¹¹⁹

The **USAID CVE activities in East Africa (i.e., G-Youth, KTI-Eastleigh, and SYLP)** were implemented in Somalia and areas of Kenya with large Somali populations deemed at risk for recruitment to Al Shabaab. When detailing the context and background for the activities, this evaluation did explicitly discuss the drivers of violent extremism, such as lack of civic engagement; absence of self-efficacy; “lack of belief in legitimate avenues for youth identity and engagement, such as *youth associations*; a distorted sense of *identity*; and a belief that *violence* is permitted or even encouraged by Islam to address grievances.”¹²⁰ These drivers informed the programming approach taken in these countries.

The overall findings for these activities were identical with respect to youth perceptions of civic engagement (e.g., participation in community meetings, issues raised with authorities), self-efficacy (e.g., how much an individual can do to solve community problems), and satisfaction with government decision-making and extremist violence. Youth in the treatment groups had statistically significant, higher senses of civic engagement, senses of optimism about their future, and self-efficacy relative to youth in comparison groups from their cities. This was not consistent for female youth, who often reported lower scores compared to males.

Interestingly, madrassa students and graduates reported lower rates of engagement, self-efficacy, and sense of preparedness for the job market compared with graduates from secular schools. Finally, despite mixed results with respect to civic engagement, identity, and self-efficacy, the evaluation demonstrated that youth from the treatment and control groups in all locations largely rejected the use of violence in the name of Islam, without any statistical differences between treatment and control. Investigating the connections between access to education and differences in education service delivery (e.g., madrassa versus secular) and CVE outcomes should be a continued research priority for Goal 3.

Furthermore, **SYLP** reported that participants reporting no income dropped from 88.7 percent before the intervention to 14.6 percent after the intervention. SYLP participants noted an increase in their mean income as well. The intervention activities that appeared to be linked most closely with these important findings included the formation of a savings cooperative society that trained youth on the importance of savings and pooling resources, youth groups as fora for information exchange, and (contributing more broadly toward social cohesion at a community level) involvement of networks of community members from different ethnic groups in the formation of district peace committees.

Finally, the **YEPS** evaluation focused on an activity in Timor-Leste, a nation that had experienced conflict throughout the 2000s. YEPS aimed to enhance youth engagement with communities and the government through training on debating techniques and soft skills such as public speaking, and through a youth-produced radio show focusing on issues such as peace and reconciliation, domestic violence, conflict resolution, and leadership. The radio show is an innovative component designed to teach youth how to address and resolve conflict peaceably—on topics youth have identified as of most interest to them. Based on focus group discussions with a range of beneficiaries, the evaluation found that the activity had inculcated communication, self-confidence, public speaking, leadership, and problem-solving skills in youth as a result of the training and radio shows. However, YEPS was meant to connect “youth with national decision makers to explore and discuss salient issues and drivers of youth related conflict.” The

¹¹⁹ See: UNESCO, *Teaching and Learning: Achieving Quality for All*; Walker et al., “Effects of Growth Restriction in Early Childhood on Growth, IQ, and Cognition at Age 11 to 12 Years and the Benefits of Nutritional Supplementation and Psychosocial Stimulation,” 36-41.; Walker et al., “Effects of Early Childhood Psychosocial Stimulation and Nutritional Supplementation on Cognition and Education in Growth-Stunted Jamaican Children: Prospective Cohort Study,” 1804-1807.; Gertler et al., “Labor Market Returns to Early Childhood Stimulation Intervention in Jamaica,” 998-1001.

¹²⁰ Swedberg and Reisman, *Mid-Term Evaluation of Three Countering Violent Extremism Activities*, 15.

evaluation found that most youth indicated that they had not substantially engaged with community decision-makers or Timorese government officials. Vocational opportunities were still lacking for these youth, and youth felt their potential was limited. This suggests a critical gap that, if bridged, could contribute toward violence prevention and CVE.

F. Combination of Youth Empowerment/Social Integration and WfD Components

The reviewed evaluations measured a range of outcomes directly or indirectly related to learners' access to quality education, including youth engagement and skills building (i.e., Timor-Leste YEPS, East Africa CVE), crime levels and victimization (i.e., Guatemala CARSI, Honduras CARSI, the Caribbean CBSI), citizens' sense of security and perceptions of neighborhood disorder (i.e., Guatemala CARSI, Honduras CARSI, CBSI), and cooperation and collaboration between youth and decision-makers (i.e., Timor-Leste YEPS).

Most interventions that sought to enhance youth outcomes demonstrated positive effects on outcomes such as leadership, problem solving, and conflict resolution skills (i.e., Timor-Leste YEPS, El Salvador CARSI, Guatemala CARSI, East Africa CVE).

3. Improved Programming

Guiding Questions	Results Summary
<p>What operational or management practices are employed in EiCC activities to address challenges of implementation in conflict and crisis contexts (e.g. risk-informed programming, feedback loops)? What types of feedback loops/processes used are being used in EiCC activities? Under what conditions are EiCC activities able to adapt or pivot due to changes in context? Are there case studies? What transition strategies exist between humanitarian and development programming, and how effective are these strategies in ensuring continued delivery of relevant, quality education programming?</p>	<p>The review found that situational analyses in conflict/crisis settings provided the most information about the barriers that conflicts/crises may be producing to learners' access to education, and how such barriers could be addressed in subsequent activity design. Most situational analyses focused on understanding how conflict/violence within the community has influenced primary school-aged children's access to basic education. All situational analyses engaged with local actors in either implementation of assessment or usage of resultant data to inform future programming responses.</p>
<ul style="list-style-type: none"> ➤ <i>Situational analyses of emergency/conflict-affected regions provide real-time, nuanced data to improve activity design and implementation. A majority of the situational analyses focused on understanding how conflict/violence within the community has influenced access of primary school-aged children to basic education. Such analyses relied on similar data collection approaches to gain perceptions of a range of school, community, and national-level stakeholders to ensure activity design and implementation remained relevant to target populations.</i> 	

A. Situational Analyses: An Overview

- *Improved education programming that reflects changing contexts or beneficiary needs is due in part to making immediate use of assessment or evaluation data. Situational analyses of conflict/crisis best represent the subset of studies that attempt to capture—in real time—an honest and accurate picture of how a conflict or crisis is interacting with*

learners' access to basic education, and that provide recommendations on how to address these barriers in precarious and ever-changing situations.

- *The timing (e.g., prior to activity implementation) and the design of such assessments, which are intended to be rapid and continuous, are likely why situational analyses or rapid risk analyses are particularly useful in crisis and conflict settings.*

Five studies were categorized as situational analyses of emergency/conflict-affected regions. These analyses were specifically related to documenting how children and youth could gain and maintain equitable access to quality education. Three main measures were analyzed:

- Main outcomes of interest (same types of student, teacher, community-level outcomes),
- Type of quantitative and qualitative data collection methodologies used, and
- Level of engagement with local actors to implement the assessment or use the resultant data.

Situational analyses, beneficiary outcomes, and process results are the three main types of evaluation results presented in this collection of evaluations for Goal 3. All situational analyses focused on school-aged children affected by conflict, and all analyses surveyed this target group (in addition to teachers, school system administrators, parents, and the community) to obtain a holistic perspective on the impact of emergency/conflict on children's access to learning. A majority of these situational analyses also conducted secondary data analysis, in addition to completing qualitative key informant interviews. Also, most of these assessments engaged with local actors in implementing the assessment and/or provided resultant data to country teams to inform the design of education interventions in crisis settings.

B. Case Studies Demonstrating Value of Situational Analyses

An excellent example is the **Mali Rapid Education and Risk Analysis (RERA)** conducted in the Gao region of the country, which illustrates how a detailed risk assessment can inform activity design and improve implementation in an ongoing and immediate way. Longer-term, "traditional" evaluations, such as the CLP evaluation in Yemen, also demonstrate how evaluation data can capture and document adaptation of an activity based upon changing contexts and needs. The ultimate result of these assessments is to inform subsequent activity design to address the barriers learners face in accessing quality education in these contexts.

The **Gao RERA** was designed to identify potential communities where USAID could intervene, based on the needs of the communities (e.g., training and employment opportunities for young people), the number of potential beneficiaries, and the security realities. RERA was also designed to uncover the populations' perceptions of the causes of the conflict, the perceived risks associated with accessing education in the conflict context, and the potential consequences of the conflict on education, educational needs, and the expectations of the populations. Finally, RERA was also designed to identify the factors, stakeholders, and local mechanisms of division, and of cohesion and resilience, which would in turn inform the design, implementation, and management of USAID interventions in the region.

The results of RERA pointed to the strong negative impacts of the conflict on the education system in Mali, including closed schools, damaged education infrastructure, threats to student safety during travel to and from school and sometimes while at school, teacher absences, and community mistrust of teachers and school officials. Perceptions of security were strongly influenced by gender, with female students reporting greater fears of abuse, harassment, and sexual assault. Finally, the assessment pointed to the potentially powerful roles youth and community members could play in ensuring the safety of learners, maintenance of school facilities, and supervision/policing of communities and learning environments.

This analysis helped inform the suite of interventions implemented in the Gao region of Mali including training to enhance the psycho-social health of students, return-to-school activities, the construction and rehabilitation of schools, curricular content and decisions around language of instruction, youth vocational training, and the distribution of school kits. In addition, this analysis also informed the way in which target communities were selected, engaged, and retained in the interventions by USAID, including a process of social negotiations, resilience activities, and the promoting of peace and the sustainability of achievements. In this way, a real-time and rapid assessment approach such as the RERA can directly contribute toward improved programming.

Another good example of improved programming comes from the evaluation of the **CLP intervention in Yemen**. CLP was initially designed to encompass a short, rapid-response activity, a medium-term stability activity, and a cluster school/training center activity lasting between 6 and 18 months, serving communities in specific and limited geographic areas to increase stability in a violence-torn country. However, political instability within Yemen resulted in activity delays and major changes in CLP leadership, and caused a rebirth of this activity. According to the evaluation report, this led to shifts in focus from short-term stabilization to the realization of longer-term development goals and to strengthened cross-sectoral linkages among agriculture, health, and governance. This was achieved through work with school-based Father-Mother Councils, community library activities, and rehabilitation of school facilities.

Longer-term activities included teacher training, working with the government to improve curricula, assistance for the establishment of accelerated learning for primary-grade school children, and financing the construction of new schools. These efforts quickly led to positive impacts across different sectors.¹²¹ The M&E systems—which included continuous feedback loops, a strong marker of effective, adaptive, relevant programming—further evolved and improved the linkages to the Yemen Monitoring and Evaluation Project (YMEP).

Some assessments, such as the **Nigeria Community Education and Conflict Assessment (CECA)**, clarified the perceptions and demands of beneficiaries for education and highlighted the impact of internal and external threats of violence on the safety of learning environments. This assessment is a strong example of how such an analytic approach can extrapolate important programmatic recommendations from community perceptions of how the conflict context is impacting education in a community. Study participants reported that a lack of security forced many families and their children to move frequently from place to place as they feared being attacked (either on their way to/from school or at school)—and this movement, in turn, results in deteriorating school conditions (i.e., closure of some schools, and teacher shortages, especially in schools receiving displaced persons).

To ensure safer learning conditions, the Nigeria CECA assessment pointed toward the importance of “raising awareness on school attacks or bombs” and preventing gender-based violence through community sensitization activities. Each of these was chosen as an important curricular or extracurricular topic that school-level stakeholders would like children to learn more about. Community mobilization efforts also proved essential in negotiations with jihadis, in accompanying students to/from schools to ensure safer journeys, and in maintenance of operational schools. The assessment also pointed toward the possibility of recruiting resident volunteer teachers to gradually reduce the teacher shortage. Such an activity would mobilize the community, ensure the community’s trust in these teachers, and improve the safety and well-being of the teachers themselves as they would be drawn from the local community.

¹²¹ Gurevich et al., *Performance Evaluation of the Education Program of the Community Livelihoods Project (CLP) Final Report*, 3.

4. Safer Learning Environments for Children and Youth

Guiding Questions

To what degree, and how effectively, are EICC activities addressing different categories of threats to learners' access to education, such as internal threats including school-related gender based violence (SRGBV); external threats such as gang violence and schools caught in cross-fire or under attack; environmental threats related to physical health (e.g. infectious diseases, malnutrition) and; environmental threats related to natural hazards (e.g. infrastructural damage, school closures)?

- *Practically all evaluations reviewed under Goal 3 pointed to the significant impact that internal, external, and environmental threats posed to creating and maintaining safe learning environments for children and youth.*
- *Safer learning environments for children and youth are best maintained through community engagement and mobilization, curricular improvements, and infrastructural improvements.*

Results Summary

The review found that in most evaluations reviewed, respondents mentioned the effects of evolving internal or external threats of violence—such as school attacks and gender-based violence—on school attendance, and community perceptions of how the evolving conflict context is impacting education, and this yielded a plethora of recommendations. This suggests continued assessments are a worthwhile focus for USAID's education investments in crisis and conflict strategy.

A. An Overview of Child-Friendly Spaces and Safer Learning Environments

- *Several programmatic elements emerged as effective regarding safer learning environments for children and youth: community sensitization to the importance of education as a pathway to peace building, introduction of conflict-sensitive curricula on topics such as gender-based violence and interethnic integration in schools, school rehabilitation efforts, and the selection of community volunteers to address teacher shortage issues.*

Education experts in the humanitarian sector advocate schools as child-friendly spaces. In the context of an emergency, education can serve as a protection mechanism that “involves the ability of schools to provide a safe and secure space that promotes the well-being of learners, teachers, and other education personnel.”¹²² In particular, “schools and child-friendly spaces protect children both physically and psychologically during an emergency. They offer protection against exploitation and harm, and create a sense of normality and routine which is crucial to the healing process following distressing experiences.”¹²³ The humanitarian community externally popularizes the notion that education automatically protects children, yet the history of emergencies contradicts the idea that schools necessarily make children safe from violence and abuse.

Most evaluations reviewed under Goal 3 demonstrated that children and youth can be exposed to a number of internal, external, and environmental threats that reduce their sense of physical and emotional safety in learning environments. **The evaluation results also demonstrated that real-time, ongoing risk analyses and assessments are a worthwhile methodological approach for USAID's education activities in evolving crisis and conflict settings.** While the evaluation sample is small, and the activities were contextualized to uniquely address the needs of a range of conflict

¹²² Dryden-Peterson, *Refugee Education: A Global Review*, 32.

¹²³ Global Education Cluster, *Education: An Essential Component of Humanitarian Response*.

contexts, the following factors were most beneficial: community sensitization to the importance of education as a pathway to immediate improved safety (and possibly longer-term peace building at the community level, as in the case of the Macedonia evaluation), introduction of conflict-sensitive curricula on certain topics (e.g., gender-based violence and interethnic integration in schools), school rehabilitation efforts, and training/support of community volunteers to address teacher shortages.

B. Case Studies Demonstrating Effective Interventions Targeting Safer School Environments

Evaluated activities addressed external and environmental threats. **Yemen CLP**, for example, was undertaken during armed internal conflict. In its early stages, USAID/Yemen and CLP designed and began to implement the intervention, then rapidly revised strategies and activities to address the rapidly evolving internal conflict in Yemen, both to promote stability and to build a foundation for long-term development. The evaluation found that the key elements of these activities that contributed to safer learning environments were school rehabilitation efforts and, most significantly, the implementation of the Yemen Early Grade Reading Approach (YEGRA).

The school rehabilitation efforts included improving sanitary facilities for girls and improving physical access for children with disabilities; additionally, CLP distributed over 21,953 desks to schools. Community support for CLP educational interventions was facilitated by the introduction and strengthening of Father-Mother Councils.¹²⁴ Communities could choose whether they wanted fathers only, mothers only, or joint councils. These councils were responsible for addressing safety-related issues within schools, including rehabilitation efforts, teacher absenteeism and education quality, and student dropout rates. These programmatic elements—the school rehabilitation activities and community mobilization efforts—were in large part successful because of the “generally applicable” nature of the YEGRA activities; they had no religious or political affiliation and were adaptable to community preferences. As a result, these activities were palatable to a diverse array of parents and community members and led to safer learning environments for children and youth.

The **Performance Evaluation of Public Works Construction Activities to Increase Access to Education in Ghana** noted that the intervention was somewhat effective in creating an environmentally friendly atmosphere for learning through school construction and rehabilitation efforts. For example, although lack of running water and gender-segregated toilets remained barriers to proper use,¹²⁵ district-level education authorities were found to have positively contributed to the “availability of basic education management infrastructure, even though they were constrained by overcrowded facilities and lack of electricity and equipment.” This again points to the importance of capacity building at a local level to ensure ownership and sustainability.

Interestingly, the evaluation also found that only one of the implementing partners facilitated participation of local stakeholders in the maintenance of constructed schools. This is likely why communities did not have favorable impressions of the education facilities that were constructed for them; “when probed about how much these addressed their needs, many stakeholders voiced issues about the structures, such as [kindergartens] were not big enough to absorb current demand, and toilet blocks were not practical for use by male and female students.” This finding highlights the critical importance of listening to school-level stakeholders and community voices, and

¹²⁴ Gurevich, et al., *Performance Evaluation of the Education Program of the Community Livelihoods Project (CLP) Final Report*, i.

¹²⁵ Rodriguez et al., *Performance Evaluation of Public Works Construction Activities to Increase Access to Education in Ghana*, 2.

understanding their perceptions, in ensuring learning environments that are appropriate, safer, and more accessible for children and youth.

There were no safety-related findings (i.e., incidents of violence, abuse, or exploitation in or around schools) expressly stated in the evaluation. The evaluation measured the construction's influence on school attendance. Because the study lacked a viable comparison group, the evaluators could not attribute the increase in access to education to the Ghana Public Works Construction intervention. However, because the intervention's construction of kindergartens coincided with an increase in student attendance, and the evaluators could not find any other probable cause, the evaluation inferred that these activities likely contributed to increasing access to education for these kindergarten-aged children.

Finally, **IIEP in Macedonia** was a unique activity in that it was the only one that addressed interethnic issues related to historical conflict and peace building. Specifically, the activity was designed to address the impact of lingering ethnic tensions for ethnic Albanians and other minority ethnic groups in the educational system. USAID/Macedonia sought to build a broad public understanding of the benefits of an integrated educational system in Macedonia through community outreach, capacity building of school management and teachers, creation of demonstration schools, and providing incentives to schools and communities. Further, in its efforts to target all primary and secondary schools across Macedonia, the activity sought to build the capacity of the Ministry of Education and Science, the Bureau for Development of Education, the State Education Inspectorate, and the Vocational Education and Training Center to support interethnic integration in education (IIE) activities.¹²⁶

The evaluation did not report whether IIE activities directly led to improved safety, but instead focused on activities that identified and rectified cultural and ethnic stereotypes caused by conflict and historical mistrust. The evaluation found that two activities in particular contributed to reducing some students' and parents' fears of violence or clashes with members of different ethnic groups at schools.¹²⁷

One key activity was establishing a fund for school refurbishments—offered as an incentive for schools to participate in ethnic integration activities. With these funds, schools designed their own IIE activities and involved school management, local officials, parents, and community members in whatever ways were most appropriate and meaningful to them. As demonstrated in other evaluations and assessments reviewed for this synthesis, community mobilization was a key factor in ensuring safer learning environments. A second key activity was media campaigns targeting local communities and school-level stakeholders, through which students could report their own stories of how interethnic education activities positively impacted their education experiences. The evaluation report hypothesized that this, in turn, could lead to dissipating the cultural and ethnic biases that create mistrust and ethnic tension within the school. Future studies should investigate whether such perceptions in turn contribute (directly or indirectly) to safer learning environments in crisis and conflict settings.

¹²⁶ Naskova et al., *Midterm Evaluation of USAID/Macedonia's Interethnic Integration in Education Project*, vi.

¹²⁷ In emergencies, protracted displacements, and other crisis and conflict affected settings, children attending school often suffer harm that exacerbates existing trauma and psychosocial issues. Sarah Dryden-Peterson, a recognized global expert on education in emergencies, explains that the United Nations High Commissioner for Refugees (UNHCR) has acknowledged the problem for nearly a decade. Inside schools children may be subjected to sexual exploitation, corporal punishment, attacks by armed groups, discrimination, and bullying. Typically, classmates and teachers that perpetrate mistreatment against students do so because of national, ethnic, or linguistic differences.

5. Equity

Guiding Questions

How is equity defined, measured and monitored in EiCC activities? In ensuring access to education, to what degree, and how effectively, do EiCC activities address the specific and unique needs of individual children and youth in crisis and conflict contexts (e.g. displaced populations, marginalized groups)? For whom are these activities most effective? Under what conditions are activities able to structure and deliver a range of educational services (e.g. accelerated/alternative education programs)? Under what conditions are such activities most effective? Which types of teaching and learning strategies are employed most effectively, and for whom, in conflict and crisis contexts?

Results Summary

The review found that equity was predominantly measured through disparities in key learning outcomes by gender and socio-economic disadvantage. The review also found that inequity and inequality were often used interchangeably. Reducing disparities were best addressed through intensive cross sectoral interventions. The effectiveness and relevance of interventions for different groups of beneficiaries is highly dependent on context, but rigorous subgroup comparisons are not often feasible. USAID should provide ongoing training opportunities to education officers and implementing partners to better understand how to design, implement, monitor and evaluate activities through an equity lens.

- *Equity, defined broadly, is best addressed through cross-sectoral interventions. Equity was most often defined and measured by how vulnerable or at-risk a child was. Vulnerabilities were broadly defined and included geographic location, IDP status, gender, whether a child came from a broken home, whether a child was an orphan or came from a marginalized group, disability status, sexual orientation, and SES.*
- *The vulnerabilities most commonly addressed in evaluations reviewed under Goal 3 included gender and marginalization.*
- *CBE activities had a significant impact on reducing inequities in access to education, as they focus on providing education to learners who cannot attend classes at mainstream schools for a variety of reasons.*
- *In crisis and conflict settings, CBE is a particularly suitable approach for providing access to female learners where travel to and from schools can pose threats to their safety.*

A. Equity and Vulnerabilities: An Overview

- *In the evaluations reviewed under Goal 3, equity was usually measured by how vulnerable or at-risk a child was. The synthesis recognized that interventions have differential impacts on students with respect to gender, age, type of school, or even type of teacher. Nonetheless, articulating those impacts was a challenge, and the majority of the learning questions guiding this synthesis could not be answered.*

Vulnerabilities included a range of factors such as geographic location, IDP status, gender, whether a child came from a broken home, whether a child was an orphan or came from a minority ethnic/tribal group, disability status, sexual orientation, and SES. The vulnerabilities most commonly addressed in the evaluations reviewed under Goal 3 included gender and marginalization.

Within the evaluations reviewed under Goal 3, issues of equity were most often measured in relation to gender and categories of marginalization (i.e., SCOPSO, NEI in Nigeria). Marginalization varies across the studies but can include ethnic/tribal background, socio-economic disadvantage, or area of

residence/locality. Frequently, evaluations include questions for key informants on the degree to which activities take into account women and other marginalized groups (e.g., CBSI). Programmatic recommendations emerging from these evaluations also frequently suggest that USAID should engage communities and local governance units to support and manage activities targeting education access for OVCs, support implementing partners to analyze access of marginalized subgroups or populations to education activities, and create or strengthen channels that would facilitate increased access of these groups to quality education.

B. Case Studies Demonstrating Effective Interventions Targeting Equity

Evaluations of activities such as CARSI and SDPP point to the value of intensive cross-sectoral interventions in addressing such vulnerabilities. Such studies also point to the importance of detailed analysis strategies (i.e., subgroup analyses) to help USAID (and implementing partners) better understand the differential impacts of an intervention on a range of beneficiaries.

As an example of an intervention that addressed equity issues, **CARSI** targeted children “of broken homes” who had been exposed to violence in either the home or the surrounding community, and provided counseling services to these children. The activity is unique in that it addresses the needs of at-risk youth using a cross-sectoral approach, wherein elements of education and workforce development, economic growth and employment, public health, and governance interventions are employed. The regional evaluation demonstrated that CARSI effectively worked with school leadership (particularly in El Salvador) to reduce or eliminate home suspension policies so that at-risk children would be able to attend school without interruption. This in turn lessened the amount of time these children spent at home or in the community where they would be exposed to violence and instability.

There were also reported decreases in the number of children and youth joining gangs. The evaluation also found that sensitization of school leadership and teachers could help flag domestic violence or abuse scenarios that were negatively impacting students’ attendance and performance in school, for male and female students. For example, in rural areas sampled in the CARSI evaluation, incest was a frequent cause of teenage pregnancy among girl students. School directors and teachers were trained to detect signs of abuse in the home, in the hope of creating safer learning environments where female students felt comfortable to divulge threats to their welfare and well-being. The activity also worked closely with local authorities to provide allowances to young teenage mothers to continue to attend school.

The evaluations of **the SDPP interventions** (discussed in greater detail later in this report) also demonstrate how evaluations can address equity through rigorous analysis and reporting strategies. SDPP interventions were not designed to have different impacts on different subgroups of students (other than students at risk of dropout), conducted substantial subgroup analyses. Subgroup comparisons help evaluators and implementers to better understand whether there are differential impacts on students with certain characteristics (e.g., female versus male students, older versus younger students), for different types of schools (e.g., schools with a higher percentage of at-risk students, rural schools), or for different types of teachers (e.g., teachers with greater versus less training and professional experience). For example, the SDPP evaluations found that dropout, promotion, survival, and transition rates for learners varied most significantly by gender across the primary and secondary cycles, and by geography (proxy for rural/urban locality). By better understanding the differential rates at which an intervention may impact different target groups, USAID and implementing partners can enhance the relevance and effectiveness of an activity. Indeed, this was the approach taken with SDPP. Such detailed quantitative analyses were not found in many evaluations reviewed under Goal 3, but were conducted mostly in technical evaluation reports.

6. Social and Emotional Learning

Guiding Questions

For whom are the EiCC activities most effective? Under what conditions are these activities most effective? What further research is needed to establish effective delivery mechanisms for strengthening SEL in crisis and conflict contexts in terms of self-awareness (e.g. identifying emotions, accurate self-perception, recognizing strengths, self-confidence, self-efficacy), self-management (e.g. impulse control, stress management, self-discipline, self-motivation, goal-setting, organizational skills), social awareness (e.g. perspective taking, empathy, appreciating diversity, respect for others), relationship skills (e.g. communication, social engagement, team work, relationship building), and responsible decision making (e.g., identifying problems, analyzing situations, solving problems, evaluating, reflecting, ethical responsibility)?

- *The SEL approach has shown substantial evidence for positive outcomes in high-income countries. SEL has also been shown to contribute to improved academic performance and success later in life (academic attainment, crime prevention, employment, etc.).*
- *The International Rescue Committee (IRC) and New York University (NYU) have produced the only rigorous evidence for SEL in a fragile state (DRC) which demonstrated mixed results on SEL outcomes and some inconsistent gains in students' reading and math performance.*

A. SEL in Crisis and Conflict Settings: An Overview

- *SEL has high anecdotal value among beneficiaries in crisis-affected contexts. In an emergency, education serves as a protective mechanism against exploitation and harm and promotes the well-being of learners, teachers, and other education personnel. SEL builds social and emotional skills and contributes to a safe, protective learning environment.*

In crisis and conflict contexts, education serves as a protective mechanism against exploitation and harm and promotes the well-being of learners, their families, and teachers. SEL fosters academic performance and students' personal development by reinforcing social and emotional skills and establishing a safe, protective learning environment. There is substantial evidence that the SEL approach results in positive outcomes in high-income countries. IRC and NYU have produced the only rigorous evidence on SEL in a fragile state (DRC): an impact evaluation that demonstrated mixed results on SEL outcomes and some inconsistent gains in students' reading and math performance.

The quality of the relationships that students have with their classmates and teachers matters because it influences achievement and behavior by affecting the learning environment's quality. Qualitative evidence on teachers in United Nations Relief and Works Agency for Palestine Refugees (UNRWA) schools highlights the role that instructors play in facilitating a positive atmosphere beneficial to pupils. Husein

Results Summary

The review found that evidence from crisis-affected settings remains limited, without conclusive findings on students' improved performance or personal development. Further research on the effectiveness of SEL activities in crisis and conflict contexts is needed to understand which interventions are most successful in enhancing self-awareness, self-management, social awareness, relationship skills and responsible decision making.

Abdul-Hamid and others at the World Bank interviewed 74 students from UNRWA schools to understand how they performed well academically despite disadvantageous conditions. The students credit their teachers who are also Palestinian refugees. That shared identity means the teachers understand the challenges their students confront at home and school. The students claim their teachers provide solid academic instruction with an eye toward socio-emotional development. This socio-emotional aspect of the pedagogy encourages students to put more effort into their studies.¹²⁸ Abdul-Hamid describes the students' views:

“Empathy, respect, and trust were common themes in the students’ descriptions of effective teachers and hard teachers. Students provided numerous examples of teachers providing them with personal guidance, acting as confidants, and helping them to address particularly difficult moments in their lives.”¹²⁹

The Palestinian students feel protected by their teachers. The sense of protection has not occurred by accident. Abdul-Hamid and others cite UNRWA for its commitment to protection and socio-emotional values in its conception of quality education.¹³⁰ While not explicitly stated by the World Bank, UNRWA uses a SEL approach in its schools.

B. Case Studies Demonstrating the Effectiveness of SEL in Crisis and Conflict Settings

Essentially, SEL has a dual concentration in how it molds students. Joseph Durlak and others clarify that SEL strives to augment academic performance and students' personal development by building social and emotional skills and establishing a safe, protective learning environment.¹³¹ The SEL approach has shown substantial evidence for positive outcomes in high-income countries: In a systematic review of 213 SEL activities, the “SEL participants demonstrated significantly improved social and emotional skills, attitudes, behavior, and academic performance that reflected an 11-percentile-point gain in achievement”¹³² relative to control groups. However, the evidence from crisis-affected contexts remains limited.

IRC and NYU have produced the only rigorous evidence for SEL in a fragile state in evaluating the **OPEQ** activity in the DRC. IRC implemented its Learning in a Healing Classroom (LHC) curriculum for 480,000 students in Katanga, North Kivu, and South Kivu. The curriculum used a SEL model that heavily emphasized math and reading while providing teachers with in-service training and one-on-one mentoring. The impact evaluation analyzed how the LHC curriculum and its teacher training changed student outcomes for reading, mathematics, perceptions about teachers and schools, and experiences with mental health issues and bullying. Using two cohorts, the results presented findings for a Katanga cohort over two years against a control group. During the second year, the Katanga cohort receiving the treatment was expanded to additional schools. In contrast, the Kivu cohort was only evaluated over one year against its control group.

Indicators for SEL well-being were based on measurements of students' perceptions of their schools and teachers being caring and supportive (using two survey instruments), as well measurements of their perceptions of schools as cooperative and predictable environments (using a tool developed by NYU). Relative to its control group after one year, the Katanga cohort had statistically significant improved perceptions of caring and supportive schools and teachers, but also statistically significant worsened perceptions of cooperative and predictable school environments. There were no significant differences

¹²⁸ Abdul-Hamid et al., *Learning in the Face of Adversity: The UNRWA Education Program for Palestine Refugees*, 37-39.

¹²⁹ Ibid, 39.

¹³⁰ Ibid, 41-42.

¹³¹ Durlak et al., “The Impact of Enhancing Students' Social and Emotional Learning,” 407.

¹³² Ibid, 405.

after two years for the Katanga cohort. The Kivu cohort demonstrated statistically significant improved perceptions of both the caring and supportive schools and teachers and the cooperative and predictable school environments.¹³³ In addition, the evaluation assessed internal threats to student safety (as categorized by USAID’s Safer Learning Environment framework): physical and emotional bullying of students by peers, with a scale for victimization, aggression, and social skills. A mental health questionnaire gathered data from students for conduct and emotional issues. For both cohorts, there were no significant results.¹³⁴

The SEL results expose a complex reality. The LHC curriculum could have failed to produce consistent, positive improvements for SEL due to how well teachers taught the daily lessons (teachers only received weekly lesson plans). Also, the scale of violence and trauma that students face in their homes and communities might have counteracted the influence of LHC on SEL. IRC contends that the SEL component of LHC should become more explicit and research should determine avenues for targeting mental health and bullying more effectively.¹³⁵

7. School Dropout Prevention

Note About This Theme

The SDPP and its evaluations do not fit neatly into any USAID Education Goal areas, but, as instructed by the Office of Education for the sake of this particular study, they were included in Goal 3 given the linkages between dropout prevention and Goal 3’s priority to increase learner access to education. This is because learning from these methodologically rigorous and well-executed studies (in four diverse countries) can inform the design and implementation of interventions to increase access in conflict and crisis settings. Outside the scope of this study, SDPP is not formally categorized as a Goal 3 activity.

Results Summary

The review found varying success in student outcomes, (i.e., attendance, dropout rates and grade promotion rates), increasing teaching capacity (effectiveness in dropout prevention, teachers’ sense of responsibility and self-efficacy), and school administration capacity. Activities focused on academic and social support, combined with additional enrichment activities for at-risk students as well as changes in teacher practices to improve student attitudes and behaviors, translating into increased student engagement and reduced school dropout rates.

- *School dropout prevention activities require a comprehensive, contextualized suite of interventions to significantly increase student attendance and grade promotion, and to reduce dropout rates.*

A. SDPP: An Overview

- *USAID’s SDPP intervention model included academic and social support for students, combined with additional enrichment activities for at-risk students and changes in teacher practices, to improve student attitudes and behavior.*

SDPP is a multicountry (Cambodia, India, Tajikistan, and Timor-Leste) program designed to reduce student dropout rates and better understand successful means of reducing dropout rates in primary and

¹³³ Aber et al., *Final Report on the Impact of the OPEQ Intervention in the Democratic Republic of Congo*, 27.

¹³⁴ Ibid, 28.

¹³⁵ The DRC evaluation also measured changes in students’ reading and math performance, using the EGRA and Early Grade Mathematics Assessment (EGMA), respectively. While the treatment group demonstrated higher reading and math scores after one year of participating in the activity, statistically significant differences were not found after the second year. The evaluation report hypothesized that the loss of significance after two years signals that the intervention became less effective for reading as IRC expanded the activity.

secondary school. The intervention focuses on academic and social support, combined with additional enrichment activities for at-risk students and changes in teacher practices to improve student attitudes and behavior, translating into increased student engagement and reduced school dropout rates. These dropout prevention activities targeted a range of grades (i.e., seventh through ninth grades in Cambodia, fifth grade in India, ninth grade in Tajikistan, and fourth through sixth grades in Timor-Leste) and provided different enrichment activities (i.e., computer classes in Cambodia; language arts and crafts in India; after-school tutoring in Tajikistan; and songs, games, and crafts in Timor-Leste) targeted toward the key beneficiary group.

B. SDPP Evaluation Results

- *Shifts in student attitudes and behaviors translated into increased student engagement and ultimately reduced school dropout rates. SDPP demonstrated varying success in increasing teaching capacity (defined as teachers' effectiveness in dropout prevention, sense of responsibility, and self-efficacy), school administration capacity, and student outcomes (i.e., attendance, dropout rates, and grade promotion).*

SDPP outcome results measured program completion; they also included indicators on teaching capacity, school administration capacity, and at-risk student performance/dropout status. In measuring teaching capacity, effectiveness of teacher dropout prevention practices was found to have made a statistically significant positive impact in Cambodia and Timor-Leste, but no impact in India or Tajikistan. The evaluation report indicated this may have been due to higher teacher dropout prevention practice baseline scores or self-reporting bias.

Teacher self-efficacy was found to be statistically significant in Cambodia and Timor-Leste but had no impact in India or Tajikistan. The indicator on teachers' sense of responsibility had the greatest variation in outcomes. In Cambodia, SDPP demonstrated a statistically significant improvement in teachers' sense of responsibility, but it showed no impact in Timor-Leste. It also showed no impact in Tajikistan, but the evaluation report suggested that this was due to high levels at baseline. Interestingly, the SDPP in India showed a statistically significant negative impact on teachers' sense of responsibility. It would be valuable to unpack these findings in future studies.

Similarly, the intervention had mixed outcomes for changes in school administrator dropout prevention practices, with only half the countries (Cambodia and Tajikistan) illustrating statistically significant positive effects. There were no effects on school administrator self-efficacy across all four country interventions. There were mixed results for at-risk student performance indicators such as at-risk students' emotional, cognitive, and behavioral attitudes toward school. The interventions did yield consistently positive effects on students' perception of parental and teacher support, except in Timor-Leste.

The SDPP interventions had mixed effects on student attendance, student dropout rates, and grade progression for at-risk students and students overall across the intervention countries.

8. Crosscutting Themes

Researched Themes	Results Summary
<ul style="list-style-type: none">• Gender• Disability• ICT• Innovative finance• Scaling and sustainability	<p>The majority of the evaluations reviewed under Goal 3 analyzed results with a gender lens but were not designed with a gender-responsive theory of change or results framework. Few evaluations made reference to people with disabilities, while innovative financing was mentioned to the extent of linking it to concerns about sustainability and future programming. A substantial number of evaluations indicated the use of information and communications technology (ICT) to increase beneficiary engagement. The majority of evaluations that mentioned sustainability discussed barriers to sustainability and scalability due to the unstable, low-resource regions and at-risk target population addressed by the evaluated activities.</p>

A. Gender¹³⁶

- *Gender is not yet a key influencer in the design, implementation, and evaluation of education interventions. Most activities evaluated under Goal 3 analyzed results with a gender lens, though most did not explicitly design the interventions with a gender-responsive theory of change or results framework.*
- *Understanding how gender interacts with different education outcomes, such as learners' safety, school enrollment and attendance, and student achievement, is key to enhancing intervention results.*

Gender is often a key contributing factor toward learners' access to education, including their enrollment, attendance, grade promotion, dropout rates, and achievement. Understanding how and why gender plays a role in learners' educational experiences in crisis and conflict settings, and the cultural, social, security, and economic factors that may underpin these relationships, is key to enhancing intervention results. While the majority of activities evaluated under Goal 3 disaggregated evaluation results by gender, most interventions were not designed explicitly with a gender-responsive theory of change or results framework. This aligns with findings from other studies that have found that USAID should provide education officers and implementing partners with ongoing training on how to design, implement, monitor, and evaluate gender-sensitive education activities¹³⁷

The majority of the evaluation reports reviewed did not specify whether the activity was designed with a gender-responsive theory of change or results framework. However, all evaluations reviewed under Goal 3 included some level of data disaggregation by gender, except for the Indonesia OVC evaluation. For example, the evaluation of the Northern Education Initiative in Nigeria found that a higher percentage of female teachers positively influenced the educational experiences of female learners, which in turn led to higher performance of female learners (compared with male learners). According to the evaluation

¹³⁶ Guiding questions include: Are evaluations adequately examining gender dynamics? To what degree do evaluations find differential results for male versus female beneficiaries, in terms of both activity access and outcomes? What can be learned from evaluations about the relationships between gender and livelihood and advancement opportunities? What contextual and implementation factors help explain gendered outcomes?

¹³⁷ Burde et al., *What Works to Promote Children's Educational Access, Quality of Learning, and Wellbeing in Crisis-Affected Contexts*, vii.

report, these positive outcomes also contributed toward increased community/social support for girls' education.

In terms of what can be learned from the relationship between school safety (and safety in general) and gendered outcomes (particularly related to gender-based violence affecting school enrollment, attendance, retention, and learning outcomes for girls), the body of evidence from the Goal 3 synthesis shows that schools can serve both as a protective safe haven from external violence in the community and as a threat to student safety and a prime place of recruitment for gangs and violent extremism groups. There were many ways this was illustrated. Overall, the assessments illustrated many ways schools serve as different threats for boys and girls; predominantly boys were recruited as soldiers for gangs or radical causes, whereas girls were subjected to various forms of gender-based violence, including sexual assault and social discrimination for early-age pregnancy.

For example, the Education Emergency Support Activity (EESA) Rapid Risk Assessment in Mali illustrated that the community greatly feared the boy students would be recruited to become child soldiers and young girls would be kidnapped and raped, forced into early marriage, abused sexually, or made to wear veils.¹³⁸ A different example lies in the CARSI Honduras assessment, which found that youth vulnerability to crime and violence is mostly driven by “lack of adequate education, high levels of under or unemployment, high levels of domestic violence (youths being victims of physical, sexual, and psychological violence at home or subject to social discrimination at school by their peers).”¹³⁹ In such circumstances, education is also often viewed as a protective environment, sheltering youth (both boys and girls) against a range of negative outcomes.

B. Disability¹⁴⁰

- *Disability, defined broadly, was rarely examined in the reviewed evaluations of Goal 3 interventions. Disability was typically mentioned, if at all, as one risk factor under a broader umbrella of vulnerabilities related to addressing equity through the intervention. Only 5 of the 28 evaluations reviewed under Goal 3 made any reference to disabilities.*
- *Increased access to education for students with disabilities was often due to a multipronged approach at local and national levels to raise awareness about the rights and needs of children with disabilities among parents, teachers, and communities and ensure an enabling education policy environment.*
- *USAID should conduct a systematic and in-depth review of the degree to which education activities implement inclusive education approaches.*

In the evaluations reviewed under Goal 3, disability is often described as a risk factor under a broader umbrella of vulnerabilities related to equity. Five of the 28 evaluations reviewed under Goal 3 made any reference to disabilities. Only one, the Indonesia OVC intervention, directly focused on a range of disabilities (hearing impairment, learning, intellectual, and multiple). Other evaluations, such as OPEQ and Ghana Public Works, sought to improve access to education for students with physical disabilities through school infrastructure activities, like building disability ramps.

¹³⁸ CERIPS, *Education Emergency Support Activity (EESA) USAID Project Rapid Education Risk Assessment (RERA)*, 8.

¹³⁹ Berk-Seligson et al., *Impact Evaluation: Honduras Country Report*, 84.

¹⁴⁰ Guiding questions include: To what degree do activities address populations with disabilities and their needs? How is disability treated in evaluations, included in the evaluation questions? Included as a sampling stratum? Is there a separate data collection module? Discussed separately in findings/conclusions/recommendations?

Indonesia OVC was the key evaluation that focused on disabilities. This intervention resulted in 300 percent increased enrollment of children with vision impairment and promoted a policy shift toward vision impaired children’s accessing schooling in the first phase of the activity. In the second phase, the activity trained 241 teachers and supported the development and passage of 22 policies. It also established two satellite early intervention and resource centers for people with disabilities in Jakarta and piloted a preservice university program to strengthen the capacity of new teachers to teach children with visual impairment.¹⁴¹

The evaluation report indicated these achievements were due to a multipronged approach tackling program and local/national policy to raise awareness about the rights and needs of children with disabilities. Increased access to schooling for students with disabilities was attributed to a reported increase in parents’ awareness of the type of disability their child has, what their child needs, and what rights their child has; an ongoing program of professional development and training for teachers of students with disabilities; and provincial and national campaigns to strengthen the education policies supporting students with disabilities.

Interestingly, the performance evaluation of the CLP Education Program indicated that CLP “accidentally” supported children with disabilities. This intervention is a well-tested example of an activity targeting equity in terms of both gender and disability. Over 1,000 schools were rehabilitated in high-visibility urban and post-conflict locations—in about 100 of these schools, the infrastructural improvements included improved sanitary facilities for girls and physical access for children with disabilities.

Part of the evaluation included focus groups with educators who noted that younger siblings who participated in the YEGRA intervention would also help their older siblings to learn how to read: “even though not intended for children with disabilities, YEGRA has been valuable in helping these children to engage in learning.”¹⁴² Community support for students with disabilities was further enhanced through strengthened Father-Mother Councils, again pointing to the critical role that families play in ensuring children’s access to school in crisis and conflict settings.

C. Information and Communications Technology¹⁴³

- *The impact of ICT on learner outcomes is inconsistent. Although the perception of ICT among education stakeholders is mostly positive, actual ICT impact on learner outcomes (e.g., academic achievement, engagement in classroom instruction) varies widely from positive to negative to no impact at all.*
- *Fourteen of the 28 evaluations reviewed under Goal 3 made some mention of ICT in the activity design, but the evaluation reports contained few details on the effectiveness or impact of ICT on access to education and other outcomes.*
- *USAID should systematically evaluate the value added of ICT to improve access to education in crisis and conflict settings. In particular, the usefulness of ICT-based approaches versus face-to-face trainings and interventions (a learning question for this synthesis that could not be answered due to lack of evidence) should be examined.*

¹⁴¹ GRM International, *Evaluation of the Opportunities for Vulnerable Children Program Indonesia*, 11.

¹⁴² Gurevich et al., *Performance Evaluation of the Education Program of the Community Livelihoods Project (CLP) Final Report*, 1.

¹⁴³ Guiding questions include: To what degree do activities utilize ICT in intervention modalities, and what are the main uses? Is there any evidence of the effectiveness of ICT-based approaches versus face-to-face training and interventions?

The number of reliable and methodologically rigorous studies that have been conducted on the impacts of ICT in educational settings within developing countries is small.¹⁴⁴ Fourteen of the 28 evaluations reviewed under Goal 3 made some mention of ICT in the activity design. Most of the evaluations that referenced ICT stated that ICT was a means to increase student engagement with school and promote school retention (youth demand for ICT was high in activities such as CARSI), but the value added of such approaches was not explicitly measured. The Honduras A Ganar evaluation noted that computer labs were provided for students to use, but the evaluation report did not describe the effectiveness or value added of this activity. The Afghanistan BELT evaluation noted the potential of ICT to increase the efficiency and timeliness of educational material and textbook distribution and recommended that USAID explore ways to use ICT in future education activities.

Other interventions were particularly creative in their use of ICT, most notably Youth Engagement to Promote Stability (YEPS) in Timor-Leste, which used a radio station activity as a mechanism to transform how youth engage with the national government and community leaders to promote peace and reconciliation. New seasons of two radio programs targeting youth, the news program “Babadok Rebenta!” and the radio drama “Karau Dikur ba Dame” were developed in partnership with targeted youth and then broadcast on community radio stations in all thirteen districts.¹⁴⁵ Over 24 episodes were produced and broadcast throughout Timor-Leste on issues identified by and relevant to youth themselves. Education-focused issues included students’ dropping out of school, unemployment, suicide and emotional well-being, the role youth can play in governance and decentralization, and awareness of disabilities.

Some challenges emerged, including unreliable radio signals and inappropriate timing of the program. However, the evaluation suggested that listeners had positive responses to these shows, and that the shows were effective in changing attitudes of community members, and in some instances changing behaviors of community members as a result of increased information/knowledge on a topic. The evaluation did not assess whether these radio activities had any impact on education outcomes such as access to school.

D. Innovative Financing Mechanisms¹⁴⁶

- *Innovative financing mechanisms are key to ensuring education activity sustainability. Eleven of the 28 evaluations reviewed under Goal 3 mentioned innovative financing to some extent, usually as a recommendation for future consideration to enhance activity sustainability.*
- *USAID should support rigorous and systematic evaluations of public-private partnerships to refurbish or enhance school infrastructure, the different level of education inputs that*

¹⁴⁴ In a systematic review on the impact of ICT on key educational outcomes in developing country contexts, Tolani, McCormac, and Zimmerman (2009) found that impacts of ICT on learner outcomes vary, whether positive, negative, or no impact at all. The perception of ICT impacts among stakeholders, however, is mostly positive, but whether ICT can meet these expectations is dependent upon how such solutions are implemented. Extant evaluations using qualitative methodologies often rely on self-reported data, without validation or triangulation across multiple sources. This can lead to an inflation, or positive bias, of the effectiveness of ICT in educational settings. Further, the impact of ICT is also dependent upon exogenous factors, such as teacher training and support, classroom management techniques, and support from school leadership. Successful school integration depends heavily on effective and integrated leadership at the school, and at regional and national levels; support systems across sectors (including professional development, infrastructure maintenance, etc.); and curricular content that is relevant to needs and interests of teachers and students.

¹⁴⁵ Stein, Kelsi, *Youth Engagement to Promote Stability (YEPS) Final Evaluation*, 4.

¹⁴⁶ Guiding questions include: To what degree have the use of private versus public monies been included as part of project design? What have been any innovative financing schemes, if any, and what have the results been?

are necessary, and the cost-effectiveness of such partnerships in conflict and crisis settings.

To foster and maintain activity sustainability, there is a growing incentive that activity design should include some form of innovative financing mechanisms. Eleven of the 28 evaluations reviewed under Goal 3 mentioned innovative financing to some extent, usually as a recommendation for future consideration, to enhance activity sustainability. For example, the LETS Jordan evaluation recommended cost sharing contributions in the activity design so that for every dollar the school receives, a percentage of that amount must be raised locally by the school or community, which could include labor or volunteer time.¹⁴⁷ The CBSI evaluation (focused on access to justice services) recommended income generation strategies, which included fee-based mediation for the private sector and access to justice models in which some users contribute toward costs as a form of cost sharing.¹⁴⁸

Only two reviewed evaluations (Macedonia IIEP and Senegal EBD) noted that activities implemented some form of cost sharing. In Macedonia IIEP, the cost-sharing requirement was at the municipality level—each municipality was required to cost share at least 10 percent of the school renovation costs, either monetarily or in kind. In the end, the cost sharing totaled approximately 20 percent of the overall activity budget. The evaluation found that cost sharing distributed ownership of the activity in a way that contributed toward sustainability. It also provided “tangible examples of the benefits of ... cooperation to stakeholders both in and outside of the renovated schools.”¹⁴⁹

The Senegal EBD evaluation found that public-private partnerships were an effective mechanism for raising funds at a national level toward improving access to high-quality education for all children. These partnerships were also designed to strengthen the connections between the education system and companies, such that learners could graduate with the skill sets and knowledge that Senegalese companies prioritized.¹⁵⁰

The public-private partnership component demonstrated strong positive results, as it:

- Raised between \$1 million and \$1.5 million in cash and in-kind contributions.
- Put in place a high-level task force chaired by the director of USAID to work on the creation of a foundation. It was composed of a number of private sector corporations such as CNES and CNP, multilateral agencies such as the World Bank, foundations, banks, the Ministries of Education and Finance, and USAID.
- Conducted a benchmarking workshop in the U.S. for eight members of this task force to see how to ensure the involvement of the private sector in education and to learn from best practices in this respect.
- Obtained signatures of over 20 partnership protocols with the private sector (Sonatel Foundation, Ecobank, Next, NSIA, and Satrec, among others), solidifying the connections between the education and private sectors.
- Set up school visits with partner business and obtained “pledges of support” at a regional level.

Public-private partnerships as a means of improving education service delivery, learners’ access to education, and learning outcomes, particularly in conflict and crisis settings, are a promising and necessary avenue for USAID to explore. USAID should support rigorous and systematic evaluations of public-

¹⁴⁷ Moubayed et al., *Final Performance Evaluation USAID/Jordan Learning Environment Technical Support Program*, 24.

¹⁴⁸ Social Impact, *Assessment of the Caribbean Basin Security Initiative*, 46.

¹⁴⁹ Naskova et al., *Midterm Evaluation of USAID/Macedonia’s Interethnic Integration in Education Project*, 23.

¹⁵⁰ Ngom, Diagne, and Fall, *USAID Basic Education project Mid-Term Evaluation Final Report*, xii.

private partnerships to refurbish or enhance school infrastructure, the different levels of education inputs that are necessary, and the cost-effectiveness of such partnerships in conflict and crisis settings.

E. Scaling and Sustainability¹⁵¹

- *Scaling and sustainability of education interventions are often blocked due to financial constraints. Most evaluations reviewed under Goal 3 made some mention of possible barriers to activity sustainability given the contexts in which the activities were implemented (unstable, low-resource areas, targeting high-risk children and youth).*
- *The predominant barrier facing Goal 3 education activities is financial sustainability, followed by lack of government and community ownership of the intervention.*
- *USAID should consider longer-term evaluations that visit activity sites and beneficiaries one or more years after the activity has concluded to assess effectiveness of sustainability mechanisms and activities.*

Activity sustainability encompasses several factors including financial, organization, and technical components. Most evaluations reviewed under Goal 3 made some mention of possible barriers to activity sustainability given the contexts in which the activities were implemented (unstable, low-resource areas, targeting high-risk children and youth). The predominant barrier to activity scale-up and sustainability is financial sustainability.

The second most common constraint to activity sustainability is government and community ownership of the intervention, as illustrated in the Ghana Public Works evaluation. The evaluation report noted that government officials and local beneficiaries were involved in the ongoing ownership and maintenance of the constructed facilities. However, observations made during the evaluation also indicated that while certain structures were built, there is “no reference to the quality, durability or sustainability of the works completed.” Further, there was no explicit linkage between the constructed structures and the construction standards/codes (within either the U.S. or Ghana).¹⁵² This suggests the need for more meaningful involvement and oversight by communities and government agencies at local and national levels to ensure the quality and sustainability of school construction efforts. This could, in turn, influence learners’ continued access to learning environments and the safety of such environments.

Other activities, like the Nigeria RARA, addressed sustainability through building local technical skill capacity. For example, the activity formed a Reading Technical Working Group that built the capacity of state education officials through materials development, training on effective literacy instruction through mentoring by master trainers, and creating tangible action plans to improve the scale-up of reading initiatives in the states. Such activities also enhanced connections among government actors.¹⁵³

Similarly, Macedonia IIEP strove for sustainability through a multipronged approach, targeting the education system at a local and national level. IIEP included activities to train and engage teachers and worked across sectors to develop institutional frameworks and policies to support IIE sustainability.

¹⁵¹ Guiding questions include: How many of the evaluated activities have been scaled and what have been the results of the scaling processes? With which kinds of partners? How effective have they been at scale? How sustainable? Do evaluations of activities that are promising but have not been scaled make recommendations concerning scaling? What patterns exist if any in these recommendations on future scaling? To what extent do the activities focus on sustainability? How do they define it and measure it? To what extent is sustainability a focus of evaluations? How do they define it and measure it? What are the findings and recommendations relating to sustainability?

¹⁵² Rodriguez et al., *Performance Evaluation of Public Works Construction Activities to Increase Access to Education in Ghana*, 91.

¹⁵³ RTI International, *Nigeria Reading and Access Research Activity (RARA): Results of an Approach to Improve Early Grade Reading in Hausa in Bauchi and Sokoto States*, 107-108.

However, as documented in the evaluation report, activities were often conducted in parallel with other formal curricular and extracurricular activities.¹⁵⁴ This led to challenges and inefficiencies in achieving sustainability and in assessing the effectiveness of efforts to enhance activity sustainability.

Other evaluations discussed the need to better integrate sustainability architecture into future activity design and implementation. As an example of best practice, the evaluation of NEI in Nigeria discussed how the activity team planned to use “lessons learned to better to strategize and fine-tune the activity for greater impact in the remaining life of the activity.” Such an approach can also inform USAID, the Government of Nigeria, and implementing partners in the design of follow-on activities that are more relevant and effective to beneficiaries.¹⁵⁵ This suggests the importance of evaluation timing and timely dissemination of results to implementing partners, particularly in crisis and conflict settings where the nature of beneficiary needs may rapidly shift.

¹⁵⁴ Naskova et al., *Midterm Evaluation of USAID/Macedonia’s Interethnic Integration in Education Project*, ix.

¹⁵⁵ Larcom, et al., *Northern Education Initiative (NEI) Project Mid-Term Performance Evaluation Final Report*, 3.

RECOMMENDATIONS

Goal I

Ensure that pre- and post-test learning outcomes data are available from all activities and results are presented in a consistent manner across reports. For reading outcomes, data from the reviewed evaluation reports were difficult to synthesize due to a lack of complete reading data. For example, Djibouti Projet AIDE had a baseline but no midline or endline. Nigeria NEI initially planned to measure learning outcomes but experienced delays and then the activity was not able to collect data. Jamaica ETP had some issues with baseline reading data collection and was unable to follow up with corrections. There was also substantial variation in terms of how the reviewed evaluations reported their overall results. Philippines Basa and Rwanda L3 provided complete results with raw scores, zero scores, and performance categories, while Mozambique ApaL and Nigeria RARA reported zero scores but no performance category percentages. Some activities reported effect sizes while others only provided statistical significance testing.

Conduct further analyses of learning outcomes, most of which can only be carried out through secondary analyses of the data sets. While the goal of the synthesis was to examine results from evaluation reports, some additional analyses took place through the Secondary Analysis and Results Tracking activity, which reported on results achieved by country and for the entire Agency. However, these results, while useful, are limited to descriptive analyses involving ORF. Much more could be done to address topics such as benchmarking and target setting, annual yearly growth, effect sizes, zero scores, performance by gender, fluency in different languages, reading comprehension, correlations between subtasks, and internal consistency reliability. Some of this work has taken place on individual activities, but meta-analyses across multiple activities would be valuable for drawing lessons learned and improving programming in a variety of contexts.

Conduct further studies of differences in reading scores between boys and girls, progress by girls in closing gaps with boys over time, and the reasons behind these issues. Quantitative studies across activities could include more in-depth analyses of issues such as generally higher reading scores for boys compared to girls, narrowing of gaps between boys and girls (such as on Mozambique ApaL), and instances of girls outperforming boys. Quantitative studies could include issues such as bias in textbooks, preference by teachers for boys over girls, and differences between boys and girls on use of time outside of school. Girls' education activities would comprise an important part of these studies.

Study and prioritize avoiding production delays and distribution problems with teaching and learning materials, especially textbooks. The extent of the problem of production delays across all activities could be studied, with themes drawn out on the extent of the problem and the reasons for why it happens. For example, it would be expected that activities that need to produce textbooks in multiple languages—such as Ethiopia READ II—would have more production delays than those producing textbooks for a single national or regional language, but this might or might not be the case. Similarly, there appears to be a need to study why the distribution of books is generally more efficient in some countries compared to others. A more detailed examination could be undertaken of how these delays and problems have affected teacher training and classroom instruction, particularly since many interventions are time-sensitive and the duration of the activity support for a set of schools is sometimes only two years.

Continue with strong emphasis on in-service teacher training workshops but strive for greater application of best practices in terms of duration, frequency, and follow-up. For the activities that had success with the workshops, such as Indonesia PRIORITAS, as well as those activities

that did not have strong success, several issues need to be studied so that lessons learned can be applied to other activities. These issues could include impacts from different doses and durations of training workshops, and the extent of follow-up after the workshops in the schools. There is a general lack of evidence on whether the concepts and skills in the workshops were adequately learned by participants, whether the new techniques were applied in the field, and whether the teachers continued to apply these methods after activity support ended. Other factors related to these issues, such as support from the districts and within the schools, should be studied in depth.

Promote the design and implementation of district-level teacher coaching and school-level mentoring models by building on the limited evidence of success and expanding their implementation to most activities. Coaching and mentoring models, while not often implemented, have had success in a limited number of activities, such as in Kyrgyz Republic QRP, which offers an example of success for other activities. In some cases, the coaching and mentoring models were not implemented because in part because of the activity design. In other cases, the models were not implemented with fidelity due to issues such as over-reliance on cascade training, lack of coordination at the relevant ministry, and activity levels for overseeing the coaches and mentors, overburdened ministry specialists without enough time to conduct field visits, and inadequate funding for logistical necessities such as transportation. Perhaps the most commonly cited problem with these initiatives was a lack of follow-up training and support for the coaches and mentors.

Sharpen community reading initiatives through greater clarity in design and capitalize on evidence of successful implementation models—including training activities and post-training follow-up—and evaluations. In addition to the community engagement issues, there were also problems with unclear objectives and lack of a clear relationship—or an indirect relationship—with children’s reading. However, a few of the initiatives—such as Philippines Basa, and Malawi EGRA—promoted community engagement in reading and provided evidence of reading outside of the school day, and improved student learning. More in-depth studies of these activities, with lessons learned, could provide information needed so that the models could be attempted and/or replicated in other countries. Quantitative data provided by relevant ministry and activity staff on numbers of participants affected, materials used, training sessions organized, and visits to the communities would be helpful from a programming perspective. Similarly, it would be important to gather quantitative information to examine the relationship between community reading initiatives and student learning.

Make greater efforts to support ministries of education in building M&E systems and coordinating with activity M&E systems so that useful indicators are developed, tracked, and used. In some extreme cases, such as Liberia LTTP 2, the activity team made efforts to boost the Ministry of Education’s M&E system, which was nonfunctional. In the case of other activities—such as on Mozambique ApaL and Rwanda L3—ministry M&E systems existed, but the activities’ efforts did not result in building the capacity of ministry counterparts to use data collected in the field. Without relevant and up-to-date information collection and processing by the government, the possibilities to undertake evidence-based policymaking and planning are very limited. One of the issues cited in the evaluation reports reviewed was logistics, not unlike the problem with district-level teacher coaching. It should be possible to study the development and implementation of government M&E systems and discuss ways of improving them so that activities are monitored and information is used.

Increase the focus on activities in the crosscutting areas of disability, innovative finance, and information technology. Almost any additional analyses would add new information to the body of evidence on crosscutting issues, except gender (cited above) which has benefitted from several analyses. Some activities have made efforts, such as Indonesia PRIORITAS, to help increase coverage and the quality of services for students with disabilities. Similarly, Rwanda L3 had an IT component involving interactive radio instruction, but these initiatives have not had proper evaluations. The Tanzania Bridge-IT

activity was retrofitted into a reading activity before it reached fruition, though perhaps there are some lessons learned from the IT-related work that took place. Much more could be done, especially through qualitative methods, to identify the reasons behind the relative lack of programming, and the quality of the initiatives, in these areas.

Give more attention to issues of capacity-building, sustainability, and scaling up in evaluations, including post-evaluation studies, and build on validated methods to improve performance. Unanswered questions include the degree to which capacity building has led to sustainability, or have activities not focused enough on capacity-building so that initiatives will be sustainable after the end of the activity. With scaling up, there is almost no evidence that activities used an internationally validated method (such as MSI's *Scaling Up Management Framework* and *Scaling Up Toolkit*). Additional studies on scaling up across activities—using internationally tested and validated tools as a reference point—would likely provide insights on whether activities are including a scaling up framework from the beginning of the activity. These studies would also help identify where there is room for improvement so that successful initiatives can be designed with scaling in mind, and whether those initiatives can then be expanded for broader coverage and benefit to populations.

Goal 2

Workforce Development

Continue to experiment with innovative ways of linking entrepreneurial skills development and livelihood support to workforce development programming, while taking account of beneficiary characteristics and market dynamics. Employers are increasingly requiring workers to be more flexible and proactive—two key characteristics of entrepreneurship education. At the same time, young people will need to be proactive about creating their own livelihoods in changing 21st century labor markets. USAID should continue to explore ways of integrating entrepreneurship development into technical/vocational training, as well as offering flexible, on-demand technical training to young entrepreneurs. An example of a new context for accessing youth in need of these supports are the youth *bunges* (associations) launched by the Kenya Yes Youth Can Project.

Conduct more strategic analyses of assets and deficiencies within country workforce development systems to shape activity design. Workforce development activities could be strengthened by developing more explicit and specific theories of change about systems-level changes that a USAID activity is seeking to effectuate. These changes could involve improved capacity of strategically-selected workforce institutions, service providers, and policy-making entities at the country level to be able to achieve particular goals. The changes could also focus on improving systems-level relationships, practices, and coordination among key actors in a country. However, regardless of which systems change is sought, the activity should explicitly define it and the intervention measures selected to effect it, and why these measures. Currently, most reviewed evaluations reveal vague and diffuse effects on country workforce development systems, such as “more private sector engagement,” without diagnosis as to what has been blocking that goal to begin with and how the activity will change that.

Examine the links between employment, gender, and violence prevention to more meaningfully integrate workforce development programming into youth violence prevention and CVE programming. Research by Mercy Corps¹⁵⁶ and others¹⁵⁷ has shown that there

¹⁵⁶ Mercy Corps, *Youth & Consequences: Unemployment, Injustice, and Violence.*; Mercy Corps, *Motivations and Empty Promises: Voices of Former Boko Haram Combatants and Nigerian Youth.*

¹⁵⁷ International Labour Organization, UN Peacebuilding Support Office, UN Development Programme, and World Bank. *Employment Programmes and Peace: A Joint Statement on an Analytical Framework, Emerging Principles for Action and Next Steps.*

is not a one-to-one match between youth employment/livelihoods and participation in violent extremism or gangs or in holding peaceful attitudes and beliefs. Therefore, as part of each violence prevention and CVE activity design, USAID should attempt to elucidate the relationship between youth economic opportunity, skills and character development, and the propensity to participate in violence. In many cases, this understanding will need to be developed in the course of activity implementation, and can and should be built into activity design and adaptive learning efforts. Success in workforce outcomes should be tracked against measures of peaceful attitudes and behavior towards developing holistic activity theories of change.

Require comparative benchmarks against which employment and earnings outcome results can be interpreted and evaluated. All workforce activities should be required to track participants' labor market and/or self-employment outcomes, and these data should no longer be permitted to be presented without comparative baselines. Baselines can be any number of valid comparisons such as country-level employment rates for similar populations (if reliable), or employment rates of comparable activities for the population, or for a control group. Pre- and post-activity comparisons for employment are less meaningful because presumably all participants take part in a workforce development activity only if they were unemployed (or significantly underemployed), and even in the absence of the activities they would have had some level of success. Employment and self-employment data sources can be diverse (e.g., they can include consumption or asset accumulation as a proxy for earnings), but the important factors are that they are consistent and include a valid comparison.

Support implementers to improve reporting of sex-disaggregated enrollment and program completion data, and utilize this data to ensure that interventions are optimally matched to the ability of beneficiaries. An activity need not be rigid nor lengthy, but interventions or combinations of interventions should be well-defined with clear completion standards against which outcomes should be tracked. Evaluations should report on enrollment, completion, and outcomes, and provide explanatory judgments for the observed results.

Develop workforce development activities that respond directly and substantively to the most urgent gender-based challenges to labor market success. Most activities will likely focus on improving women's economic empowerment, but in some country contexts young men face greater or different challenges such as social stigma. Regardless, the gendered focus of USAID programming should be strategic, intentional, made explicit, and rigorously tested.

Higher Education

Improve the relevance of higher education institutions to country development efforts by expanding results monitoring of their extension services with end users. Extension services may include research, training, and/or tool development or piloting with community constituencies, community-based organizations and the private sector. USAID should foster incentives to strengthen the feedback loop between university researchers and users of their research to build the responsiveness of the higher education system to development needs.

Sharpen activity theories of change in conflict-affected countries that show how higher education investments can contribute to gender equality, stabilization and peace-building. There is great potential for the strengthening of higher education systems to contribute to peace building, such as through reducing marginalization of population subgroups with more equitable access to higher education, fostering community peace dialogues, and conducting evidence-based research into the specific drivers of conflict in countries.

Develop and utilize measures for tracking improved quality of core higher education services including teaching/learning, student services, and research. Although most, if not all,

higher education activities are focused on improving quality, there are very few, if any, objective measures used in the evaluations reviewed for this synthesis that enable USAID to determine success. Although lessons may be drawn from K-12 education experience, the higher education context is sufficiently different, especially in its research role, that unique measures are needed. The employability of youth graduates is an important metric and its use should be expanded.

Expand comprehensive efforts to enhance women’s higher education leadership in strategic fields such as agricultural sciences and technology. USAID has conducted at least two innovative activities linking secondary-level girls’ recruitment to higher education programs through community engagement and support to female faculty. This systematic, gender-sensitive approach is strategic and comprehensive and therefore worth testing more rigorously for possible replication.

Improve the sustainability of higher education quality improvement efforts through increased participation of research end users. Community and private-sector partnerships should be included in research grant making to enhance the sustainability of these efforts.

Goal 3

Strengthen institutional capacity at local and national levels to ensure improved access to basic education. Enhancing the capacities of individuals (teachers, school administrators, subnational and national officials) and institutions are commonly used approaches to improve access to education, and education infrastructure and rehabilitation. USAID should provide continuous training to Education Officers and implementing partners on its Human and Institutional Capacity Development Handbook, including the Behavior Engineering Model for effective capacity strengthening. This guidance can then be more explicitly integrated during activity design, and better reflected in activity evaluation plans.

Tailor violence prevention and CVE activities to the unique and diverse needs of target populations. While this recommendation seems self-evident, there is not a one-size-fits-all solution to violence prevention and CVE. Indeed, the Goal 3 evaluations reviewed for this synthesis demonstrate that the most effective interventions were those uniquely responsive to a range of contexts and cross-sectoral in nature. The evaluations also demonstrated the critical and positive role families, churches, and schools play in ensuring at-risk children and youth stay in school, engage in socially positive activities, and decrease their likelihood of participating in criminal activity. Investigating the connections between access to education and differences in education service delivery (e.g., madrassa versus secular) and CVE outcomes should also continue to be a research priority for Goal 3.

Prioritize situational analyses to assess demand, feasibility, and sustainability for education activities in emergency and conflict-affected settings. The situational analyses this study reviewed clearly demonstrate the value of rapid assessments, which relied on mixed methods to capture a range of perspectives with enough detail to influence strategy and implementation. USAID should prioritize ongoing, lower-cost assessments to inform activity design and management and to augment more traditional evaluation approaches. Scopes of work for future Goal 3 evaluations should require assessments of: (a) flexibility, agility, and procurement speed; (b) intensive program management; (c) innovation, evaluation, and learning; and (d) informed risk-taking, among other variables.

Design school dropout prevention activities as a comprehensive, contextualized suite of interventions to significantly increase student attendance and grade promotion, and to reduce dropout. While not explicitly categorized under Goal 3, lessons learned about effective approaches to reduce school drop out in the SDPP studies can inform activities designed to increase access to schools in crisis and conflict settings. SDPP demonstrated varying success in improving teacher capacities and student education outcomes such as attendance, dropout and grade promotion. Each

dropout prevention activity was complex and uniquely designed to the needs of target populations in India, Cambodia, Timor-Leste and Tajikistan. USAID should continue to support interventions grounded in rigorous evidence-based reviews on effective solutions for complex problems such as student dropout for male and female learners in crisis and conflict settings.

Maintain safer learning environments for children and youth through community engagement and mobilization, curricular improvements and infrastructure improvements.

The evaluations reviewed in this synthesis demonstrate that in conflict and crisis contexts, students' physical and emotional well-being is impacted by many threats within and outside of the school, and male and female students are affected differently by such threats. However, effective solutions to these threats do exist—such as engaging community members to create safe transportation solutions for students and teachers and to manage school rehabilitation efforts, and developing conflict-sensitive curricula. Future studies should investigate whether dissipating cultural and ethnic biases (which contribute to ethnic tension within schools) in turn contribute (directly or indirectly) to safer learning environments in crisis and conflict settings.

Address equity through cross-sectoral interventions. Evaluations of activities such as the CARSII and SDPP interventions point to the value of intensive cross-sectoral interventions in addressing equity issues. USAID should continue to emphasize and provide guidance on the importance of detailed analysis strategies (i.e., subgroup analyses) to help Education Officers and implementing partners better understand the differential impacts of an intervention on a range of beneficiaries. For example, evaluators should identify and define dimensions of equity (based upon USAID's definition and the activity's theory of change) and the composition of the population within the country context. Next, disparities in key education and other outcomes between individual children and youth, as well as broader groupings of children (controlling for observable student characteristics) within the population, should be analyzed. The analysis should include distributions at aggregated levels, such as school, community, district, region and national. If an impact evaluation is being conducted, impacts should be estimated overall and for each subgroup.

Conduct further research on SEL and its effective integration into education interventions.

Currently, SEL has high anecdotal value among beneficiaries. Previous research demonstrating the value of SEL exists primarily in high-income countries. IRC and New York University have produced the only rigorous evidence for SEL in a fragile state (DRC). USAID should prioritize further research on how SEL can be used to enhance children's well-being in crisis and conflict contexts.

Emphasize the role of gender in the design, implementation, and evaluation of education interventions.

Understanding how gender interacts with different education outcomes such as learners' safety, school enrollment and attendance, and student achievement and empowerment is key to enhancing intervention results. The evaluations reviewed under Goal 3 did not demonstrate whether activities were designed with a gender-responsive theory of change in mind. USAID should support implementing partners and independent evaluation teams to understand and integrate Agency guidance on gender-sensitive theories of change.

Increase focus on disability (defined broadly) for Goal 3 interventions and their evaluations.

Only five of the 28 evaluations reviewed under Goal 3 made any reference to disabilities. Only one, the Indonesia OVC evaluation, directly focused on disabilities (specifically hearing impairment and learning, intellectual, and multiple disabilities). Others, like the DRC OPEQ and Ghana Public Works interventions, indirectly addressed equity in access to education for students with disabilities through infrastructure work like building disability ramps. USAID should conduct a systematic and in-depth review of the degree to which education activities implement inclusive education approaches in crisis and conflict settings. USAID should further support implementing partners and independent evaluation teams to understand

and integrate Agency guidance on addressing the needs of students with disabilities in conflict and crisis contexts.

Leverage innovative financing mechanisms to enhance sustainability. In crisis and conflict contexts, USAID should collaborate with national governments, other donors, and implementing partners through the Education Cluster and Protection Cluster platforms to tackle sustainability through innovative technical and financial mechanisms. USAID should support rigorous and systematic evaluations of public-private partnerships that refurbish or enhance school infrastructure, including the different level of education inputs that are necessary, and the cost-effectiveness of such partnerships in conflict and crisis settings. USAID should **consider longer-term evaluations that involve visits to activity sites and with beneficiaries one or more years after the activity has concluded**, to assess effectiveness of sustainability mechanisms and activities.

Evaluate the value added of ICT as a means to improve access to education in crisis and conflict settings. In particular, the usefulness of ICT-based approaches versus face-to-face trainings and interventions should be examined (this was a learning question for this synthesis that could not be answered due to lack of evidence).

Employ a balance of standardized indicators developed from commonly agreed upon definitions of key topics, and customized indicators. These indicators should align with activity results frameworks and theories of change. Such indicators should also be realistic and suit the difficult and resource-constrained contexts in which Goal 3 activities are implemented.

Ensure that theories of change are clear about the unique value of specific activities toward individual outcomes so that such theories can be evaluated. USAID should provide ongoing professional development opportunities for USAID Education Officers and implementing partners focused on the aforementioned technical topics as well as processes to strengthen activity monitoring and evaluation systems and activities.

Cross-Goal Recommendations

Lack of Impact Evaluations

Increase the number of experimental and quasi-experimental impact evaluations to enable judgments about the effectiveness of specific interventions to inform investment decision-making. More of such evaluations are needed across the goals. For Goal 2 on workforce development, these evaluations could be used to inform choices about the length and intensity of training, and the most effective (and ideally cost-effective) combination of training, microenterprise support, and employment services to enable female and male youth to obtain good-quality employment. For higher education programs, evaluations could elucidate the kinds of faculty professional development efforts most effective for modernizing pedagogy and producing graduates with 21st century skills.

Lack of Clear Description of Key Information in Evaluation Reports

Request that evaluations consistently include clear descriptions of key information. This information includes descriptions of interventions implemented, including duration and intensity of treatment (by hours, over time), the characteristics of the beneficiaries, and the outputs and outcomes reported. Wherever possible, evaluations should provide comparative information to make judgments about activity effectiveness more transparent and evidence-based.

Lack of Cost-Effectiveness Analyses¹⁵⁸

Increase the number of cost-effectiveness studies to improve activity design decision-making and increase financial sustainability of USAID investments. The cost-effectiveness analysis is a critical dimension for judging the effectiveness of interventions. Furthermore, collection of cost data and consideration of cost in the country context during the activity design phase is vital to ensuring interventions can be sustained within the budgets of country governments or private sector implementers. Collection of cost data is not common in USAID activities, so to do so well will require planning, coordination and transparency with implementing partners, as well as with country stakeholders responsible for sustaining programmatic outcomes.

Gender

Conduct and deepen gender analysis prior to activity design to identify gender-based dynamics affecting education quality, access to education and training, and youth labor market outcomes after activity completion. Gender analysis should include understanding the different perspectives, life circumstances and outcomes for *both* girls/young women and boys/young men, and these dynamic relationships should be captured in the activity design. The analyses should focus on increasing equity, safety and empowerment while promoting gender equality and reducing gender gaps. For example, if young women are discouraged from accessing certain types of career training, changing attitudes of families and male peers could be part of the activity design alongside empowerment activities for the young women. Similarly, if materials and training modules have built-in gender biases, then activities should work with governments to design a process for making revisions, including piloting modified materials to ensure that gender biases have been removed prior to scaling up production and distribution. USAID should support implementing partners and independent evaluation teams to understand and integrate Agency guidance on gender-sensitive theories of change.

Disability

Expand efforts to create inclusive learning and work environments for people with disabilities. Two suggested approaches to implement this expansion could be: to require a component on improving access for populations of people with disabilities in very large education activities; and to create a small number of pilot activities around key disability categories that would bear experiences that could be replicated in larger activities. Further guidance on how to define and measure “special needs” in individual studies and additional evaluations that focus not only on specialized education (efforts that focus on a specific type of disability) but also on disability inclusive education (related to changes to the overall education system for the full spectrum of disabilities) could also enrich the body of evidence produced by USAID on education programming related to disability.

Information and Communications Technology

Further research on the value added of ICT is needed, as well as possible uses of technology in the field to gather information during monitoring. In this synthesis, the evaluation results showed that ICT-related solutions were used for instructional purposes, scaling of EMIS systems, employment services, education management, professional networking, agricultural research, and as a means to increase student engagement with school and promote school retention. However, a range of technical and design difficulties in getting the systems to work as intended were reported and

¹⁵⁸ For more information about evaluations that referenced cost-effectiveness analysis, see: Thomaz Alvares de Azevedo and Sean Kelly, *Supplemental Topics from the Synthesis of USAID-Funded: Evaluations: Education Sector, 2013–2016* (USAID, January 2018).

maintenance was cited as an ongoing issue with computers (although cellphone technology seems promising). The study team recommends conducting qualitative research to identify the technical issues hampering the adoption of ICT-related solutions, quantitative research to measure the impact and cost-effectiveness of these investments in comparison to alternatives such as face-to-face training, and further consideration of the usefulness of the selected technologies given the intervention modality and environment where it would be deployed during the activity design phase. More effort could also be put into examining IT applications and promoting technology that would make information collection and processing more efficient and effective for adaptive management.

Innovative Finance

Experiment with innovative finance that aims to build resources and sustainability for programming for disadvantaged populations. A starting point may be to conduct cost-effectiveness or unit cost analyses so that the level of finance for activities is known in advance and funds can be solicited using accurate data. Innovative financing schemes may be able to leverage badly needed private sector funds to complement public and donor funding, and incentivize service provider innovation. Social impact bonds, development impact bonds, results-based financing, as well as student loan schemes and employer pay-for-service agreements, may help unleash new funding sources for education activities. In crisis and conflict contexts, USAID could collaborate with national governments, other donors, and implementing partners through the Education Cluster and Protection Cluster platforms to discuss how to best tackle sustainability. Private sector partners could also be invited to discuss different finance mechanisms for testing. For entrepreneurship promotion, an increasingly important component for Goal 2 programming, USAID could augment financing for entrepreneurs through angel investing and crowd-sourcing. USAID's Development Credit Authority is available to reduce the risk of such schemes. Given the novelty of such approaches in the development field, the study team recommends targeted and well-designed tests of selected modalities.

CONCLUSION

At the beginning of this decade, USAID, Congress and the development community came together to address the need for a more strategic approach to development.^{159,160} As part of the Agency's response, a series of institutional reforms started in 2010 with the aim of focusing USAID's programming on clear, strategic goals. The USAID Education Strategy for 2011-2015, dubbed "Opportunity Through Learning," was a significant component of this redesign effort, resulted in substantial changes to education programming, and paved the way for other sectors to focus their programming.¹⁶¹

As the end of the 2011 Education Strategy period approaches, it was a propitious time to take stock of what has been learned. However, information about activities mentioned in this study is limited to the information included in the evaluation reports and may not be exhaustive. The review of the 69 evaluations included in this synthesis produced three types of findings: (1) positive findings that can be used in future programming; (2) inconclusive findings for which further study is needed; and (3) negative findings from interventions that should most likely not be repeated. The strength of evidence from these evaluations varied, with different levels of quality as judged by evaluation principles. The number of evaluations associated with each Education Strategy Goal provided a medium-sized body of evidence. While most evaluations reviewed for this synthesis were conducted in sub-Saharan Africa, there was a lack of robust evaluations from regions such as Europe and Central Asia, the Middle East and North

¹⁵⁹ US Senate Hearing 111-806, "Statement of Senator Patrick J. Leahy."

¹⁶⁰ Unger, *The Shape of U.S. Global Development Reforms*.

¹⁶¹ Willis, *Agendas, Actions, and Accountability in International Development: A Case Study of the USAID Education Strategy for 2011-2015*.

Africa, and South Asia. Along with increasing the quality of evaluation reports, these kinds of voids in geographical areas should be addressed to ensure both quantity and consistency in coverage for evaluations of education activities.

Finally, while the Education Strategy started in 2011, the Implementation Guidance took over a year to be finalized. In addition, the contracting mechanisms used to support Goal 1 (Assistance to Basic Education All Children Reading), Goal 2 (YouthPower) and Goal 3 (Assistance to Basic Education: Access for All) were only awarded in 2014–2015, and implementation will continue until 2019. Thus, findings and lessons learned related to the Education Strategy will continue to unfold.

ANNEX I: STATEMENT OF WORK

Education Evaluation Syntheses – Goals 1 and 3

1. Activity Description

Building on recent efforts to synthesize what is being learned from evaluations that USAID commissions,¹⁶² the Education Office in the Bureau for Economic Growth, Education, and Environment (E3/ED) is commissioning syntheses of evaluation findings related to the three Goals in the USAID Education Strategy. Products developed under this activity will address topics of interest to E3/ED and the Agency’s education officers worldwide related to Goal 1 “Improved reading skills for 100 million children in primary grades” and Goal 3 “Increased equitable access to education in crisis and conflict environments for 15 million learners.”

2. Existing Information Sources

E3/ED already has an inventory of recent education sector evaluations produced by the Bureau or by overseas Missions. Older evaluations, should the Bureau decide to examine a longer time period, can be accessed through the Agency’s Development Experience Clearinghouse (DEC). Annual Performance Plan and Report (PPR) documents may be useful for determining whether evaluations reported as having been completed in previous years are all available in the DEC.

3. Activity Purpose, Audiences, and Intended Uses

Purpose

E3/ED intends that the analytic products that result from this activity will support evidence-based decision making by ensuring that findings from sets of evaluations on topics of interest to the Office are accessible to USAID staff. E3/ED’s initial intent was to focus this activity on two topics related to Goal 2 of the Education Strategy: higher education and youth workforce development. Ensuing internal discussions led to an expansion of the scope to also include syntheses of evaluation findings on topics under Goals 1 and 3 under a common approach that could be applied across these three goals, and will be implemented across two mechanisms: the E3 Analytics and Evaluation Project and Reading and Access Evaluation. This activity will comprise two main phases. In phase 1, the quality of evaluation reports will be reviewed. In the phase 2, findings and lessons learned from a subset of evaluations that met quality standards identified in the first phase will be extracted and synthesized. It is expected that up to 80 evaluation reports published between 2013 and 2016 across all three Goals will be reviewed under phase 1, with only a subset of those reports included in phase 2.

Audiences

The primary audience for the products to be developed under this activity are E3/ED and Mission staff as well as implementing and country partner organizations that plan and deliver education and workforce development programs and related support services.

¹⁶² These efforts include the annual E3 Sectoral Synthesis of Evaluation Findings (https://www.usaid.gov/sites/default/files/documents/1865/E3_Sectoral_Synthesis_Report.pdf) and an evaluation synthesis from the Bureau for Food Security (BFS) that focuses on what has been learned from the Feed the Future initiative (<https://agrilinks.org/sites/default/files/resource/files/Final%20KDAD%20Evaluation%20Synthesis.pdf>).

Intended Uses

Two main reports are expected to be produced. The first report will be based on a standardized Evaluation Quality Protocol, and E3/ED will use its findings to determine topics on which it will develop additional guidance, products, and presentations to improve the quality of evidence generated for USAID-funded activities in the education sector. The second report will be based on a standardized Findings and Lessons Protocol, and E3/ED and Mission staff working the education sector may use the synthesized findings and lessons learned to inform future USAID education programming worldwide related to each of the three Education Strategy Goals. Work performed by Reading and Access Evaluation under this activity should focus on Goals 1 and 3.

4. Synthesis Topics

E3/ED will confirm the topics for which findings and lessons learned will be extracted and synthesized. Tentative topics related to all three Education Strategy Goals are provided below:

Goal 1 – Early Grade Reading

- Topic 1: Teacher training (pre-service and in-service)
- Topic 2: Materials development, production, distribution, utilization
- Topic 3: Parent/community engagement/support/education/mobilization
- Topic 4: Systems/policy/government capacity strengthening

Goal 2 – Workforce Development

- Topic 1: Training
- Topic 2: Entrepreneurship
- Topic 3: Private sector involvement
- Topic 4: Systems/policy/government capacity strengthening
- Topic 5: Youth engagement

Goal 2 – Higher Education

- Topic 1: Training
- Topic 2: Private sector involvement
- Topic 3: Systems/policy/government capacity strengthening
- Topic 4: Youth engagement

Goal 3 – Education in Conflict Settings

- Topic 1: Training
- Topic 2: Parent/community engagement/support/education/mobilization
- Topic 3: Systems/policy/government capacity strengthening
- Topic 4: Direct service delivery

5. Gender and Disability Considerations

Participation in the education system and educational outcomes vary considerably across countries, and can be substantially affected by gender and disability status. Therefore, it is expected that the syntheses prepared under this activity will report education-related findings by gender and disability status, when such information is available in the reviewed evaluations.

6. Activity Tasks

Evaluation Quality Protocol

In initial discussions about this activity, E3/ED requested that the synthesis team develop a preliminary framework for an approach to assessing the quality of evaluations to be examined under this activity. The framework the synthesis team prepared highlighted several core principles for consideration, including:

- Be consistent with USAID Evaluation Policy;
- Not be biased in favor of any particular evaluation design type, as it is expected that impact evaluations, performance evaluations, and qualitative evaluations will be reviewed; and
- Be amenable to a heterogeneous set of evaluation questions, ranging from the effectiveness of project/activity to the project/activity implementation and sustainability to the continued relevance of Agency assistance where circumstances may have shifted.

Pursuant to these recommendations, E3/ED requested that the synthesis team develop and pilot test an Evaluation Quality Protocol, which it will then pilot test in collaboration with E3/ED and incorporate feedback as appropriate. The protocol may also be shared with external audiences, such as the Comparative and International Education Society (CIES) annual conference, for additional feedback. This protocol should be used by a team of expert reviewers to identify which evaluation reports will have findings and lessons extracted and included in the syntheses.

The criteria for inclusion of reports in phase 2 will be developed by the synthesis team in collaboration with E3/ED. Possible criteria include the strength of the conceptual framing, openness, and transparency; robustness of methodology; cultural appropriateness; and the validity, reliability, and cogency of the evaluation. Data collected for phase 1 should also be analyzed and the main findings of this phase summarized in a 10- to 15-page report assessing the quality of the evaluations by Education Strategy Goal. Work performed under this activity should focus on the review of evaluations that have Goal 1 and/or Goal 3 as their primary Education Strategy Goals. Results may also be disaggregated by geographic areas, pilots/scale-ups, whether conflict/crisis affected, and country income level.

Findings and Lessons Protocol

USAID experts in education subsectors will identify specific topics under each Education Strategy Goal about which findings and lessons learned will be synthesized (Section 4 provides a preliminary list of those topics). While the topics of the syntheses to be produced are expected to vary, E3/ED expects that the synthesis for each topic will be developed using a common outline, approach, and standards for documenting findings and lessons learned. Following initial discussions with E3/ED, the synthesis team will be expected to prepare a Findings and Lessons Protocol to extract findings (by gender and disability status when possible) and lessons learned (e.g., successes and challenges) about these topics from the evaluation reports. Existing E3 and BFS evaluation syntheses will be examined as potential templates, but final decision on the common outline, approach, and standards will be determined in collaboration with E3/ED. A Findings and Lessons Synthesis Report that addresses the agreed-upon topics under each Education Strategy Goal will be produced collaboratively by the two mechanisms. Work performed under this activity should focus on the extraction and synthesis of findings and lessons from evaluations that have Goals 1 and/or 3 as their primary Education Strategy Goals.

Team Selection and Training

Once E3/ED has selected the synthesis topics and approved the Evaluation Quality and the Findings and Lessons Protocols, the synthesis team will conduct training exercises for the reviewers. Different teams of reviewers might be used to apply each protocol. The selection and training of the reviewers for the Evaluation Quality Protocol should take into account the “Reviewing Evaluations for the Evaluation Synthesis Initiative” memorandum prepared by E3/ED, which suggests crowdsourcing the reviews in order to assess the quality of the evaluations while disseminating the evaluation quality criteria. The selection and training of reviewers for the Findings and Lessons Protocol should consider reviewers who are subject matter experts.

Implementation

Following E3/ED’s approval of the Evaluation Quality Protocol and associated training, the synthesis team – in collaboration with E3/ED – will develop a systematic process for the review of the USAID evaluation reports. This process may include efforts to publicize the framework and quality criteria with key partners in the broader education and evaluation community. The implementation of the Evaluation Quality Protocol will result in the selection of the evaluations that will be subjected to the Findings and Lessons Protocol, based on criteria to be agreed with E3/ED as well as a summary report about the quality of the evaluations by Education Strategy Goal.

Following E3/ED’s approval of the Findings and Lessons Protocol and associated training, the synthesis team – in collaboration with USAID staff in the topical fields on which the synthesis volumes will focus – will extract findings and other relevant data from topical sets of USAID evaluation reports produced between 2013 and 2016, and identify lessons for future programming, as relevant. This phase may also include the preparation of key findings and lessons summaries for each synthesis topic that would serve as a precursor to the preparation of a synthesis report and reviewed collaboratively by the synthesis team and USAID to highlight and prioritize the findings by topical area and identify any gaps in the summaries that may need to be addressed before a synthesis report is prepared.

Draft and Final Reports

The synthesis team will prepare drafts reports summarizing the main findings, conclusions, and recommendations from its implementation of the Evaluation Quality and Findings and Lessons Protocols. Based on USAID review and comments on such drafts, the synthesis team will prepare final versions for E3/ED’s approval.

Dissemination Plan and Implementation

E3/ED will prepare dissemination plans for the reports produced under this activity, with inputs from the synthesis team as required. Thus, the schedule and budget for this activity should include time and resources for the synthesis team’s involvement at the dissemination phase. The dissemination strategy should consider how the study products will be utilized by the identified audiences, and incorporate follow-up interviews as appropriate to determine and share actual instances of utilization.

7. Deliverables

A preliminary list of deliverables anticipated under this activity is provided below. The synthesis team, in consultation with E3/ED, will develop a Work Plan that will detail specific deliverables to be prepared under this activity with corresponding due dates. While products produced under this activity will focus on evaluations that have Goals 1 and/or 3 as their primary Education Strategy Goals, E3/ED may require that reports focusing on Goals 1, 2, and 3 be consolidated or summarized in one Evaluation Quality

Report and one Findings and Lessons Synthesis report to be prepared across the two implementing mechanisms.

1. Draft Activity Work Plan, including draft Evaluation Quality Protocol and draft Finding and Lessons Protocol
2. Final Activity Work Plan, including final Evaluation Quality Protocol and final Findings and Lessons Protocol
3. Draft Evaluation Quality Report, including draft dissemination plan
4. Final Evaluation Quality Report, including final dissemination plan and lessons learned about the evaluation quality review process and protocols
5. Draft Findings and Lessons Synthesis Report, including draft dissemination plan
6. Final Findings and Lessons Synthesis Report, including final dissemination plan and lessons learned about the findings and lessons review process and protocols

8. Team Composition

A research team led by Management Systems International (MSI) is expected to conduct this study across two mechanisms: the E3 Analytics and Evaluation Project, which is implemented by MSI in partnership with Development and Training Services and NORC at the University of Chicago; and the Reading and Access Evaluation project, which is implemented by NORC with MSI as a subcontractor. The review of evaluations and corresponding products related to Goal 2 will be conducted through the E3 Analytics and Evaluation Project while evaluations and corresponding products related to Goals 1 and 3 will be funded under the Reading and Access Evaluation project. Design, analysis, reporting, and dissemination efforts should be carried out across both mechanisms.

Separate Work Plans should be produced for the activities conducted under the E3 Analytics and Evaluation Project and the Reading and Access Evaluation project. These Work Plans should propose a team and organizational approach to managing this activity, for E3/ED review and approval. It is recommended that the teams include at a minimum an overall Team Leader, an Activity Coordinator, designated Goal Leads for Education Strategy Goals 1 and 3, as well as a sufficient number of mid- or senior-level Technical Advisors necessary to complete the tasks described in this SOW. It is expected that MSI will engage NORC at the University of Chicago to provide technical assistance and reviews of draft products at key points in this study.

9. USAID Participation

It is anticipated that E3/ED technical staff with expertise in the topics selected for examination under this activity will play an active role in developing the focus topics, reviewing study products, and developing lessons for future programming that will be incorporated into final syntheses volume(s). The exact nature of USAID staff participation will be further elaborated through discussions between E3/ED and the synthesis team, and may vary somewhat from topic to topic. In addition, through such discussions, E3/ED and the synthesis team will explore what roles implementing partners with which E3/ED collaborates may play in the topical areas to be covered.

10. Scheduling and Logistics

The tasks under this activity to be carried out by the Reading and Access Evaluation project will be completed between approximately July 2016 and December 2017, with the timeline for subsequent dissemination tasks to be discussed with E3/ED. In its Work Plan, the project team will propose a detailed schedule for implementation of the required tasks for E3/ED's approval.

11. Reporting Requirements

Reporting requirements will be finalized during discussions between E3/ED and the synthesis team concerning the synthesis topics, and will be incorporated into the final Work Plan.

12. Budget

The Reading and Access project team responding to this SOW will propose in its Work Plan an estimated budget to complete the tasks described in the Work Plan, for USAID's approval.

ANNEX 2: EVALUATION REPORTS REVIEWED BY ACTIVITY AND EDUCATION STRATEGY GOAL

Country	Activity Name	Goal	DEC Link
Dominican Republic	Effective Schools Program (ESP)	1	http://pdf.usaid.gov/pdf_docs/PA00KM95.pdf
Ethiopia	Improving Quality of Primary Education Program (IQPEP)	1	http://pdf.usaid.gov/pdf_docs/PA00K6KW.pdf
Ethiopia	Reading for Ethiopia's Achievement Developed Institutional Improvement (READ II)	1	http://pdf.usaid.gov/pdf_docs/PA00M563.pdf
Ghana	Partnership for Accountable Governance in Education (PAGE)	1	http://pdf.usaid.gov/pdf_docs/pbaaa020.pdf
Guatemala	Classroom Education Reform Project (REAULA)	1	http://pdf.usaid.gov/pdf_docs/PA00JP35.pdf
Indonesia	Prioritizing Reforms, Innovations and Opportunities for Reaching Indonesia's Teachers, Administrators, and Students (PRIORITAS)	1	http://pdf.usaid.gov/pdf_docs/PA00KJXV.pdf
Jamaica	Jamaica Education Transformation Project	1	http://pdf.usaid.gov/pdf_docs/pdacx310.pdf
Kyrgyz Republic	USAID Quality Reading Project (QRP) for the Kyrgyz Republic	1	http://pdf.usaid.gov/pdf_docs/PA00M399.pdf
Liberia	Liberia Teacher Training Program Phase II (LTTP II)	1	http://pdf.usaid.gov/pdf_docs/PA00JNC4.pdf
Malawi	Malawi Early Grade Reading Activity (MEGRA) [Performance Evaluation]	1	http://pdf.usaid.gov/pdf_docs/PA00KBNS.pdf
Malawi	Malawi Early Grade Reading Activity (MEGRA) [Impact Evaluation]	1	http://pdf.usaid.gov/pdf_docs/PA00KVBP.pdf
Mozambique	USAID/Aprender A Ler (ApaL) project	1	http://pdf.usaid.gov/pdf_docs/PA00M5D4.pdf
Philippines	Basa Pilipinas (Basa) [Impact Evaluation]	1	http://pdf.usaid.gov/pdf_docs/pa00mbm8.pdf
Philippines	Basa Pilipinas (Basa) [Performance Evaluation]	1	http://pdf.usaid.gov/pdf_docs/PA00MGD2.pdf
Philippines	Whole School Reading Program (WSRP)	1	http://pdf.usaid.gov/pdf_docs/PA00M37Q.pdf
Rwanda	Literacy, Language and Learning Initiative (L3)	1	http://pdf.usaid.gov/pdf_docs/PA00KZS9.pdf
Tanzania	Bridge-IT	1	http://pdf.usaid.gov/pdf_docs/PA00JSSH.pdf
Zambia	Time to Learn (TTL)	1	http://pdf.usaid.gov/pdf_docs/pa00kzlk.pdf
Afghanistan	Afghanistan Workforce Development Program (AWDP)	2	http://pdf.usaid.gov/pdf_docs/PA00K48W.pdf
Armenia	Junior Achievement of Armenia (JAA)	2	http://pdf.usaid.gov/pdf_docs/PA00TJH.pdf
Azerbaijan	Youth Business Leadership Project (YBLP)	2	http://pdf.usaid.gov/pdf_docs/PA00K9M6.pdf
Georgia	Georgia Economic Prosperity Initiative (EPI) project	2	http://pdf.usaid.gov/pdf_docs/pdacy472.pdf
Georgia	Georgia Education Management Project (EMP)	2	http://pdf.usaid.gov/pdf_docs/Pdacu911.pdf

Country	Activity Name	Goal	DEC Link
Honduras	Mejorando la Educacion para Trabajar, Aprender y Superarse (Proyecto METAS)	2	http://pdf.usaid.gov/pdf_docs/PA00K2H2.pdf
Indonesia	Indonesia University Partnerships (UP) program	2	http://pdf.usaid.gov/pdf_docs/PA00JRCZ.pdf
Iraq	USAID/Iraq Opportunities Project (Foras)	2	http://pdf.usaid.gov/pdf_docs/PA00KNDR.pdf
Jordan	Youth for the Future (Y4F)	2	http://pdf.usaid.gov/pdf_docs/PA00KN57.pdf
Kenya	Teacher Education and Professional Development (TEPD) project	2	http://pdf.usaid.gov/pdf_docs/pdacx751.pdf
Kenya	Yes Youth Can! (YYC)	2	http://pdf.usaid.gov/pdf_docs/PA00JZQX.pdf
Lebanon	University Scholarship Program (USP)	2	http://pdf.usaid.gov/pdf_docs/PA00KS5T.pdf
Liberia	Advancing Youth Project (AYP)	2	http://pdf.usaid.gov/pdf_docs/PA00M9N5.pdf
Morocco	Favorable Opportunities to reinforce Self-Advancement for Today's Youth (FORSATY)	2	http://pdf.usaid.gov/pdf_docs/PA00KZ43.pdf
Nepal	Education for Income Generation Program (EIG)	2	http://pdf.usaid.gov/pdf_docs/pbaaa002.pdf
Nicaragua	Education for Success (EFS)	2	http://pdf.usaid.gov/pdf_docs/PA00M5CN.pdf
Pakistan	Fulbright Student Program	2	http://pdf.usaid.gov/pdf_docs/PA00JTW5.pdf
Pakistan	Pakistan-United States Science & Technology Cooperation (S&T)	2	http://pdf.usaid.gov/pdf_docs/PA00K48G.pdf
Paraguay	Women's Leadership Program in Paraguay (WLPP)	2	http://pdf.usaid.gov/pdf_docs/PA00KNGF.pdf
Rwanda	Akazi Kanoze (AK) Youth Livelihoods Project	2	http://pdf.usaid.gov/pdf_docs/PA00MGDQ.pdf
Rwanda	Women's Leadership Program (WLP) Rwanda	2	http://pdf.usaid.gov/pdf_docs/PA00KPDP.pdf
Vietnam	Higher Engineering Education Alliance Program (HEEAP)	2	http://pdf.usaid.gov/pdf_docs/pdacx675.pdf
Afghanistan	Assessment of Learning Outcomes and Social Effects (ALSE) project	3	http://aalims.org/uploads/Burde_AALIMS.pdf
Afghanistan	Basic Education, Learning, and Training (BELT) project	3	http://pdf.usaid.gov/pdf_docs/PA00MB4J.pdf
Cambodia	School Dropout Prevention Pilot (SDPP) Project - Cambodia	3	http://schooldropoutprevention.com/wp-content/uploads/2016/01/Cambodia_Findings_Report-Volume_1.pdf
DRC	Opportunities for Equitable Access to Quality Basic Education (OPEQ_	3	https://www.rescue.org/sites/default/files/document/642/ed-opportunitiesforequitableaccesstoqualitybasiceducation.pdf
DRC	Youth and Alternative Education in North Kivu	3	https://eccnetwork.net/wp-content/uploads/1.27.17_DRCFinal.Links_.pdf
El Salvador	Rapid Education and Risk Analysis El Salvador	3	http://pdf.usaid.gov/pdf_docs/pa00mq56.pdf
El Salvador	U.S. Government's (USG) Central America Regional Security Initiative (CARSI)	3	http://pdf.usaid.gov/pdf_docs/pbaah621.pdf
Ghana	Public Works Construction Activities	3	http://pdf.usaid.gov/pdf_docs/PA00KH8Z.pdf
Guatemala	U.S. Government's (USG) Central America Regional Security Initiative (CARSI)	3	http://pdf.usaid.gov/pdf_docs/pbaah623.pdf

Country	Activity Name	Goal	DEC Link
Honduras	U.S. Government's (USG) Central America Regional Security Initiative (CARSI)	3	https://www.vanderbilt.edu/lapop/carsi/CARSI_Honduras_v1_Formatted_W_02.16.16.pdf
India	School Dropout Prevention Pilot (SDPP) Project - India	3	http://schooldropoutprevention.com/wp-content/uploads/2016/01/India_Findings_Report-Volume_1.pdf
Jordan	Learning Environment Technical Support (LETS)	3	http://pdf.usaid.gov/pdf_docs/PA00K1QB.pdf
Macedonia	Interethnic Integration in Education Project (IIEP)	3	http://pdf.usaid.gov/pdf_docs/PA00K15Q.pdf
Mali	Education Emergency Support Activity (EESA)	3	https://eccnetwork.net/wp-content/uploads/Mali-EESA-Rapid-Education-Risk-Analysis-2016.pdf
Mali	Education Recovery Support Act	3	https://eccnetwork.net/wp-content/uploads/RERA-Mali-Jan-2016.pdf
Nigeria	USAID Nigeria Education Crisis Response Project	3	http://pdf.usaid.gov/pdf_docs/PA00KNPC.pdf
Senegal	USAID Basic Education (USAID/EDB) project	3	http://pdf.usaid.gov/pdf_docs/pdacx672.pdf
Tajikistan	School Dropout Prevention Pilot (SDPP) Project - Tajikistan	3	http://schooldropoutprevention.com/wp-content/uploads/2016/01/Tajikistan_Findings_Report-Volume_1.pdf
Timor-Leste	School Dropout Prevention Pilot (SDPP) Project - Timor-Leste	3	http://pdf.usaid.gov/pdf_docs/PBAAD881.pdf
Timor-Leste	Youth Engagement to Promote Stability (YEPS)	3	http://pdf.usaid.gov/pdf_docs/PA00KZDJ.pdf
Yemen	Yemen Community Livelihoods Project (CLP)	3	http://pdf.usaid.gov/pdf_docs/PA00K5M7.pdf
Djibouti	Projet AIDE	1 & 2	http://pdf.usaid.gov/pdf_docs/pdacy251.pdf
Ethiopia	School-Community Partnership Serving Orphan and Vulnerable Children Affected by HIV/AIDS (SCOPSO)	1 & 3	http://pdf.usaid.gov/pdf_docs/pbaaa329.pdf
Indonesia	Opportunities for Vulnerable Children (OVC)	1 & 3	http://pdf.usaid.gov/pdf_docs/PA00JM2M.pdf
Nigeria	Nigeria Reading and Access Research Activity (RARA)	1 & 3	http://pdf.usaid.gov/pdf_docs/PA00KVMI.pdf
Nigeria	Northern Education Initiative (NEI)	1 & 3	http://pdf.usaid.gov/pdf_docs/pdacy473.pdf
Caribbean	The Caribbean Basin Security Initiative (CBSI)	2 & 3	http://pdf.usaid.gov/pdf_docs/PA00MHCT.pdf
East Africa	USAID Countering Violent Extremism (CVE) projects: Kenya Transition Initiative-Eastleigh (KTI-E); Kenya Garissa Youth Project (G-Youth); Somali Youth Livelihoods Program (SYLP)	2 & 3	http://pdf.usaid.gov/pdf_docs/pdacx479.pdf
Honduras	A Ganar	2 & 3	http://pdf.usaid.gov/pdf_docs/PA00KTBW.pdf

ANNEX 3: GOAL I SUPPORTING INFORMATION

TABLE 2: GOAL I INTERVENTION COMPONENTS

Country – Activity	Learning Outcomes	Support for Classroom Instruction	Teacher Training	Community Engagement	Policy and Systems
Djibouti - Projet AIDE	✓	✓	✓	✓	✓
Dominican - Republic ESP	✓	✓	✓	✓	✓
Ethiopia - IQPEP	✓	✓	✓		✓
Ethiopia - READ II	✓	✓	✓		
Ethiopia - SCOPSO				✓	✓
Ghana - PAGE	✓			✓	✓
Guatemala - REAULA		✓	✓	✓	✓
Indonesia - OVC			✓		✓
Indonesia - PRIORITAS	✓	✓	✓	✓	✓
Jamaica - ETP	✓	✓	✓	✓	
Kyrgyzstan - QRP	✓	✓	✓	✓	✓
Liberia - LTTP2			✓	✓	✓
Malawi - MEGRA	✓	✓	✓	✓	✓
Mozambique - ApaL	✓	✓	✓		✓
Nigeria - NEI				✓	✓
Nigeria - RARA	✓	✓	✓	✓	
Philippines - Basa	✓	✓	✓	✓	
Philippines - WSRP	✓	✓	✓		
Rwanda - L3	✓	✓	✓	✓	✓
Tanzania - Bridge-IT	✓	✓	✓		✓
Zambia - TTL		✓	✓	✓	✓

TABLE 3: LEARNING OUTCOMES

Country – Activity	Overall Results	Disaggregated Results	Subtask Results
Djibouti - Projet AIDE	✓	✓	✓
Dominican Republic - ESP	✓		
Ethiopia - IQPEP	✓	✓	✓
Ethiopia - READ II	✓		
Ethiopia - SCOPSO			
Ghana - PAGE	✓		✓
Guatemala - REAULA			
Indonesia - OVC			
Indonesia - PRIORITAS	✓	✓	✓
Jamaica - ETP	✓		✓
Kyrgyzstan - QRP	✓	✓	✓
Liberia - LTTP II			
Malawi - MEGRA	✓	✓	✓
Mozambique - ApaL	✓	✓	✓
Nigeria - NEI			
Nigeria - RARA	✓	✓	✓

Country – Activity	Overall Results	Disaggregated Results	Subtask Results
Philippines - Basa	✓	✓	✓
Philippines - WSRP	✓	✓	✓
Rwanda - L3	✓	✓	✓
Tanzania - Bridge-IT	✓		
Zambia - TTL			

TABLE 4: SUPPORT CLASSROOM INSTRUCTION

Country – Activity	Time on Task	Materials Distribution	Quality of Materials	Instructional Guides	Language Appropriate
Djibouti - Projet AIDE		✓	✓		
Dominican Republic - ESP		✓			
Ethiopia - IQPEP	✓	✓		✓	
Ethiopia - READ II		✓			
Ethiopia - SCOPSO					
Ghana - PAGE					
Guatemala - REAULA			✓		
Indonesia - OVC					
Indonesia - PRIORITAS		✓			
Jamaica - ETP	✓	✓	✓	✓	
Kyrgyzstan - QRP		✓	✓		
Liberia - LTTP2				✓	
Malawi - MEGRA	✓	✓	✓	✓	✓
Mozambique - ApaL	✓	✓	✓	✓	
Nigeria - NEI					
Nigeria - RARA	✓	✓	✓	✓	✓
Philippines - Basa	✓	✓	✓	✓	✓
Philippines - WSRP				✓	
Rwanda - L3		✓	✓	✓	✓
Tanzania - Bridge-IT				✓	
Zambia - TTL		✓	✓	✓	✓

TABLE 5: TEACHER TRAINING

Country – Activity	Pre-service	In-service	In-school Mentoring	District-Level Coaching	Pedagogy
Djibouti - Projet AIDE		✓	✓	✓	
Dominican Republic - ESP		✓	✓		
Ethiopia - IQPEP	✓	✓	✓	✓	✓
Ethiopia - READ II		✓			
Ethiopia - SCOPSO					
Ghana - PAGE					
Guatemala - REAULA	✓				
Indonesia - OVC	✓	✓			
Indonesia - PRIORITAS	✓	✓	✓		
Jamaica - ETP		✓	✓		
Kyrgyzstan - QRP		✓	✓	✓	
Liberia - LTTP2	✓	✓			

Country – Activity	Pre-service	In-service	In-school Mentoring	District-Level Coaching	Pedagogy
Malawi - MEGRA		✓	✓		✓
Mozambique - ApaL		✓	✓	✓	✓
Nigeria - NEI					
Nigeria - RARA		✓	✓	✓	✓
Philippines - Basa		✓	✓		
Philippines - WSRP		✓	✓		✓
Rwanda - L3	✓	✓	✓		✓
Tanzania - Bridge-IT		✓			✓
Zambia - TTL		✓	✓	✓	✓

TABLE 6: COMMUNITY ENGAGEMENT

Country – Activity	PTAs/SMCs	Parental Engagement	Extracurricular Activities	Information for Parents
Djibouti - Projet AIDE	✓			
Dominican Republic - ESP	✓	✓	✓	
Ethiopia - IQPEP				
Ethiopia - READ II				
Ethiopia - SCOPSO	✓			
Ghana - PAGE	✓		✓	✓
Guatemala - REAULA		✓		
Indonesia - OVC				
Indonesia - PRIORITAS	✓			
Jamaica - ETP	✓	✓		
Kyrgyzstan - QRP			✓	
Liberia - LTTP2	✓			✓
Malawi - MEGRA	✓	✓	✓	
Mozambique - ApaL				
Nigeria - NEI	✓			✓
Nigeria - RARA	✓	✓	✓	
Philippines - Basa	✓			
Philippines - WSRP				
Rwanda - L3	✓	✓		
Tanzania - Bridge-IT				
Zambia - TTL	✓	✓		

TABLE 7: POLICY AND SYSTEMS

Country – Activity	M&E Systems	Policy and Data	Capacity Building
Djibouti - Projet AIDE	✓	✓	✓
Dominican Republic - ESP	✓	✓	✓
Ethiopia - IQPEP	✓		✓
Ethiopia - READ II			
Ethiopia - SCOPSO			✓
Ghana - PAGE		✓	✓
Guatemala - REAULA		✓	✓
Indonesia - OVC			✓

Country – Activity	M&E Systems	Policy and Data	Capacity Building
Indonesia - PRIORITAS		✓	✓
Jamaica - ETP			
Kyrgyzstan - QRP	✓		✓
Liberia - LTTP2	✓	✓	✓
Malawi - MEGRA	✓	✓	✓
Mozambique - ApaL		✓	
Nigeria - NEI	✓		✓
Nigeria - RARA			
Philippines - Basa			
Philippines - WSRP			
Rwanda - L3	✓	✓	✓
Tanzania - Bridge-IT			✓
Zambia - TTL	✓		✓

TABLE 8: CROSSCUTTING

Country – Activity	Gender	Disability	Innovative Financing	Scaling and Sustainability	Information Technology
Djibouti - Projet AIDE	✓			✓	✓
Dominican Republic - ESP		✓		✓	
Ethiopia - IQPEP	✓				✓
Ethiopia - READ II				✓	
Ethiopia - SCOPSO					
Ghana - PAGE	✓			✓	
Guatemala - REAULA				✓	
Indonesia - OVC	✓	✓		✓	
Indonesia - PRIORITAS	✓	✓		✓	
Jamaica - ETP		✓		✓	
Kyrgyzstan - QRP	✓			✓	
Liberia - LTTP2				✓	✓
Malawi - MEGRA	✓	✓	✓	✓	
Mozambique - ApaL	✓			✓	
Nigeria - NEI	✓			✓	✓
Nigeria - RARA	✓				
Philippines - Basa	✓			✓	
Philippines - WSRP	✓				
Rwanda - L3	✓				✓
Tanzania - Bridge-IT				✓	✓
Zambia - TTL				✓	

ANNEX 4: GOAL 2 SUPPORTING INFORMATION

TABLE 9: TYPES OF WFD SERVICES

Country – Activity	Skills Training	Employment Services	Entrepreneurship Promotion	Institutional Change
Afghanistan - AWDP	✓	✓		✓
Armenia - JA	✓	✓	✓	✓
Azerbaijan - YBLP	✓	✓		✓
Djibouti - Projet AIDE	✓	✓		✓
Georgia - EPI	✓	✓	✓	✓
Honduras - Proyecto METAS	✓	✓		✓
Iraq - Foras	✓	✓	✓	✓
Jordan - Y4F	✓	✓		✓
Liberia - AYP	✓	✓		✓
Nepal - EIG	✓		✓	✓
Nicaragua - EFS	✓	✓		✓
Rwanda - AK	✓	✓	✓	✓

TABLE 10: WFD SKILLS TRAINING

Country – Activity	Technical Skills	Soft Skills or Life Skills	Workforce Readiness	Vocational Skills	Basic Skills
Afghanistan - AWDP	✓	✓		✓	
Armenia - JA		✓			
Azerbaijan - YBLP		✓	✓		
Djibouti - Projet AIDE	✓			✓	
Georgia - EIP				✓	
Honduras - Proyecto METAS					✓
Iraq - Foras	✓				
Jordan - Y4F	✓	✓	✓		
Liberia - AYP		✓	✓	✓	✓
Nepal - EIG		✓		✓	✓
Nicaragua - EFS		✓		✓	✓
Rwanda - AK		✓	✓	✓	✓

TABLE 11: WFD EMPLOYMENT SERVICES

Country – Activity	Internships	Mentoring	Apprenticeships	Job Placement	Career Services	Job Fairs	Networking
Afghanistan - AWDP				✓			
Armenia - JA		✓				✓	✓
Azerbaijan - YBLP	✓	✓					✓
Djibouti - Projet AIDE	✓			✓			
Georgia - EPI							✓
Honduras - Proyecto METAS	✓		✓	✓			
Iraq - Foras	✓		✓	✓		✓	
Jordan - Y4F					✓		
Liberia - AYP	✓						
Nepal - EIG							✓
Nicaragua - EFS				✓			
Rwanda - AK		✓					

TABLE 12: WFD ENTREPRENEURSHIP PROMOTION

Country – Activity	Entrepreneurships Skills Development	Business Coaching	Access to Finance
Afghanistan - AWDP			
Armenia - JA	✓	✓	✓
Azerbaijan - YBLP			
Djibouti - Projet AIDE			
Georgia - EPI	✓		✓
Honduras - Proyecto METAS			
Iraq - Foras	✓		✓
Jordan - Y4F			
Liberia - AYP			
Nepal - EIG	✓		✓
Nicaragua - EFS			
Rwanda - AK	✓	✓	

TABLE 13: WFD SYSTEMS AND INSTITUTIONAL CHANGE

Country – Activity	PPPs	Cost Share	Demand-Driven Curriculum	Industry Associations	Labor Market Assessment	WfD Authority	LMIS	Policy Reform	TVET Capacity Building	Professional Development
Afghanistan - AWDP	✓		✓		✓				✓	✓
Armenia - JA	✓									✓
Azerbaijan – YBLP	✓	✓		✓						
Djibouti - Projet AIDE	✓								✓	✓
Georgia - EPI	✓			✓	✓					✓
Honduras - Proyecto METAS	✓			✓						✓
Iraq - Foras	✓	✓	✓	✓	✓		✓			✓
Jordan - Y4F	✓		✓	✓						✓
Liberia - AYP	✓		✓	✓						
Nepal - EIG	✓		✓		✓					
Nicaragua - EFS	✓							✓		✓
Rwanda - AK ¹⁶³	✓		✓		✓	✓		✓	✓	

¹⁶³ Akazi Kanoze’s evaluation was a narrowly tailored randomized control trial that did not discuss its institutional change efforts. The MasterCard Foundation has detailed these results in its [Skills at Scale](#) report (2017).

TABLE 14: TYPES OF VIOLENCE PREVENTION AND CVE WFD* SERVICES

Country – Activity	Skills Training	Employment Services	Entrepreneurship Promotion	Youth Civic Engagement	Other Peace-Building Activities ¹⁶⁴
Caribbean - CBSI [Dominican Republic - Alerta Joven]	✓	✓	✓		
Caribbean - CBSI [Guyana - SKYE]	✓	✓	✓		
Caribbean - CBSI [Jamaica - Empowering Jamaica's Youth (JA)]	✓		✓		
East Africa - CVE [Kenya - KTI-E]		✓	✓	✓	✓
East Africa - CVE [Kenya - G-Youth]	✓	✓		✓	
East Africa - CVE [Somalia - SYLP]	✓	✓	✓		
Honduras - A Ganar	✓	✓	✓	✓	
Kenya - YYC			✓	✓	✓
Morocco - FORSATY	✓	✓		✓	

TABLE 15: VIOLENCE PREVENTION AND CVE WFD SKILLS TRAINING

Country – Activity	Technical	Soft Skills or Life Skills	Workforce Readiness	Vocational	Basic Skills
Caribbean - CBSI [Dominican Republic - Alerta Joven]				✓	✓
Caribbean - CBSI [Guyana - SKYE]		✓	✓		
Caribbean - CBSI [Jamaica - Empowering Jamaica's Youth (JA)]			✓		
East Africa - CVE projects [Kenya - KTI-E]					
East Africa - CVE projects [Kenya - G-Youth]	✓		✓		
East Africa - CVE projects [Somalia - SYLP]	✓			✓	
Honduras - A Ganar	✓	✓	✓	✓	
Kenya - YYC					
Morocco - FORSATY		✓	✓	✓	✓

¹⁶⁴ Other peace-building activities include public debates on extremism, social messaging, interfaith dialogue, voter registration.

TABLE 16: VIOLENCE PREVENTION AND CVE WFD ENTREPRENEURSHIP PROMOTION

Country – Activity	Entrepreneurships Skills Development	Access to Finance
Caribbean - CBSI [Dominican Republic - Alerta Joven]		
Caribbean - CBSI [Guyana - SKYE]	✓	
Caribbean - CBSI [Jamaica - Empowering Jamaica's Youth (JA)]	✓	
East Africa - CVE [Kenya - KTI-E]	✓	✓
East Africa - CVE [Kenya - G-Youth]	✓	
East Africa - CVE [Somalia - SYLP]	✓	
Honduras – A Ganar	✓	✓
Kenya – YYC	✓	✓
Morocco - FORSATY		

TABLE 17: VIOLENCE PREVENTION AND CVE WFD EMPLOYMENT SERVICES

Country – Activity	Internships	Mentoring	Apprenticeships	Job Placement	Career Services
Caribbean - CBSI [Dominican Republic - Alerta Joven]				✓	
Caribbean - CBSI [Guyana - SKYE]		✓			
Caribbean - CBSI [Jamaica - Empowering Jamaica's Youth (JA)]					
East Africa - CVE [Kenya - KTI-E]					
East Africa - CVE [Kenya - G-Youth]	✓				
East Africa - CVE [Somalia - SYLP]	✓		✓	✓	
Honduras – A Ganar	✓	✓		✓	
Kenya – YYC					
Morocco - FORSATY	✓			✓	✓

TABLE 18: HIGHER EDUCATION ACTIVITIES

Country – Activity	Faculty Training	Scholarships	Institutional Capacity Building	Research Capacity Building	Partnerships	External Training
Georgia - EMP	✓		✓		✓	✓
Indonesia - UP	✓		✓	✓	✓	
Kenya - TEPP	✓		✓			
Lebanon - USP		✓				
Pakistan - Fulbright Student Program		✓				
Pakistan - S&T				✓		
Paraguay - WLPP	✓	✓	✓		✓	✓
Rwanda - WLP	✓		✓	✓	✓	✓
Vietnam - HEEAP	✓		✓			

TABLE 19: GOAL 2 RANGE OF ENROLLMENT AND COMPLETIONS FOR ACTIVITIES

Country – Activity	Enrollment			Completion		
	Total	Men	Women	Total	Men	Women
Afghanistan - AWDP				9,022	5,864	3,158
Armenia - JA	Not monitored					
Azerbaijan - YBLP	105	50	55			
Djibouti - Projet AIDE				311	60	251
Georgia - EPI	Not monitored					
Honduras - Proyecto METAS	35000			8140		
Iraq - Foras	175,000			50,000	33,500	16,500
Jordan - Y4F				3,873		
Liberia - AYP	22,256	9,348	12,908	3,103	1,117	1,986
Nepal - EIG	74,917	58,435	45,580			
Nicaragua - EFS						
Rwanda - AK	18,288	8,865	9,423			

ANNEX 5: GOAL 3 SUPPORTING INFORMATION

TABLE 20: INTERVENTIONS BY FACTORS RELATED TO INDIVIDUAL RADICALIZATION AND RECRUITMENT

Country – Activity	Push Factors				Pull Factors			
	High Levels of Social Marginalization/ Fragmentation	Poorly Governed or Ungoverned Areas	Government Repression/ Human Rights Violations	Endemic Corruption and Elite Impunity	Cultural Threat Perceptions	Access to Material Resources	Social Status/ Respect from Peers/ Sense of Belonging	Self-Esteem or Personal Empowerment/ Adventure
Caribbean - CBSI	✓	✓			✓	✓	✓	✓
East Africa - CVE	✓	✓					✓	✓
El Salvador - CARSI	✓	✓	✓	✓	✓	✓	✓	✓
Guatemala - CARSI	✓	✓	✓	✓	✓	✓	✓	✓
Honduras - A Ganar	✓	✓		✓			✓	✓
Honduras - CARSI	✓	✓	✓	✓	✓	✓	✓	✓
Macedonia - IIEP	✓	✓	✓		✓		✓	
Timor-Leste - YEPS	✓	✓	✓				✓	✓
Yemen - CLP	✓	✓			✓	✓	✓	

TABLE 21: INTERVENTION APPLICATIONS OF USAID’S CORE BODY OF PROGRAM PRINCIPLES (FOR ANALYSIS, PLANNING, AND DESIGN)

Country – Activity	Focus on Drivers of Violent Extremism and Insurgency	Promote Inclusive Country Ownership	Exercise Selectivity	Take a Coordinated and Integrated Approach	Tailor and Coordinate Communications	Think Locally and Bring an Entrepreneurial Approach	Consider Transnational Strategies
Caribbean - CBSI		✓	✓	✓			
East Africa - CVE	✓	✓		✓			✓
El Salvador - CARSI	✓	✓	✓		✓		✓
Guatemala - CARSI	✓	✓	✓		✓		✓
Honduras - A Ganar	✓	✓	✓				
Honduras - CARSI	✓	✓	✓		✓		✓
Macedonia - IIEP	✓	✓		✓	✓		
Timor-Leste - YEPS	✓	✓		✓	✓		
Yemen - CLP		✓		✓		✓	

TABLE 22: INTERVENTION APPLICATIONS FOR OPERATIONAL AND MANAGEMENT RESPONSIVENESS

Country – Activity	Flexibility, Agility, and Procurement Speed	Intensive Program Management	Innovation, Evaluation, and Learning	Informed Risk-Taking and Experimentation	Establish and Empower a Steering Committee
Caribbean - CBSI		✓			✓
East Africa - CVE			✓		
El Salvador - CARSI	✓		✓		✓
Guatemala - CARSI	✓		✓		✓
Honduras - A Ganar	✓	✓	✓		
Honduras - CARSI	✓		✓		✓
Macedonia - IIEP	✓	✓	✓	✓	✓
Timor-Leste - YEPS			✓	✓	✓
Yemen - CLP	✓	✓	✓		✓

TABLE 23: GOAL 3 EVALUATIONS ADDRESSING SPECIFIC EICC PRIORITY AREAS AND CROSSCUTTING TOPICS

Country – Activity	Topics of Interest						Crosscutting					
	Safer Learning	Equity	Improved Programming	Strengthened Institutional Capacity	Countering Violent Extremism	Social Emotional Learning	Access	Gender	Disability	Innovative Financing	Scaling and Sustainability	Information Technology
Afghanistan - ALSE		✓	✓	✓		✓	✓	✓			✓	
Afghanistan - BELT		✓	✓	✓			✓	✓			✓	✓
Cambodia - SDPP		✓	✓	✓		✓	✓	✓			✓	✓
Caribbean - CBSI	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
DRC - OPEQ	✓	✓	✓			✓	✓	✓			✓	
DRC - Youth and Alternative Education in North Kivu	✓	✓	✓	✓		✓	✓	✓			✓	
East Africa - CVE	✓	✓	✓	✓	✓	✓	✓	✓		✓		✓
El Salvador - CARSI	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓
El Salvador - RERA	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	
Ethiopia - SCOPSO	✓	✓	✓	✓		✓	✓	✓			✓	
Ghana - PWC	✓	✓	✓	✓			✓	✓	✓		✓	✓
Guatemala - CARSI	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓
Honduras - A Ganar	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓
Honduras - CARSI	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	
India - SDPP		✓	✓	✓		✓	✓	✓			✓	✓
Indonesia - OVC	✓	✓		✓		✓	✓		✓		✓	
Jordan - LETS	✓	✓	✓	✓		✓	✓	✓		✓	✓	
Macedonia - IIEP	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mali - EESA	✓	✓	✓	✓	✓	✓	✓	✓				
Mali (Gao Region) - RERA	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	
Nigeria - CECA	✓	✓	✓	✓	✓	✓	✓	✓				
Nigeria - NEI		✓	✓	✓			✓	✓			✓	
Nigeria - RARA		✓	✓	✓			✓	✓			✓	
Senegal - EDB	✓	✓	✓	✓			✓	✓		✓	✓	✓
Tajikistan - SDPP		✓	✓	✓		✓	✓	✓			✓	✓
Timor-Leste - SDPP		✓	✓	✓		✓	✓	✓		✓	✓	✓
Timor-Leste - YEPS	✓	✓		✓		✓	✓	✓				✓
Yemen - CLP	✓	✓	✓	✓			✓	✓	✓	✓	✓	

TABLE 24: NEEDS ASSESSMENT BY OUTCOME, DATA COLLECTION METHODOLOGY, ENGAGEMENT OF LOCAL ACTORS

Country – Activity	Main outcomes of interest (same types of students, teacher, community level outcomes?)	Do they employ similar or different quant/qual data collection methodologies?	Do they engage with local actors in implementation of the assessment or usage of the resultant data?
DRC - Youth and Alternative Education in North Kivu	To understand the demand for alternative education in conflict affected Northern Kivu DRC.	200 young people, 25 key informant interviews. Then 23 government officials, donors, and national and international nongovernmental actors in the DRC education sector.	They do engage with local education actors to understand policy framework for alternative education
El Salvador - RERA	Education sector (schools, education staff, learners, families, and school communities)	Qualitative situation analysis (secondary data and key informant interviews at national level with primary data from a limited purposive sample of school communities in high risk municipalities.	USAID Mission to El Salvador and the USAID Goal 3 Education Team in Washington asked the USAID ECCN to directly manage a RERA in El Salvador.
Mali - EESA	Malian School aged children/Malian education system/community impacted by severe food crisis during the armed rebellion in March 2012.	Interview survey (school administrators, teachers, parents, students)	
Mali (Gao Region) - RERA	Total of 7,450 children aged 9-14 who previously dropped out or never attended school. And 2,100 young people aged 15-24 to complete work readiness program to find employment. (school system, community, context of violence)	Analysis of secondary data (Education cluster) and primary data (interviews and focus groups) collected by partner NGO's.	Engaged with local NGO's to implement assessment
Nigeria - CECA	IDP youth and children in Northern states of Nigeria	Focus groups with parents, teachers, IDP youth, in depth interviews with host community members segregated by age/sex/ displacement status	USAID/Nigeria mission funded as key component of project design of implementation in Northern Nigeria states.

ANNEX 6: 2017 REVISED VERSION OF THE EVALUATION QUALITY ASSESSMENT TOOL ¹⁶⁵

Principle of Quality	Impact Evaluations	Performance Evaluations		Review Result	Overall Conclusion	Notes/Justification
		Quantitative	Qualitative			
Conceptual framing		[1] Are the research/evaluation questions included in the report?		yes/no	adequate/not adequate	
		[2] Does the report include research/evaluation hypotheses?		yes/no		
		[3] Are the evaluation questions appropriate for the intervention's conceptual framework (logframe/theory of change/ results framework)?		yes/partial/no/not applicable		
		[4] Does the report acknowledge/draw upon existing relevant research?		yes/partial/no		
		[5] Does the report explain the local context in sufficient detail?		yes/partial/no		
Openness and transparency		[6] Is the report open about study limitations with the implementation of the evaluation , such as issues faced during data collection that might affect the study's design?		yes/partial/no	adequate/not adequate	
		[7] Is the report open about study limitations due to issues with the implementation of the intervention being evaluated?		yes/partial/no/not applicable		
		[8] Does the discussion about the findings refer to relevant contextual factors or methodological considerations?		yes/no/not applicable		
		[9] Is the report open about potential influence due to the study team composition?		yes/partial/no		
Cultural appropriateness		[10] Does the report list steps taken to ensure that study questions and methodology are informed by local stakeholders, are culturally relevant and contextually appropriate?		yes/no	adequate/not adequate	
		[11] Does the report list steps to address and document that data collection tools were developed/adapted with participation of relevant local stakeholders and are culturally appropriate?		yes/partial/no		
		[12] Does the report list steps taken to validate findings/conclusions/recommendations with local stakeholders as part of the evaluation ?		yes/no		
		[13] Was the study designed to take into account locally relevant stratifiers, such as political, social, ethnic, religious, geographical or sex/gender phenomena during data collection and data analysis?		yes/partial/no		

¹⁶⁵ The evaluation quality assessment tool developed as part of the Assessment of the Quality of USAID-Funded Evaluations in the Education Sector, 2013-2016, was revised upon completion of the review, based on the comments from reviewers. This version reflects these revisions.

Principle of Quality	Impact Evaluations	Performance Evaluations		Review Result	Overall Conclusion	Notes/Justification
		Quantitative	Qualitative			
Robustness of methodology	[14] Is the methodology explained in sufficient detail?			yes/partial/no	adequate/not adequate	
	[15] Is the methodology appropriate for answering posed study questions?			yes/partial/no/not applicable		
	[16] Does the counterfactual meet standards of rigor?			yes/no/not applicable		
	[17] Does the report include information from multiple data sources and how the data were triangulated?			yes/partial/no/not applicable		
	[18] Does the report mention steps taken to mitigate common threats to the integrity of the evaluation (such as non-equivalence at baseline, non-compliance, spillover, systematic attrition) or common biases (confounding bias, selection bias, experimenter bias, etc)?			yes/partial/no		
	[19] For the quantitative research methods used, are the sampling approach and sample size calculations presented in sufficient detail (to include, at a minimum, type of analysis, MDES, alpha and beta)?			yes/partial/no/not applicable		
			[20] For the qualitative research methods used, is the sampling approach described in sufficient detail? (at a minimum, a rationale for the sample size and method of sample selection) and is it appropriate for the study objectives?	yes/partial/no/not applicable		
Validity	[21] Do indicators used in the evaluation capture the construct or phenomenon being investigated?			yes/partial/no/not applicable	adequate/not adequate	
	[22] Were the sampling conducted in such a way such that the results are generalizable to the population of beneficiaries reached through the activity?			yes/partial/no/not applicable		
	[23] Does the report allude to whether the study findings may have been biased by the activity of doing the study itself?			yes/no		
	[24] Does the report address the external validity of findings?			yes/partial/no/not applicable		
	[25] Were all data collection tools piloted with representatives of target populations prior to beginning of the data collection?			yes/partial/no		
	[26] Are confidence intervals reported around point estimates?			yes/no/not applicable		
	[27] Is treatment effect presented in terms of effect size?			yes/no/not applicable		

Principle of Quality	Impact Evaluations	Performance Evaluations		Review Result	Overall Conclusion	Notes/Justification
		Quantitative	Qualitative			
Reliability	[28] Does the report list steps taken to ensure that data were collected with a high degree of reliability?			yes/partial/no	adequate/not adequate	
	[29] Does the report adequately address missing data/non-response?			yes/partial/no		
Cogency	[30] Are all the study questions, including sub-questions, answered?			yes/no/not applicable	adequate/not adequate	
	[31] Does the Executive Summary include answers to all of the study questions?			yes/no		
	[32] Is the report accessible to the audiences for whom the report indicates it is written (e.g., minimizing technical jargon if intended to the general public)?			yes/no		
	[33] Are conclusions based on findings and are the findings related to the evaluation questions?			yes/partial/no/not applicable		
	[34] Is the narrative in the report supported by charts, maps and infographics that help non-technical audiences easily understand the study findings?			yes/partial/no		

Item Description and Source

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Conceptual Framing: Study questions included	[1] Are the research/evaluation questions included in the report? [All evaluation types]	yes/no	All research/evaluation questions must be phrased as questions; it is not enough that they be inferable from the stated objectives of the study. Questions must be clearly stated and be answerable through the reported research methods.	ADS 201 maa: Evaluation reports should adequately address all evaluation questions included in the SOW, or the evaluation questions subsequently revised and documented in consultation and agreement with USAID.
Conceptual Framing: Study hypotheses included	[2] Does the report include research/evaluation hypotheses? [All evaluation types]	yes/no	Research/evaluation hypotheses must be explicitly described; it is not enough that they be inferable from the stated objectives of the study.	BE2, Checklist: Does the study outline a hypothesis?
Conceptual Framing: Study questions appropriate given the intervention's conceptual framework	[3] Are the evaluation questions appropriate for the intervention's conceptual framework (logframe/theory of change/ results framework)? [IE, Perf. Quant, Perf. Qual]	yes/partial/no/NA	All research/evaluation questions should be based on the intervention's conceptual framework. "Partial" score could be given when some, but not all, listed evaluation questions correspond to the intervention's conceptual framework. "NA" score should be given to research studies that do not evaluate a specific intervention.	BE2, Checklist: Does the study pose an appropriate research question?
Conceptual Framing: Study acknowledges/draws upon existing country-specific research	[4] Does the report acknowledge/draw upon existing relevant research? [All evaluation types]	yes/partial/no	Studies should build on existing research, both local and funded by international donors. The report should specify how questions, methodology, tools and analysis plans are informed by prior research. "Partial" score could be given when only some of the questions are informed by existing knowledge.	BE2, Checklist: Does the study acknowledge existing research?
Conceptual Framing: Local context provided allows non-experts appreciate relevance of the study	[5] Does the report explain the local context in sufficient detail? [All evaluation types]	yes/partial/no	The local context should be explained in enough detail for a general audience to be able to appreciate the relevance of the intervention being evaluated. "Partial" score could be given when some, but not all, elements of the intervention have corresponding contextual information.	USAID Evaluation Policy, page 8: Evaluation reports should include sufficient local and global contextual information so that the external validity and relevance of the evaluation can be assessed.
Conceptual Framing: Conclusion	Conceptual framing: Conclusion [All evaluation types]	adequate/not	Adequate: Overall, this evaluation demonstrates adherence to principles of conceptual framing Not Adequate: This evaluation contains major deficiencies in demonstrating adherence to principles of conceptual framing or provides insufficient information for determining this	

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Conceptual Framing: Justification	Conceptual framing: Notes/Justification [All evaluation types]		For instance: <i>“The authors acknowledge existing research and make clear how their analyses sit within the context of existing work. They provide a theoretical framework in the report, where they outline their major assumptions. The study also poses specific research questions.”</i>	
Openness and Transparency: Open about limitations to implementing the study	[6] Is the report open about study limitations with the implementation of the evaluation , such as issues faced during data collection that might affect the study’s design? [All evaluation types]	yes/partial/no	Limitations to the implementation of the evaluation should be clearly presented. Clarity around study limitations is particularly important if they directly impact the evaluator’s ability to credibly and effectively answer an evaluation question or impact generalizability of the findings (i.e., if data collection was successful but more expensive or inconvenient than anticipated, it is not a limitation). “Partial” score could be given if the report mentions limitations without discussing them in detail.	BE2, page 17: The study should also clearly state the sample size.
Openness and Transparency: Open about limitations to implementing the intervention	[7] Is the report open about study limitations due to issues with the implementation of the intervention being evaluated? [IE, Perf. Quant, Perf. Qual]	yes/partial/no/NA	Limitations to the implementation of the intervention being evaluated should be clearly presented, such as delays or changes that may compromise the integrity of the evaluation design. “Partial” score could be given if the report mentions imitations without discussing them in detail. “NA” score should be given to research studies that do not evaluate a specific intervention.	BE2, page 17: An important sign of quality is whether the author is being self-critical; being open about limitations.
Openness and Transparency: Alternative interpretations of the findings included	[8] Does the discussion about the findings refer to relevant contextual factors or methodological considerations? [All evaluation types]	yes/no/NA	The evaluation report should balance the presentation of the findings with a discussion contextualizing them and/or addressing how they might be affected by methodological decisions. This discussion might include broaching alternative explanations for the findings. If some findings yield inconsistencies with others, this should be discussed as well. “NA” score should be given if individually findings were not conducive with discussion about contextual or methodological considerations and collectively they are not contradictory.	BE2, page 17: An important sign of quality is whether the author is being self-critical; being open about (...) alternative interpretations and pointing out inconsistencies with other results.

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Openness and Transparency: Open about potential biases due to the study team composition	[9] Is the report open about potential influence due to the study team composition? [All evaluation types]	yes/partial/ no	USAID encourages study teams to include at least one evaluation specialist, host country team members, and a team leader who is external to USAID. USAID also requires that evaluation team members certify their independence by signing statements disclosing any conflict of interest or fiduciary involvement with the project or program they will evaluate. It is expected that an evaluation will indicate that such forms, or their equivalent, are on file and available or are provided in an evaluation annex. "Partial" score could be given if some, but not all, these recommendations are followed.	BE2, Checklist: Does the researcher acknowledge their own subjectivity in the process of the research?
Openness and Transparency: Conclusion	Openness and transparency: Conclusion [All evaluation types]	adequate/n ot	Adequate: Overall, this evaluation demonstrates adherence to principles of openness/transparency Not Adequate: This evaluation contains major deficiencies in demonstrating adherence to principles of openness/transparency or provides insufficient information for determining this	
Openness and Transparency: Justification	Openness and transparency: Notes/Justification [All evaluation types]		For instance: <i>"The authors are transparent about the design and methods that have been employed in the evaluation as well as the data (and resulting sample) that have been gathered and analyzed. This allows for the study to be repeated and corroborated."</i>	
Cultural Appropriateness: Study questions and methodology informed by local stakeholders	[10] Does the report list steps taken to ensure that study questions and methodology are informed by local stakeholders, are culturally relevant and contextually appropriate? [All evaluation types]	yes/no	The evaluation questions and methodology should be informed by relevant local stakeholders. This could be done during in-country design workshops as well as through meeting with the ministry or other relevant stakeholders.	ADS 201sae: Is there reasonable assurance that the data collection methods being used do not produce systematically biased data.

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Cultural Appropriateness: Data collection tools developed with participation of local stakeholders	<p>[11] Does the report list steps to address and document that data collection tools were developed/ adapted with participation of relevant local stakeholders and are culturally appropriate?</p> <p>[All evaluation types]</p>	yes/partial/ no	<p>The report should describe whether tools have been developed to suit the local context, such as whether the tool was developed by international experts and then merely translated into a local language or whether local knowledge has been used effectively in the adaptation of the tool to reflect resources relevant to the context, such as including support from host country experts. Quality control of translators (back-translation) is recommended. "Partial" score could be given if some, but not all tools suit the local context.</p>	<p>BE2, page 20: For all research designs, it is important to consider the extent to which the measures/ instruments/ variables used in the study suit local contexts. The reviewer should note whether measures have been developed to suit the local context: does the study, for instance, merely translate into a local language or recognize that a test developed in a specific linguistic area may not be automatically suitable to a local context with translation or because of multiple socio-linguistic processes? The reviewer should also note whether local knowledge has been used effectively in the adaptation of measures to reflect resources relevant to the context; for example, are the instruments designed with support and recognition from the local community?</p>
Cultural Appropriateness: Findings/conclusions/ recommendations validated with local stakeholders	<p>[12] Does the report list steps taken to validate findings/conclusions/ recommendations with local stakeholders as part of the evaluation?</p> <p>[All evaluation types]</p>	yes/no	<p>Findings, conclusions and recommendations must be communicated to the appropriate audiences in a culturally and contextually suitable way prior to finalization of the report, in order to validate accuracy of conclusions and help inform recommendations. Steps to validate these with local stakeholders may include in-country presentations and workshops conducted during the evaluation (instead of as dissemination studies after the evaluation was concluded).</p>	<p>EGRA Toolkit, 2nd edition, page 122: Results must be communicated to the appropriate audiences in a culturally and contextually suitable way in order to support understanding and action.</p>

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Cultural Appropriateness: Findings disaggregated by locally relevant stratifiers	[13] Was the study designed to take into account locally relevant stratifiers, such as political, social, ethnic, religious, geographical or sex/gender phenomena during data collection and data analysis? [All evaluation types]	yes/partial/no	The extent to which a study takes into account locally relevant stratifiers has considerable bearing on the study's design, its analytical strategy and the interpretation of its findings. Being informed by locally relevant stratifiers might include making cross-cultural or cross-linguistic comparisons part of the analytical strategy or ensuring that knowledge of the local context is used in the interpretation of differential effects between groups. "Partial" score should be assigned when the study is purposeful with considering variable impacts on gender but not any other stratifiers.	BE2, page 20: This includes the extent to which the analysis includes locally relevant social stratifiers (for example, socio-economic status, gender, rural-urban differences, etc.) and influences which may affect interpretation of results.
Cultural Appropriateness: Conclusion	Cultural appropriateness: Conclusion [All evaluation types]	adequate/not	Adequate: Overall, this evaluation demonstrates adherence to principles of cultural appropriateness. Not Adequate: This evaluation contains major deficiencies in demonstrating adherence to principles of cultural appropriateness or provides insufficient information for determining this.	
Cultural Appropriateness: Justification	Cultural appropriateness: Notes/Justification [All evaluation types]		For instance: "The evaluation describes systematic processes used to check for the cultural relevance of measurement items (for example, in the absence of lists of age-specific words for Bangla-speaking children, a list was created of words that fit two criteria: they should be known to grade 1 or 2 children but unknown to preschoolers, and they should be used in the storybooks). Thus, the instrument used is culturally sensitive. The analysis is also culturally sensitive, as it discusses the factors that undermine or promote educational outcomes within the Bangladeshi context. The study discusses the use of two supply-and-demand side interventions – a school-only grant and a school grant plus an education allowance – which the authors discuss in relevance to the context, where grants are used to provide key inputs to schools while the education allowance provides a conditional monetary incentive for out-of-school children to attend school."	

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Robustness of Methodology: Methodology explained in detail	<p>[14] Is the methodology explained in sufficient detail?</p> <p>[All evaluation types]</p>	yes/partial/no	<p>USAID requires that an evaluation report identifies the study design, data collection methods and data analysis techniques used. It is common to include the methodology description in the body of the report under a methodology section with a longer and more detailed methods annex.</p> <p>The description of methods must indicate: how respondents were selected,</p> <ul style="list-style-type: none"> • what types of interviews were conducted; • with whom they were conducted (e.g., key informant interviews, individual interviews with beneficiaries, group interviews) and; • detailed information on the kinds of analyses that were conducted (e.g., correlations, regressions, content analysis, pattern analysis). <p>“Partial” score could be given if some, but not all elements mentioned (design, data collection methods and data analysis techniques) were described in sufficient detail.</p>	ADS 201 maa: Evaluation methodology should be explained in detail and sources of information properly identified.
Robustness of Methodology: Methodology appropriate for answering posed study questions	<p>[15] Is the methodology appropriate for answering posed study questions?</p> <p>[All evaluation types]</p>	yes/partial/no/NA	<p>USAID recognizes that the methodology used to address the posed questions may be defined in the issued Scope of Work for the evaluation. USAID also recognizes that different designs are more or less appropriate to answering different research questions, and that the selection of method (or methods) for a particular evaluation also balances cost, feasibility, and the level of rigor needed to inform specific decisions. Assessing the appropriateness of the chosen methodology may be further complicated when the evaluation includes a variety of questions that require a mixed-method approach; for such evaluations, the assessment of the methodology must include the review of the evaluation design vis-a-vis each stated study questions.</p> <p>“Partial” score could be given if the methodology proposed is appropriate for some, but not all posed questions. “NA” score should be given if the study does not pose research/evaluation questions.</p>	USAID Evaluation Policy, page 8: evaluation should principally consider the appropriateness of the evaluation design for answering the evaluation questions as well as balance cost, feasibility, and the level of rigor needed to inform specific decisions.

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Robustness of Methodology: Counterfactual meet standards of rigor	[16] Does the counterfactual meet standards of rigor? [IE]	yes/no/NA	Measuring what would have happened in the absence of an intervention is a requirement for establishing a causal relationship. A counterfactual can be created in a number of ways, from simply using respondents from a geographically close unit as comparison group to using statistical analysis to compensate for the potential selection biases of non-randomization to randomly assigning subjects to treatment(s) and control groups. Considerations about its rigor may include a review of information in the report about baseline equivalence, differential attrition, etc. "NA" score should be given if the evaluation is not an Impact Evaluation.	USAID Evaluation Policy, page 3: Impact evaluations measure the change in a development outcome that is attributable to a defined intervention; impact evaluations are based on models of cause and effect and require a credible and rigorously defined counterfactual to control for factors other than the intervention that might account for the observed change. (...) Performance evaluations encompass a broad range of evaluation methods. They often incorporate before-after comparisons, but generally lack a rigorously defined counterfactual.
Robustness of Methodology: Data triangulation described as part of methodology	[17] Does the report include information from multiple data sources and how the data were triangulated? [All evaluation types]	yes/partial/no/NA	Typically, stronger bodies of evidence are likely to emerge if similar findings are obtained from different types of data (e.g., tests, interviews, observations) and respondent types (e.g., students, parents, teachers). It is important that contradictory data be taken into account when discussing the findings. "Partial" score could be given if data from different sources are presented but the findings don't connect them into a coherent narrative. "NA" score should be given if the evaluation does not use multiple data sources.	CASP, Qualitative Checklist: To what extent contradictory data are taken into account?
Robustness of Methodology: Addressed internal validity, either threats to inference or common biases	[18] Does the report mention steps taken to mitigate common threats to the integrity of the evaluation (such as non-equivalence at baseline, non-compliance, spillover, systematic attrition) or common biases (confounding bias, selection bias, experimenter bias, etc)? [All evaluation types]	yes/partial/no	USAID Evaluation Policy requires that evaluation reports address methodologically common limitations, such as when there is a disjunction between the treatment that is assigned and the treatment that is received (non-compliance). "Partial" score could be given if some, but not all threats or biases identified are discussed.	USAID Evaluation Policy, page 10: Evaluation reports that include the original statement of work, a full description of methodology (or methodologies) used, as well as the limitations in the inferences that can be drawn.

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Robustness of Methodology: Described sampling approach and parameters used to compute sample size	[19] For the quantitative research methods used, are the sampling approach and sample size calculations presented in sufficient detail (to include, at a minimum, type of analysis, MDES, alpha and beta)? [IE, Perf. Quantitative]	yes/partial/ no/NA	Details of power calculation should be included in either the main body of the report or in an annex. This should include the parameters used in the power function that relates power (beta) to its determinants: (1) level of significance (alpha), (2) minimum detectable effect size (MDES) or minimum detectable impact (MDI), (3) and the sample size. "Partial" score could be given if the description of the sample size calculations presents only some of the parameters used. "NA" score could be given if the evaluation/research used only qualitative research methods	JPAL's Running Randomized Evaluations, page 271: A power function relates power to its determinants: (1) level of significance, (2) MDE size, (3) the unexplained variance of the outcome of interest, (4) allocation fractions, (5) and the sample size.
Robustness of Methodology: Described sampling approach to collect qualitative data	[20] For the qualitative research methods used, is the sampling approach described in sufficient detail? (at a minimum, a rationale for the sample size and method of sample selection) and is it appropriate for the study objectives? [Perf. Qualitative]	yes/partial/ no/NA	Researchers/evaluators should provide a description of the sampling frame and potential issues with it, if any. This should include an explanation of how the participants were selected, whether these participants were the most appropriate to provide access to the type of knowledge sought by the study, whether there was a point at which incoming data produced little or no new information (saturation) as well as any discussions around recruitment, such as why some people might have chosen not to take part in the study. "Partial" score should be given if only some of these elements were discussed. "NA" score should be given if this study did not use qualitative research methods.	CASP, Qualitative Checklist: Recommended considerations about "If the researcher has explained how the participants were selected"; "If they explained why the participants they selected were the most appropriate to provide access to the type of knowledge sought by the study"; "If there are any discussions around recruitment (e.g. why some people chose not to take part)".
Robustness of Methodology: Conclusion	Robustness of methodology: Conclusion [All evaluation types]	adequate/n ot	Adequate: Overall, this evaluation demonstrates adherence to principles of appropriateness/rigor of chosen methodology Not Adequate: This evaluation contains major issues with the appropriateness of the chosen methodology, major deficiencies in the rigor with which it was applied or provides insufficient information for determining this	
Robustness of Methodology: Justification	Robustness of methodology: Notes/Justification [All evaluation types]		For instance: <i>"The study aims to identify and examine specific effects of receiving grants alone compared to receiving grants as well as training on student learning outcomes. The study clearly aims to establish a causal linkage between grants versus grants/training on student outcomes. The experimental design was, therefore, most appropriate to answer the research question. The study demonstrates rigorous application of the experimental technique within The Gambian setting. The authors clearly describe the interventions and adopt all the rigors of a well-applied randomization."</i>	

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Validity: Addressed construct validity of the assessment tools	[21] Do indicators used in the evaluation capture the construct or phenomenon being investigated? [IE, Perf. Quantitative]	yes/partial/ no	In order to assess the validity of the measurement, it is important to consider whether or not the chosen indicators adequately capture the concepts being measured or whether there are other dimensions central to the concepts that are being ignored, such as a labor market condition index that ignores underemployment. "Partial" scores could be given if some, but not all key indicators, adequately captured the concepts being measured.	BE2, page 24: In the case of measurement validity, it is important to repeatedly consider whether or not the indicator chosen fully captures the concept being measured. Are there other dimensions of the central concept that are being ignored?
Validity: Addressed the external validity of findings from the sample to population	[22] Were the sampling conducted in such a way such that the results are generalizable to the population of beneficiaries reached through the activity? [IE, Perf. Quantitative]	yes/partial/ no/NA	A number of characteristics of the survey design, such as timing of the assessment and absence of sampling weights, may affect the interpretation and/or calculation of population estimates. The evaluator/research may provide information about the timing of the assessment (e.g., pre-test and post-test being conducted at comparable time points in a cross-sectional design) or construction and use of sampling weights in the analysis (when different observations in a random selection process may have different probabilities of selection). "Partial" score could be given if the report mentions that the interpretation and/or calculation of some but not all population estimates took into account relevant survey design characteristics. "NA" score should be given in case this is a qualitative study.	StataCorp's Survey Data Reference Manual, page 3: In sample surveys, observations are selected through a random process, but different observations may have different probabilities of selection. Weights are equal to (or proportional to) the inverse of the probability of being sampled. Various postsampling adjustments to the weights are sometimes made, as well. A weight of w_j for the j th observation means, roughly speaking, that the j th observation represents w_j elements in the population from which the sample was drawn. Omitting weights from the analysis results in estimates that may be biased, sometimes seriously so.

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Validity: Addressed ecological validity of findings	[23] Does the report allude to whether the study findings may have been biased by the activity of doing the study itself? [All evaluation types]	yes/no	Evaluators/researchers might discuss in the report whether findings could have been influenced by the process of research itself (ecological validity) or whether participants may have changed their behavior in response to their perception of the evaluators' objective (response bias), such as when the treatment group works harder than normal in response to being part of an evaluation (Hawthorne effects). Note that the tendency of participants to give an answer to a question that is in line with social norms even if this does not accurately reflect their experience (social desirability bias) is not relevant for this question. This might include discussions about whether the implementer may have brought in irreproducible energies that accountable for the success of a pilot but that might be absent in a scale-up.	BE ² , page 25: whether the findings could have been influenced by the process of research itself (ecological validity).
Validity: Addressed the external validity of findings to other contexts	[24] Does the report address the external validity of findings? [All evaluation types]	yes/partial/ no/NA	Findings are externally valid when they are valid in contexts other than those the evaluation was conducted in. Thus, researchers/evaluators may discuss the local conditions that would make it replicable in a different context. "Partial" score could be given if the external validity of some, but not all key findings, are discussed in the report. "NA" score should be given in case this evaluation did not intend to have data from a sample extrapolated to a population.	BE2, Checklist: To what extent is the study externally valid?
Validity: Data collection tools piloted with representatives of target populations	[25] Were all data collection tools piloted with representatives of target populations prior to beginning of the data collection? [All evaluation types]	yes/partial/ no	Researchers/evaluators should describe if respondents used to pilot the data collection tools were similar to the target population of the full study. "Partial" score could be given if the report mentions that piloting was done but not with who.	EGRA Toolkit, 2nd edition, page 92: The students and schools selected for the pilot sample should be similar to the target population of the full study.
Validity: Confidence intervals reported around point estimates	[26] Are confidence intervals reported around point estimates? [IE, Perf. Quantitative]	yes/no/NA	USAID recommends that the margin of error be reported along with the findings from statistical samples. "NA" score should be given if the study does not use inferential statistical methods.	ADS 201sae: Has the margin of error been reported along with the data? (Only applicable to results obtained through statistical samples.)

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Validity: Treatment effects presented in terms of effect sizes	[27] Is treatment effect presented in terms of effect size? [IE, Perf. Quantitative]	yes/no/NA	Researchers/evaluators often record the study findings in the units of the outcome variable. To improve the comparability of effect size estimates across outcome variables and across studies, effect sizes in terms of standard deviations should also be provided, taking into consideration the study design. "NA" should be given if the study did not conduct statistical hypothesis testing (as in the case of qualitative studies).	What Works Clearinghouse Procedures and Standards, page 22: For all studies, the WWC records the study findings in the units reported by the study authors. In addition, the WWC computes and records the effect size associated with study findings on relevant outcome measures.
Validity: Conclusion	Validity: Conclusion [All evaluation types]	adequate/not	Adequate: Overall, this evaluation demonstrates adherence to principles of validity. Not Adequate: This evaluation contains major deficiencies in establishing the measurement, internal, external or ecological validity or provides insufficient information for determining this.	
Validity: Justification	Validity: Notes/Justification [All evaluation types]		For instance: <i>"The authors describe steps they took to address the validity of the study. For example, items included in the test had to relate directly to what grade 5 children would be expected to know at the start and end of the school year and statistical analyses were conducted to assess the internal consistency of questions in order to refine and adjust the assessment tools (measurement validity). In assessing learning progress of pupils in grade 5, the study included initial test scores into the estimation and controlled for background factors that may generate biases (internal validity). The study is based on longitudinal data collected from 5 provinces out of 58 in Vietnam, the generalizability of the findings is somewhat questionable (external validity), and there is no discussion of whether the findings could have been influenced by the process of research itself (ecological validity). While it could be improved, overall this study meets basic standards of scientific validity."</i>	
Reliability: Steps taken to ensure that data were reliably collected	[28] Does the report list steps taken to ensure that data were collected with a high degree of reliability? [All evaluation types]	yes/partial/no	USAID recommends that data collection methods be documented in writing to ensure that the same procedures are followed each time. The report may describe the use of data quality assurance checks such as accompaniments, back-checks and scrutiny, and these may have been conducted through spot-checking or for all questions in the data collection form. In case of paper-and-pencil data collection, double data entry report and/or double manual verification may also be mentioned in the report. Steps used in qualitative studies may include audio recording, videotaping and transcribing interviews. "Partial" score could be given if steps to ensure the reliability of some, but not all data collected, are described.	ADS 201sae: Are data collection and analysis methods documented in writing and being used to ensure the same procedures are followed each time?

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Reliability: Target and actual sample sizes reported and non-responses bias discussed	[29] Does the report adequately address missing data/non-response? [All evaluation types]	yes/partial/no	Researchers/evaluators should report the target number of respondents, the number of respondents reached, and the number of respondents who were included in the data analysis. This includes non-response in qualitative studies. For quantitative evaluations, the report may also mention using post-stratification to adjust weights for non-response. "Partial" score could be given if information about valid responses is provided to some, but not all data used in the findings.	What Works Clearinghouse Procedures and Standards, page D.4: study must report the number of students (teachers, schools, etc.) who were designated as treatment and comparison group samples and the proportion of the total sample (e.g., students, teachers, or schools in the treatment and comparison samples combined) with outcome data who were included in the impact analysis (i.e., response rates). Both overall attrition and attrition by treatment status must be reported.
Reliability: Conclusion	Reliability: Conclusion [All evaluation types]	adequate/not	Adequate: Overall, this evaluation demonstrates adherence to principles of reliability. Not Adequate: This evaluation contains major deficiencies in establishing the reliability of the measurement or provides insufficient information for determining this.	
Reliability: Justification	Reliability: Notes/Justification [All evaluation types]		For instance: <i>“This study used multiple researchers to undertake school observations and interviews; the researchers checked their own conclusions with each other and then cross-checked them against the wider analytical team to analyze between schools. The team ensured that different types of data were collected – observations, interviews and document analysis – to triangulate findings and take into account the variety of possible contexts. The authors also provide a good example of how to enhance the reliability of qualitative analysis: interviews were videotaped and transcribed.”</i>	
Cogency: Answers to all study questions, including sub-questions, included	[30] Are all the study questions, including sub-questions, answered? [All evaluation types]	yes/no/NA	The purpose of an evaluation report is to provide the evaluators’ findings and recommendations on each and every evaluation question. Accordingly, USAID expects that the answers to all evaluation questions, including any sub-questions, will be provided in the report. "NA" score could be given if no evaluation questions are provided in the report.	ADS 201 mah: Address all evaluation questions in the Statement of Work (SOW) or document approval by USAID for not addressing an evaluation question.

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Cogency: Answers to all study questions included in the Executive Summary	[31] Does the Executive Summary include answers to all of the study questions? [All evaluation types]	yes/no	The executive summary must provide an accurate representation of the main elements of the evaluation report without adding any new material information or contradicting the evaluation report in any way. As such, it is recommended that all evaluation questions/issues, including any sub-questions/issues, will be provided in the Executive Summary.	ADS 201 maa: The Executive Summary of an evaluation report should present a concise and accurate statement of the most critical elements of the report.
Cogency: Written in a language adequate to its stated audience	[32] Is the report accessible to the audiences for whom the report indicates it is written (e.g., minimizing technical jargon if intended to the general public)? [All evaluation types]	yes/no	Reports should be written in an accessible way to non-experts. Excessive use of research terminology is also undesirable; the report should favor terminology that its intended audience is expected to be familiar with.	USAID Evaluation Policy, page 10: USAID evaluations of all types will use sound social science methods and should include the following basic features: (...) Evaluation reports that are shared widely and in an accessible form with all partners and stakeholders, and with the general public.
Cogency: Connection between study questions, findings, conclusions and recommendations	[33] Are conclusions based on findings and are the findings related to the evaluation questions? [All evaluation types]	yes/partial/ no	In order to strengthen the study's conclusion validity, USAID requires that evaluation findings be based on reliable quantitative and/or qualitative data, and that conclusions and recommendations should be based on these findings. USAID also encourages evaluators to present a clear progression from Study questions to Findings to Conclusions to Recommendations (if any) in their reports, such that none of a report's conclusions and recommendations appear to lack grounding. "Partial" score could be given if some supporting data is provided for some, but not all findings.	E3 Sectoral Synthesis Checklist, question 32: Can a reader can follow a transparent path from findings to conclusions to recommendations?
Cogency: Visuals are helpful for a non-technical audience to understand the findings	[34] Is the narrative in the report supported by charts, maps and infographics that help non-technical audiences easily understand the study findings? [All evaluation types]	yes/partial/ no	Visuals must be used to facilitate understanding of the findings by general audiences. Visuals should be standalone, such that they are interpretable without the audience needing to read extra text. "Partial score" could be given if the report uses visuals to an insufficient extent.	EGRA Toolkit 2nd edition, page 120: Data visualization must be used to facilitate understanding of the findings by general audiences. Visualizations are "standalone," such that the visual is interpretable without the audience needing to read extra text.
Cogency: Conclusion	Cogency: Conclusion [All evaluation types]	adequate/n ot	Adequate: Overall, this evaluation demonstrates adherence to principles of cogency. Not Adequate: This evaluation contains major deficiencies in demonstrating adherence to principles of cogency or provides insufficient information for determining this.	

Principle/Abbrev. Item	Question	Score	Descriptor	Source
Cogency: Justification	Cogency: Notes/Justification [All evaluation types]		For instance: <i>“The evaluation contains a clear, logical argumentative thread that runs through the entire report. This links the conceptual framework for the study to the data and analysis, and, in turn, to the conclusions. The conclusions are backed up by the evaluation findings.”</i>	

ANNEX 7: REFERENCES

- Abdul-Hamid, Huesin, Harry Anthony Patrinos, Joel Reyes, Jo Kelcey, and Diaz Varela, *Learning in the Face of Adversity: The UNRWA Education Program for Palestine Refugees*. Washington, D.C.: World Bank, 2014.
- Aber, J Lawrence, Leighann Starkey, Carly Tubbs, Catalina Torrente, Brian Johnston, Sharon Wolf, Anjali Shivshanker, and Jeannie Annan, *Opportunities for Equitable Access to Quality Basic Education (OPEQ). Final Report on the Impact of the OPEQ Intervention in the Democratic Republic of Congo*, 2016.
- Adams, Saaka, Charles Nornoo, S. Dunham Rowley, and Brenda Sinclair, *Final performance evaluation of USAID/Ghana's partnership for accountable governance in education (PAGE) project: final report*, 2014.
- Alcid, Annie, *A Randomized Control Trial of Akazi Kanoze Youth in Rural Rwanda*, 2014.
- Alvares de Azevedo, Thomaz. *Assessment of the Quality of USAID-Funded Evaluations: Education Sector, 2013-2016*. USAID, January 2018. http://pdf.usaid.gov/pdf_docs/pa00srw1.pdf.
- Alvares de Azevedo, Thomaz and Sean Kelly. *Supplemental Topics from the Synthesis of USAID-Funded: Evaluations: Education Sector, 2013–2016*. USAID, January 2018.
- American Institutes for Research, *Performance evaluation of reading for Ethiopia's achievement developed institutional improvement (READ II)*, 2015.
- American Institutes for Research, *USAID quality reading project: midterm impact report for the Kyrgyz Republic*, 2016.
- Ashwill, Maximillian, Donald Méndez Quintana, and Francisco Sequeira Rankin, *Reducing risk factors for youth on Nicaragua's southern Caribbean coast: an evaluation of FADCANIC's development model for at-risk youth*, 2016.
- Auten, Sarah, Amit Dabla, and Mark Bassie, *Advancing youth project performance evaluation: final report*, 2016.
- Bahadur Karki, Yagya, Tara Nath Pande, Bhoj Raj Neupane, Khadga B. Karki, Smritee Bajracharya, *Final evaluation report: education for income generation project (EIG)*, 2013.
- Bassie, Mark, John C. Dalton, David P. Evans, Mariama Koon, Frank R. Schorn, John Y. Sumo, *Midterm Assessment of the Liberia Teacher Training Program Phase II*, 2013.
- Berk-Seligson, Susan, Diana Orcés, Georgina Pizzolito, Mitchell A. Seligson, and Carole Wilson, *Central America Regional Security Initiative (CARSI) Impact Evaluation: Guatemala Country Study*, 2014.
- Berk-Seligson Susan, Diana Orcés, Georgina Pizzolito, Mitchell A. Seligson, and Carole Wilson, *Impact Evaluation of USAID's Community-Based Crime and Violence Prevention Approach in Central America: El Salvador Country Report*, 2014.
- Berk-Seligson Susan, Diana Orcés, Georgina Pizzolito, Mitchell A. Seligson, and Carole Wilson, *Impact Evaluation: Honduras Country Report*, 2014.
- Boardman, Gerald, Dr. Dwatmadji, Johan Ceelen, and Wardhani Kusuma, *Evaluation of the Indonesia university partnerships: program: phase three – partnerships #5-#8*, 2015.
- Boyle, Philip, Ettibari Bouasla, Mhammed Abderebbi, *Mid-term evaluation: favorable opportunities to reinforce self-advancement for today's youth (FORSATY)*, 2015.

- Building Evidence in Education (BE2) Steering Committee. Assessing the Strength of Evidence in the Education Sector. 2015. Retrieved from https://www.usaid.gov/sites/default/files/documents/1865/BE2_Guidance_Note_ASE.pdf
- Burde, Dana, Amy Kapit, Matt Lisiecki, Joel Middleton, Vedrana Misic, Mohammad Amin Sadiqi, and Cyrus Samii, *Phase One Outcomes Report- The Assessment of Learning Outcomes and Social Effects on Community-Based Education: A Randomized Field Experiment in Afghanistan*, 2016.
- Burde, Dana, Ozen Guven, Jo Kelcey, Heddy Lahmann, and Khaled Al-Abbadi, *What Works to Promote Children's Educational Access, Quality of Learning, and Wellbeing in Crisis-Affected Contexts*, 2015.
- Butler, Steve, Nana Adeishvili, George Murvanidze, *Mid-Term Performance Evaluation of the Georgia Economic Prosperity Initiative*, 2013.
- Centre d'Etudes et de Recherche sur l'Information en Population et Santé (CERIPS), *Education Emergency Support Activity (EESA) USAID Project Rapid Education Risk Assessment (RERA) for the regions of Segou, Mopti, Tombouctou, Gao, and Kidal*, 2016.
- Creative Associates International, *Community Education and Conflict Assessment (CECA)*, 2015.
- Creative Associates International and Mathematica Policy Research, *Do early warning systems and student engagement activities reduce dropout?: findings from the school dropout prevention pilot program impact evaluation in Cambodia. Volume 1: main findings*, 2015.
- Creative Associates International and Mathematica Policy Research, *Do early warning systems and student engagement activities reduce dropout?: findings from the school dropout prevention pilot program impact evaluation in India. Volume 1: main findings*, 2015.
- Creative Associates International and Mathematica Policy Research, *Do early warning systems and student engagement activities reduce dropout?: findings from the school dropout prevention pilot program impact evaluation in Tajikistan. Volume 1: main findings*, 2015.
- Creative Associates International and Mathematica Policy Research, *Do early warning systems and student engagement activities reduce dropout?: findings from the school dropout prevention pilot program impact evaluation in Timor-Leste. Volume 1: main findings*, 2015.
- Dall, Frank, Lynn Losert, Agung Purwadi, Bandi, Sandra Triamoko, and Yanti Yulianti Masoken, *Indonesia: PRIORITAS Mid-term Evaluation*, 2015.
- Durlak, Joseph, Allison B. Dymnicki, Rebecca Taylor, Roger P. Weissberg, and Kriston B. Schellinger, "The Impact of Enhancing Students' Social and Emotional Learning," *Child Development*, 82, no. 1 (2011): 405-432.
- Dryden-Peterson, Sarah. *Refugee Education: A Global Review* (Geneva: United Nations High Commissioner for Refugees, 2011).
- Duthie, Mike, Mateusz Pucilowski, Erin Murphy-Graham, *A Ganar alliance impact evaluation: summative baseline report : Honduras*, 2015.
- "Education: An Essential Component of Humanitarian Response," Global Education Cluster, 1 September 2016, Retrieved from http://educationcluster.net/wp-content/uploads/2014/01/Education-An-Essential-Component-of-a-humanitarian-response_EN.pdf
- Education Development Center, *Employability Study: An Evaluation of METAS' Basic Labor Competences Training and Certification Program*, 2014.
- Education Development Center, *Literacy, Language and Learning Initiative (L3) National Fluency Assessment of Rwandan Schools: Midline Report*, 2016.

- Education Development Center, *Basa Pilipinas Program Final Outcome Evaluation Report: 2013- 2016*.
- Education in Crisis and Conflict Network. *Analysis of Indicators Used in USAID Activities in Crisis and Conflict Environments*, 2016.
- Education In Crisis & Conflict Network, *Rapid Education and Risk Analysis El Salvador Final Report*, 2016.
- Education Recovery Support Act, *Rapid Education Risk Analysis in the Gao Region*, 2016.
- Education In Crisis & Conflict Network, *USAID ECCN Alternative Education in the DRC Final Research Report*, 2016.
- Evans, David K., and Anna Popova. "What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews." *The World Bank Research Observer* 31, no. 2 (2016). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2571082
- Extension of the USAID Education Strategy to December, 2017. Retrieved from http://pdf.usaid.gov/pdf_docs/PA00MXPF.pdf
- Frischkorn, Rebecca, Zachariah Falconer-Stout, Lyn Messner, *Time to Learn Project Year 4 Performance Evaluation Report*, 2016.
- Gates, Sarah, Laura Lippman, Noel Shadowen, Holly Burke, Obed Diener, and Morrisa Malkin, *Key Soft Skills for Cross-Sectoral Youth Outcomes*. Washington, DC: USAID's YouthPower: Implementation, YouthPower Action, 2016.
- Paul Gertler, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan Chang, and Sally Grantham-McGregor, "Labor Market Returns to Early Childhood Stimulation Intervention in Jamaica," *Science* 344, no. 6187 (2014): 998-1001.
- Gilboy, Andrew, Jane Fraser, Pham Hoa, Ngo Thuan, *Mid-term evaluation of the higher engineering education alliance program (HEEAP)*, 2013.
- Glewwe, Paul W., Eric A. Hanushek, Sarah D. Humpage, and Renato Ravina. *School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010*. No. w17554. National Bureau of Economic Research, 2011.
- Graham, Jimmy, and Sean Kelly. *How Effective Are Early Grade Reading Interventions? A Review of the Evidence*. Washington, DC: World Bank, 2018.
- GRM International, *Evaluation of the opportunities for vulnerable children program Indonesia*, 2013.
- Gurevich, Robert, Joe Kitts, Lee R, Briggs, Steven Hansch, and Edward Allan, *Performance evaluation of the education program of the community livelihoods project (CLP): final report*, 2014.
- Husain, Tariq, Zameer Haider, Aftab Ismail Khan, and Babar Mufti, *Pakistan-United States science and technology cooperation (S&T) program: mid-term performance evaluation report*, 2014.
- Husein Abdul-Hamid, Harry Anthony Patrinos, Joel Reyes, Jo Kelcey, and Andrea Diaz Varela, *Learning in the Face of Adversity: The UNRWA Education Program for Palestine Refugees* Washington, DC: World Bank, 2016.
- International Labour Organization, UN Peacebuilding Support Office, UN Development Programme, and World Bank. *Employment Programmes and Peace: A Joint Statement on an Analytical Framework, Emerging Principles for Action and Next Steps*, September 2016
- Jere, Peter, Janet K. Orr, Jennifer Bisgard, and Michael Ogwa, *Performance Evaluation of the USAID/Malawi Early Grade Reading Activity (EGRA)*, 2015.

- Joseph Durlak, Allison Dymnicki, Rebecca Taylor, Roger Weissberg, and Kriston Schellinger, "The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Intervention," *Child Development* 82, no. 1, 2011.
- Kluve, Jochen, Susana Puerto, David Robalino, Jose Manuel Romero, Friederike Rother, Jonathan Stöterau, Felix Weidenkaff, and Marc Witte, *Interventions to Improve the Labour Market Outcomes of Youth: A Systematic Review of Training, Entrepreneurship Promotion, Employment Services, and Subsidized Employment Interventions*. Geneva: The International Labour Organization, 2016.
- Kraft, Richard, Andrew I. Epstein, *Final Performance Evaluation of the Improving Quality of Primary Education Program (IQPEP) Final Report*, 2014.
- Kremer, Michael, Conner Brannen, and Rachel Glennerster. "The challenge of education and learning in the developing world." *Science* 340, no. 6130 (2013): 297-300.
- Krieger, Douglas, Ghazanfar Hoti, Sarah Tirmazi, Ahmad Jameel, and Sara Azmat Zaidi, *Fulbright student program evaluation in Pakistan midterm performance evaluation report*, 2014.
- Krishnaratne, Shari, Howard White, and Ella Carpenter. "Quality education for all children? What works in education in developing countries." New Delhi: International Initiative for Impact Evaluation (3ie), Working Paper 20 (2013).
- Lancaster, Illana and Justine Mirembe, *Evaluation of HED/USAID women's leadership program-Rwanda (2012-2015)*, 2015.
- Larcom, Joan, Yinka Falola-Anoemuah, Saliyu Ingawa, and Nura Nasir, *Northern education initiative (NEI) project: mid-term performance evaluation : final report*, 2013.
- Leigh, Stuart and Edwin Ochieng, *Final performance evaluation of the teacher education and professional development project in Kenya*, 2013.
- McClellan, Iain and Ilimi Elabeh, *Projet AIDE performance evaluation 2009-2013: evaluation report final*, 2013.
- McEwan, Patrick J. "Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments." *Review of Educational Research* 85, no. 3 (2015): 353-394.
- Menendez England and Associates, *Performance evaluation of the Georgia education management project (EMP)*, 2013.
- Mercy Corps, *Youth & Consequences: Unemployment, Injustice, and Violence*, 2015.
- Mercy Corps, *Motivations and Empty Promises: Voices of Former Boko Haram Combatants and Nigerian Youth*, April 2016.
- Miller, Jason, Joanna Khater, Wassim Katerji, Nawja Andraos, Kamil Wanna, and Harvey Herr, *USAID Lebanon university scholarship program mid-term evaluation*, 2015.
- Mindling, Michael, Deon Edward-Kerr, Janet, Orr, and Winsome Francis, *Midterm performance evaluation of the USAID/Jamaica basic education project: in support of the Jamaica education transformation project*, 2013.
- Moubayed, Layla, Brenda Sinclair, Nedjma Koval, Maram Barqawi, *Final performance evaluation: USAID/Jordan learning environment technical support program*, 2014.
- Moubayed, Layla, Khaled Qubajah, Wala'a Aqrabawi, Maram Baraqawi, and Dima al-Naber, *Final performance evaluation of the USAID/Jordan youth for the future (Y4F) project*, 2014.
- Nagarajan, Geetha, Pedro Carneiro, Melissa Chiappetta, Brian Fuller, Emily Gonzalez, Brenda Sinclair, and Aaron Mapondera, *Impact Evaluation of the Early Grade Reading Activity (EGRA)*, 2015.

- Naskova, Zojka, Mathais Kjaer, Jelena Mihajlovic, and Mirjana Popovic Valjevac, Alison Miranda, Daniel Sabet, and Sarah Alvarez, *Midterm performance evaluation of USAID/Macedonia's interethnic integration in education project*, 2014.
- Ngom, El Hadji, Amadou Wade Diagne, Serigne Malick Fall, *USAID basic education project mid-term evaluation: 'a committed and successful educational community': final report*, 2013.
- NORC at the University of Chicago, *Yes youth can! impact evaluation: final report*.
- Oates, Lauren, Aimee Rose, Rachel Santos, and Zach Warren, *Report of the Verification of the Emergency Textbook Distribution of the Ministry of Education, April – July 2016*, 2016.
- Population, Health, and Development Group, *Final Evaluation Report Education for Income Generation*, 2013.
- PRIN International Consultancy and Research Services PLC, *Final performance evaluation of the school-community partnership serving orphan and vulnerable children affected by HIV/AIDS (SCOPSO) project*, 2013.
- Pritchett, Lant. The Evidence About What Works in Education: Graphs to Illustrate External Validity and Construct Validity. RISE Insights, June 2017. Retrieved from <http://www.riseprogramme.org/content/evidence-about-what-works-education-graphs-illustrate-external-validity-and-construct>
- QED Group, *USAID/Iraq foras final performance evaluation report*, 2015.
- Ramos-Mattoussi, Flavia and Venus Caballero, *External evaluation of the women's leadership program in Paraguay : evaluation report*, 2015.
- Raupp, Magda, Bruce Newman, Luis Revés, Carlos Lauchande and Edward Jay Allan, *Impact evaluation of the USAID/aprender a ler project in Mozambique: final report*, 2016.
- READ Act. Retrieved from <https://www.congress.gov/bill/115th-congress/house-bill/601>
- Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (E/CN.3/2017/2), Annex III. Retrieved from: <https://unstats.un.org/sdgs/indicators/Official%20Revised%20List%20of%20global%20SDG%20indicators.pdf>
- Roberts, Vicki, Michael Lightfoot, Manizha Wafeq, Najeeb Aryan, *Mid-term performance evaluation (April 2012-October 2013): Afghanistan workforce development program (AWDP) project* 2014.
- Rodriguez, Luis, Eric Allemano, Alberto Treves, Reynolds Kissiedo, Queenstar Ocansey, Kofi Mereku, Melissa Chiappetta, and Jean-Camille Kollmorgen, *Performance Evaluation of Public Works Construction Activities to Increase Access to Education in Ghana*, 2015.
- Room to Read, *Data-driven methods for setting reading proficiency benchmarks*, 2017.
- RTI International, *Nigeria Reading and Access Research Activity*, 2015.
- Sedere, Upali, Mamuka Shatirishvili, and Natia Gorgadze, *Final Performance Evaluation of Georgia Education Management Project*.
- Sinclair, Brenda, Jennie Campos, Marissa Kimsey, *Final Performance Evaluation of USAID/Dominican Republic's Effective Schools Program (ESP); final evaluation report*, 2015.
- Social Impact, *Assessment of the Caribbean Basin Security Initiative*, 2016.
- Social Impact, *Basa Pilipinas Impact Evaluation: Midline Report*, 2016

- Statman, James and Bahruz Babayev, *Final Performance Evaluation of the Youth Business Leadership Project (YBLP) in Azerbaijan*, 2014.
- Stein, Kelsi, *Final evaluation 2015 youth engagement to promote stability*, 2015
- Swedberg, Jeffrey and Lainie Reisman, *Mid-Term Evaluation of Three Countering Violent Extremism Projects*, 2013.
- Tumushabe, Joseph, Lydia Sandi, Johannes Mulokozi, and Maurice Egbert Sekabila, *Performance Evaluation of the BridgelT Project*, 2014.
- Turner, Mark, Alexandre Monnard, Anna Yeritsyan, and Arman Udumyan, *Mid-term performance evaluation of junior achievement of Armenia (JAA) entrepreneurship and civic activism for young people: final report*, 2014.
- UNESCO. *Teaching and Learning: Achieving Quality for All*. EFA Global Monitoring Report, 2014. Retrieved online: <http://unesdoc.unesco.org/images/0022/002256/225660e.pdf>
- Unger, Noam. The Shape of U.S. Global Development Reforms. The 2011 Brookings Blum Roundtable Policy Briefs, 2011. Retrieved from https://www.brookings.edu/wp-content/uploads/2016/07/2011_blum_us_global_development_unger.pdf
- US Senate Hearing 111-806. Statement of Senator Patrick J. Leahy. Retrieved from <https://www.gpo.gov/fdsys/pkg/CHRG-111shrg54986/html/CHRG-111shrg54986.htm>
- USAID Evaluation Policy. USAID, October 2016. Retrieved from <https://www.usaid.gov/sites/default/files/documents/1870/USAIDEvaluationPolicy.pdf>
- USAID. Implementation Guidance to the 2011 USAID Education Strategy. USAID, 2012. Retrieved from http://pdf.usaid.gov/pdf_docs/Pdact461.pdf
- USAID. VEI Policy and the Crime and Violence Prevention Field Guide, USAID 2016. Retrieved from http://pdf.usaid.gov/pdf_docs/PA00MG36.pdf
- Universidad Alberto Hurtado, Centro de Investigación y Desarrollo Educacional, *Evaluation: Education Reform in the Classroom (ReAula) Project*.
- Vinogradova, Elena and Gabriel Montero, *Whole School Reading Program Evaluation Findings*, 2013.
- Susan Walker, Susan Chang, Christine Powell, and Sally Grantham-McGregor, "Effects of Growth Restriction in Early Childhood on Growth, IQ, and Cognition at Age 11 to 12 Years and the Benefits of Nutritional Supplementation and Psychosocial Stimulation," *Journal of Pediatrics* 137, no. 1 (2000): 36-41.
- Susan Walker, Susan Chang, Christine Powell, and Sally Grantham-McGregor, "Effects of Early Childhood Psychosocial Stimulation and Nutritional Supplementation on Cognition and Education in Growth-Stunted Jamaican Children: Prospective Cohort Study," *Lancet* 366, no. 9499 (2005): 1804-1807.
- Willis, Garth. *Agendas, Actions, and Accountability in International Development: A Case Study of the USAID Education Strategy for 2011-2015*. 2014.

U.S. Agency for International Development
1300 Pennsylvania Avenue, NW
Washington, DC 20004