



EQUIP2 LESSONS LEARNED IN EDUCATION STUDENT ASSESSMENT

A Guide to Education Project Design,
Evaluation, and Implementation Based on
Experiences from EQUIP2 Projects in
Egypt, Ghana, Honduras, and Namibia

By Li-Ann Kuan, Ph.D.
American Institutes for Research



USAID
FROM THE AMERICAN PEOPLE



EQUIP2: Educational Policy, Systems Development, and Management is one of three USAID-funded Leader with Associates Cooperative Agreements under the umbrella heading Educational Quality Improvement Program (EQUIP). As a Leader with Associates mechanism, EQUIP2 accommodates buy-in awards from USAID bureaus and missions to support the goal of building education quality at the national, sub-national, and cross-community levels.

FHI 360 is the lead organization for the global EQUIP2 partnership of education and development organizations, universities, and research institutions. The partnership includes fifteen major organizations and an expanding network of regional and national associates throughout the world: Aga Khan Foundation, American Institutes for Research, CARE, Center for Collaboration and the Future of Schooling, East-West Center, Education Development Center, International Rescue Committee, Joseph P. Kennedy, Jr. Foundation, Michigan State University, Mississippi Consortium for International Development, ORC Macro, Research Triangle Institute, University of Minnesota, University of Pittsburgh Institute of International Studies in Education, Women's Commission for Refugee Women and Children.

EQUIP2 LESSONS LEARNED IN EDUCATION STUDENT ASSESSMENT

A Guide to Education Project Design, Evaluation, and
Implementation Based on Experiences from EQUIP2 Projects in
Egypt, Ghana, Honduras, and Namibia

Li-Ann Kuan, Ph.D.
American Institutes for Research

2011

This paper was made possible by the generous support of the American people through the United States Agency for International Development (USAID) under Cooperative Agreement No. GDG-A-00-03-00008-00. The contents are the responsibility of FHI 360 through the Educational Quality Improvement Program 2 (EQUIP2) and do not necessarily reflect the views of USAID or the United States Government.



CONTENTS

Acronyms	1
Introduction	3
Methodology Used to Carry Out this Review	7
The Test Development Process	11
Lessons Learned	17
Conclusions	37

ACRONYMS

BECAS	Basic Education Comprehensive Assessment System (Ghana)
BES III	Basic Education Support III (Namibia)
DNEA	Directorate of National Examinations and Assessment (Namibia)
EFA-FTI	Education for All-Fast Track Initiative
EMOE	Egyptian Ministry of Education
EQUALL	Education Quality for All
ERP	Educational Reform Program (Egypt)
GES	Ghana Examination Services
GMOE	Ghanaian Ministry of Education
HMOE	Honduran Ministry of Education
IIEP	UNESCO International Institute for Educational Planning
M&E	Monitoring and Evaluation
MIDEH	Honduran Improving Student Achievement Project (Honduras)
MOE	Ministry of Education
NCEEE	National Center for Examinations and Educational Evaluation
NIED	National Institute for Educational Development (Namibia)
NMOE	Namibian Ministry of Education
NSAT	National Student Achievement Test (Namibia)
RFA	Response for Applications
TIMSS	Trends in International Math and Science Study
UNESCO	United Nations Educational, Scientific and Cultural Organization
USAID	United States Agency for International Development

Increasingly, donors, MoEs/national education authorities, and educators are realizing the importance of measuring student learning outcomes and how such measurement can be used to transform education practices and improve student learning. Hence, national assessments are used more and more to provide an accurate representation of student performance at a given age or grade level. This performance information may then be used to identify students' needs and provide key stakeholders with an independent evaluation of what students are learning to help them make informed decisions. Also, student performance data obtained through national assessment programs may, over time, provide developing countries a means to gauge progress towards donor benchmarks and international standards. Some countries have used the results of their respective national assessments as part of their school accountability models and to generate support for targeted policy initiatives.

As part of the EQUIP2 lessons learned activity on student learning outcomes, we take a close look at projects that have attempted to help governments in developing countries institutionalize national assessment programs to obtain valid information about what students are learning in schools. Although each project reviewed was implemented differently, the procedures undertaken by the projects were similar, in part because there are specific development activities that have to be conducted to ensure that effective assessment systems are established and implemented.

This paper examines the efforts of EQUIP Associate Awards in four countries to establish a national system for measuring student learning outcomes and the lessons learned from these experiences. The EQUIP Associate Awards included in this review are: Egypt's Education Reform Program; Ghana's Basic Education Comprehensive Assessment System; Honduras' Honduran Improving Student Achievement Project; and Namibia's Basic Education Support III and National Student Achievement Test. Although the assessment development and implementation activities deviate very little between projects, the level of government involvement, local buy-in, funding, plans for sustaining the work, and subsequent outcomes attained vary tremendously by project.

EGYPT: Education Reform Support Program (ERP)**Funding:** Initial Award: \$1,871,548; Final Award: \$51,568,490**Time frame:** 2004–2011

To bring about reform at the school level, the Education Reform Program (ERP) worked to increase the level of data-driven decision making, measurement of learning outcomes, sharing of new knowledge, and more effective teaching practices to empower schools and communities to direct their own educational change process. The school-based reform effort was built on the partnership among the government, schools, and civil society to provide the skills and knowledge needed to effectively implement the Egyptian national standards for students, teachers, and schools. Student learning outcomes were intended to be used as a progress indicator to measure the effectiveness of the teacher interventions that were conducted through ERP. However, as the project evolved, the assessments were used to measure end-of-grade learning throughout Egypt as well.

GHANA: Basic Education Comprehensive Assessment System (BECAS)**Funding:** Initial Award: \$999,839; Final Award: \$1,284,168**Time frame:** 2004–2007

The Basic Education Comprehensive Assessment System (BECAS) project was designed to provide technical assistance to the Ghana Education Service (GES) to develop and implement a comprehensive educational assessment system. The standardization and alignment of student learning and achievement assessments to the Education Strategic Plan to ensure valid measures of student performance were key objectives of BECAS.

HONDURAS: Honduran Improving Student Achievement Project (MIDEH)**Funding:** Initial Award: \$9,173,629; Final Award: \$20,141,183**Time frame:** 2004–2011

The Honduran Improving Student Achievement Project (MIDEH) is composed of an integrated set of testing, research, technical support, and capacity building activities that address major components of the Honduran Ministry of Education's (MOE) national education program, including standards and assessments. With local educator involvement, technical assistance through MIDEH was provided to the government to develop content standards, teacher guides, teacher training modules, supervisor support systems, and a standardized testing system with summative and formative assessments to increase student learning outcomes for meeting Education for All (EFA-FTI) goals on student achievement. Key objectives of the project included providing technical assistance, training, and related support for the Honduran MOE, local education institutions and educators to increase student achievement in Spanish and mathematics; and reducing student failure and dropout rates while improving student flow rates and access to instruction in the 7th–9th grades. One of the vehicles to bring about change that was envisioned through MIDEH was the development of an effective national testing system predicated on the existence of an appropriate system of standards composed of aligned curriculum, benchmarks, and indicators.

NAMIBIA: Basic Education Support III (BES III)**Funding:** Initial Award: \$14,116,537; Final Award: \$13,225,437**Time frame:** 2005–2009

As the name suggests, Namibia's Basic Education Support III (BES III) project was the third installment of BES I and II, with BES III activities building upon the successes of the earlier BES work. The purpose of measuring student learning outcomes in this project was to evaluate the improvement of classroom instruction, and ultimately student learning, as a result of the BES III interventions. The project team carefully reviewed the student learning outcome data regularly to inform implementation decisions pertaining to teacher training and development of materials. The collaborating partner from the Ministry of Education on BES III was the National Institute for Educational Development (NIED).

NAMIBIA: National Student Achievement Test (NSAT)**Funding:** \$878,163**Time frame:** 2009–Present

Given BES III's successful implementation of the student assessments of student learning outcomes, MOE (which was at the time interested in establishing a census-based national assessment) took an increased interest in the BES III assessment model. Although BES III ended in 2009, USAID continued to fund technical experts in assessment to assist with Namibia's efforts to develop annual national assessments within the framework of the National Standardized Achievement Test (NNSAT) project, for grades 5 and 7 to be used as a diagnostic tool for improving education quality. The Directorate of National Examinations and Assessment (DNEA) is responsible for developing and implementing these assessments

METHODOLOGY USED TO CARRY OUT THIS REVIEW

This review used a qualitative approach to better understand the why and how of decision making, not just what, where, and when. The main methods used in the study included interviews and document reviews. It is important to point out that this review is not based on an in-depth evaluation of each Associate Award, i.e., visits were not made to each country to interview a wide variety of individuals nor was an extensive review made of documents generated by the Associate Award or related to the Associate Award.

PREPARATION OF INTERVIEW PROTOCOL, SUMMARY /MATRIX, CONSENT FORM FOR EACH PROJECT

To prepare for this review an interview protocol was developed and piloted. In addition, a summary and matrix for each country case and consent form were developed. The information for the country summary and matrix was taken primarily from the RFA and corresponding proposal to be used as a reference point during the interviews.¹ Key topics raised in the interview protocol) may be found in the textbox below.

TOPICS ADDRESSED IN INTERVIEW PROTOCOL

1. EQUIP2 project's development hypothesis (or what you wanted to accomplish related to the project goal) related to the national assessment program and the assumptions underlying the hypothesis.
2. Key project activities related to the national assessment program: what they were; why they were selected; the assumptions linked to the activities and their validity; whether the activities led to the outcomes you expected; if not, why.
3. Adequacy of time frame and funding for what the project wanted to accomplish related to the national assessment program.

¹ The summary and matrices for each project contain information on: life of project funding, project start and end dates, the country and education context, role of other donors, the project purpose and key activities.

4. Extent to which the project built in sustainability, the extent to which sustainability was achieved (or not) and why.
5. Whether the project led to outcomes that were expected and, if not, why.
6. Adjustments made, if any, to: project activities, budget, and timeframe.
7. Project monitoring and evaluation: indicators selected to assess project impact and track activity progress in the national assessment program; which were most useful and why; how the information collected was used; are there other indicators that would have been more useful.
8. Successes and challenges related to implementing the assessment program: aspects of the project that were most successful and why; biggest challenges encountered in managing the project and how addressed.
9. Ability to adapt to changing circumstances/as needed reprogram or change aspects of the program.

INTERVIEWS CARRIED OUT USING THE PROTOCOL AND SUMMARY DOCUMENTS

The protocol was used to carry out interviews of approximately an hour and a half each. Between four and seven individuals (e.g., USAID staff involved in the design of the Requests for Applications (RFAs) and in overseeing implementation of the cooperative agreements, project staff involved in implementing the Associate Award and providing backstop support from the United States, and, where possible, host-country counterparts involved in implementing the Award) were interviewed for each country. A total of 23 individuals across the projects in the four countries were interviewed (see Table 1). A majority of the interviewees were project staff based in the United States or out in the field. As needed, some individuals were interviewed a second time to clarify points from the first interview, obtain additional information that was not sought out in the interview, and/or to triangulate information obtained from other interviewees.

Table 1: Total number of people interviewed by country and affiliation²

	Total	Project Staff	MOE	USAID
Egypt	7	5	2	2
Ghana	4	3	0	1
Honduras	7	4	0	3
Namibia	5	5	0	0
Total	23	17	2	6

² Note: MOE representatives from Ghana, Honduras, and Namibia did not respond to repeated requests for interviews for this retrospective study.

After all the interviews were completed, an analysis was conducted to examine the decisions/actions that contributed to successfully implementing national assessment programs and those that acted as barriers to implementing them, and why/how these actions added or took away from the implementation process. These actions were then condensed into the following lessons learned.

THE TEST DEVELOPMENT PROCESS

ESTABLISHING LEARNING EXPECTATIONS

Developing a national assessment program can be an extremely costly endeavor that requires adequate time to mature. To ensure that the national tests provide valid and reliable measures of what students know and are able to do, there are recommended steps and procedures that should be adhered to during the test development and administration process. For instance, prior to beginning any test development effort, test developers need to be clear about the content to be measured. Thus, the subject-specific learning expectations that define what students should know and be able to do at the end of the grade have to clearly describe grade-level appropriate knowledge and skills. Since these learning expectations drive education practices throughout the system, they have to be thorough and age-appropriate for students as well as comparable to international standards, if countries want to be certain that their students are competitive in the global market.



To properly develop these learning expectations—if they don’t already exist—subject-specific panels of educators (teachers, principals, university faculty) from around the country should be brought together to either write the expectations or review and comment on expectations generated by a team of experts who are experienced in this area. If learning expectations already exist, they still need to be reviewed for “assessability” and optimal weight in a standardized testing operation. Bringing together a panel of educators from various parts of the country in the development process ensures that

the content covered in these expectations are representative of what is being taught throughout the country and is not biased towards specific regions. Ideally, there will be several rounds of review meetings before the expectations are finalized and ready for national distribution.

DEVELOP TEST ITEMS

The primary objective of the test development process is to ensure that the test items accurately and reliably measure the knowledge and skills identified in the learning expectations, as opposed to irrelevant, ancillary skills. Hence, there are specific practices that should be built into the test development process to maximize test item alignment with the learning expectations and test reliability. For example, test developers should actively involve teachers from around the country in the entire test development process. These teachers should demonstrate deep and solid content expertise in their respective subject areas, regardless of their years of teaching experience. In countries where national assessments are being employed for the first time or are relatively new, teachers have to be trained to write test items—by experienced item writers. Such trainings typically last about two weeks and involve introducing teachers to the learning expectations, to the skills and knowledge students have to demonstrate to be considered successful in having achieved the expectations, and to provide guidelines for writing test questions. Since item writing is a skill that is developed over time, teachers will initially create items that are not of sufficient quality, hence the test items that are created during the training will have to be reviewed and edited by an experienced item writer before they can be pilot tested.

CONDUCT PILOT TESTING

The items identified for pilot testing are placed on test forms and administered to a representative sample of students. Several test forms are typically created for pilot testing, since test developers want to pilot more items than they need for the actual national assessments. Hence, the test forms need to be carefully designed so that subsequent statistical comparisons can be made across items administered to different students. The sample size of students selected to participate in the pilot test will vary from country to country, and may include anywhere from hundreds of students to thousands of students, depending on the national population that will eventually participate in the national assessment. To administer the tests to students across the country in a uniform manner, test developers employ and train

individuals who serve as data collectors whose primary responsibility is to administer the pilot tests to students according to standardized procedures.

After the pilot test administration, individual student responses are entered onto an electronic spreadsheet so that psychometric analysis can be performed to inform the test developers if the test items successfully captured student learning as expected. For test items that do not perform acceptably, test developers may choose to bring back the panel of teachers who wrote the items to review the pilot test data and make edits to the items so that these items can be used on test forms for future test administrations of the national assessments. For example, if a test question has two correct answers or an ambiguous correct answer, it will be evident from the data analysis reports since high achieving students will perform poorly on this item. Under these circumstances, the test developers will explain to the teachers what was problematic about the item so that the teachers can revise the item accordingly. The newly revised items will then be sent out for pilot testing before it is included on an operational test form. Reviewing poor performing piloted items is a step that is typically skipped in the event that sufficient items pilot tested well for an operational test form.

CONSTRUCT AND ADMINISTER OPERATIONAL TEST FORMS, PROCESS STUDENT RESPONSES, AND DEVELOP SCORE REPORTS

When there are sufficient items to create operational test forms, test developers should involve a small group of teachers in the construction of the tests. The role of teachers in this activity is to serve as content reviewers before the test forms are sent out for printing. Since only a limited number of items can be included on the test forms, ensure that the items selected for the national assessment represent content that is important for students to master at the end of the grade and that they are error free so that they elicit accurate student responses. Once a test form is finalized, it can be sent for reproduction, packing, and delivery to schools.

After the test is administered, student responses are transported from the school to a central location for processing and scoring so that test reports can be subsequently produced and sent out to the schools. However, for these test reports to be accessible and actionable for educators they should not only contain the total score or percent correct score of students but also performance levels. To produce those, standards must be set by teachers to judge the levels of student performance on the test that meet the criteria

for being considered excellent, proficient, basic, or below basic for a specific grade.

Overall, the test development process may take from nine to eighteen months, depending on the staff resources available and academic schedule for pilot testing. Establishing a national assessment program, on the other hand, may take longer since it involves acquainting stakeholders with the assessment process, as well as the understanding and application of data into decision making to improve daily practices. Furthermore, building the capacity of the local staff to independently perform psychometric analyses and provide item writing training to local educators requires time to develop.

Given these considerations, the cost associated with developing a national assessment program can be considerable. The development process involves the participation of numerous educators from around the country, the production, delivery, and scoring of test forms for pilot testing and operational test administrations for every participating student, the production and delivery of test reports. Educators throughout the country must be trained to understand the score reports and what the results mean about student learning.

LESSONS LEARNED FROM THE EQUIP ASSOCIATE AWARDS

The analysis of data collected over the course of our study has clearly reinforced the overarching message that the implementation of a national assessment program requires time, staff, and financial resources, host government commitment, and local buy-in to achieve required results and sustainability.

As the following discussion will illustrate, the responsible stewardship of donor funds may require a new, two-step approach for USAID's efforts to bid projects and contract with providers who will work with developing countries on the development of their assessment systems. Each RFA opportunity should begin with a feasibility study of the readiness of a nation, including its government commitment and education infrastructure, to successfully develop and implement national assessments. The lessons learned from this study, which are described in the following sections, illustrate the features that must be examined prior to funding an assessment implementation initiative to maximize the likelihood of a successful project.

Only after the data from this feasibility study of the necessary preconditions has been collected and analyzed should the main scope of work for the project be conceptualized and bid for contract. The results of the feasibility study may indicate that this is not the right time to move forward with the main project goals at all, but that funds should instead be directed toward work that will establish and reinforce whichever preconditions are still absent. But at the very least, a feasibility study will have an impact on the way the project is conceptualized for a bid by enriching the stated project requirements for the main scope of work with conclusions that were gathered from the work of this preliminary study.

Data from our study have identified several key lessons that USAID can learn from the assessment initiatives examined. A careful consideration of these lessons will enrich USAID's outlook for future funding opportunities

and will help clearly establish the preconditions that will be necessary for the successful implementation of future assessment initiatives.

1. ESTABLISH A CLEAR DEVELOPMENT HYPOTHESIS FOR THE NATIONAL ASSESSMENT PROGRAM THAT TAKES INTO ACCOUNT THE PRECONDITIONS NECESSARY FOR SUCCESSFUL IMPLEMENTATION

The development hypothesis of a project outlines what the designers hope to achieve through a project and how they intend to attain the specified objectives. Typically, the main goals of a national assessment program are to establish what students are learning in the classroom to inform the work of educators and policymakers as they determine whether programs are effective and identify the necessary changes to improve existing education practices. Towards this end, assessment specialists are recruited to design a system for monitoring and evaluating student performance and to develop assessments that provide valid and reliable measures of what students know and can do.

This approach to establishing a national assessment program is too narrow, however, and does not take into account the preconditions that need to be in place for such a program to thrive. To sustain any national assessment effort successfully over a long period of time, USAID and its implementing partners must consider the willingness of the host country governments to make long-term financial commitments, the capacity of Ministry of Education (MOE) staff to maintain the technical quality, and the ability of local educators to understand the assessment results and their implications on curriculum/instructional planning and implementation. The absence of any of these preconditions will require a broadening of the development hypothesis to also account for the need to complete the pre-work necessary to implement successfully project goals. For example, if governments are unable to make long-term financial commitments to the assessment program, then the donor and implementing partner need to consider providing additional assistance to help the government identify/secure additional sources of funding. If teachers require training to interpret and understand the assessment results, then steps must be taken to ensure that teachers, MOE, and other key stakeholders receive adequate training.

Two of the projects (ERP and BES III) reviewed were not built around development hypotheses that were specific to establishing national assessment programs. The development hypotheses for these two projects were oriented instead toward larger education initiatives. For these projects, the assessments

were initially developed in close collaboration with the respective MOEs to serve as monitoring and evaluation (M&E) tools. However, during the course of this collaboration, the MOEs for each country decided to adopt the assessments nationally for systemwide monitoring and evaluation.

On ERP, the M&E assessments were converted into national assessments with limited consideration or planning for how the government would sustain the program over the long term. Thus, after the assessments were administered and the student responses were analyzed little was or could be done to disseminate the results to the education community or to provide training around data interpretation and use since there was no infrastructure or financial resources to support this work at the national level. As a result, there was a sense that although ERP had demonstrated tremendous success in developing technically sound assessments and building the capacity of local counterparts at the National Center for Examinations and Educational Evaluation (NCEEE), the effort to build a national assessment program could not be viewed as successful holistically because the ability to provide feedback about student learning was lacking in this initiative.

Unlike ERP, the efforts of the Namibian Ministry of Education (NMOE) to adopt the assessment model developed by BES III were more strategic and led to the funding and creation of a separate project, the National Student Achievement Test (NSAT). As part of the NSAT project, NMOE developed a separate set of objectives specific to implementing a national assessment program. Given these objectives, NMOE not only outlined key activities that need to be accomplished under NSAT, but also identified the barriers to establishing an effective national assessment program. With a clear understanding of the conditions that encourage or hinder proper implementation, NMOE has developed a fairly comprehensive implementation plan that not only takes into account the procedures for creating quality assessments, but also sources of funding, staff capacity building needs, and training for subject specialists, principals, and teachers for the sustained and successful implementation of NSAT.

The assessment programs for Ghana and Honduras were developed from RFAs that explicitly called for the development of national assessments. It is unclear if USAID took into consideration the preconditions discussed above when drafting the development hypothesis for each project. Although both assessment programs began with the same goal, i.e. to use the assessments as a monitoring tool to help donor groups track the country's progress towards Education for All–Fast Track Initiative (EFA–FTI) goals, the implementation approach taken by the projects in each country ended up being somewhat different, not because of the level of technical expertise on the part of the

implementing partners but because that the work was conceptualized differently by the USAID missions in each country.

It is unclear whether a development hypothesis was developed for the BECAS project. The evidence obtained from interviews suggests that if there was a development hypothesis, it probably was not well thought through. On page 5 of the BECAS RFA, the goals of the project were defined as “to systematize the assessments so that their information is clear, non-contradictory, and easily used for diagnosis at the level intended.” Priority on the project was given to achieving defined sustainable results through improved institutional capacity at the central, regional, and school levels. The scope of the BECAS work was narrowly conceived and not grounded in the larger context and needs of Ghanaian education. This could explain why the amount of funding and time dedicated to accomplishing the project was sparse and inadequate.

In contrast, the project in Honduras used the student assessment system to facilitate and drive the attainment of larger education goals, such as “technical assistance, training, and related support for the Honduran MOE, local education institutions, and educators to increase student achievement in Spanish and mathematics, reduce student failure and dropout rates, improve student flow rates, increase access to instruction in the 7th–9th grades and secondary education to achieve the goals and indicators of USAID’s Regional Strategy in Honduras, the Poverty Reduction Strategy, and the EFA–Fast Track Initiative” (extracted from the MIDEH RFA). Because the assessment effort was part of a larger systemic initiative, its development and implementation was rooted and grounded in other education policy initiatives in Honduras, allowing it to grow over time as a result of increased funding from USAID and the Honduran government, which recognized that specific activities could not occur or be as effective if the student assessment system was not effective.

2. BE CLEAR ABOUT THE PURPOSE OF THE NATIONAL ASSESSMENTS. TESTS THAT ARE CREATED FOR PROJECT MONITORING AND EVALUATION DO NOT NECESSARILY MAKE EFFECTIVE NATIONAL ASSESSMENTS

Prior to establishing an assessment program, it is critical to establish an understanding of what the assessments will be used to measure and for what purpose. This understanding is central to developing assessments that yield valid and reliable test score interpretations. Tests that are created for one purpose and used for another typically produce invalid results.

Of the projects reviewed, only the assessments developed through the ERP project lacked clarity of purpose. The assessments were initially developed to provide measures of improved student learning as a result of teacher training. However, during the course of the project, the Egyptian Ministry of Education (EMOE) expressed interest in using the assessments as a measure of student learning in the classroom throughout the school year. Using the same assessments for two different purposes was problematic given the different outcomes intended to be measured. Presumably, an assessment evaluating the impact of teacher training on student learning would assess skills and knowledge relevant to the constructs covered by the training. An assessment to examine what students have mastered throughout the year would assess skills and knowledge deemed essential for students at a specific grade as well as for promotion to the next grade.

According to several sources interviewed, the assessments were better suited as a national assessment for measuring student mastery of grade level content rather than as a monitoring and evaluation tool for the project to measure the impact of teacher interventions on students' critical thinking. In fact, the data showed that students attending the control schools whose teachers did not receive any training performed better than students who were part of the intervention group. There may be other factors that could have contributed to the contradictory findings, such as poor sampling procedures.

In contrast to ERP, the developers of NSAT conducted separate test development activities from BES III although Namibia's NSAT was informed by the work previously conducted on the BES III. The decision to conduct separate test development activities allowed the Directorate of National Examinations and Assessment (DNEA) to produce tests that measured the knowledge and skills that NMOE deemed important and appropriate for students in grades 5 and 7 to have at the end of the school year. The reactions from local educators to date suggest that they feel that the assessments are fair and appropriately reflect what students are expected to know.

3. ESTABLISH GOVERNMENT COMMITMENT AS A COLLABORATING PARTNER FOR THE NATIONAL ASSESSMENT PROGRAM PRIOR TO ISSUING THE RFA OR RFP. CLARIFY, AS MUCH AS POSSIBLE, HOW THE NATIONAL ASSESSMENT PROGRAM WILL BE AFFECTED IF THERE IS A CHANGE IN LEADERSHIP AT THE MINISTRY OF EDUCATION

The MOE has an integral role in most national assessments. Even when the MOE appoints an outside agency to implement the assessment, the host

country's Ministry remains responsible for identifying policy needs, the student population to be assessed, and the subject area(s) to be assessed, and in most cases curricula and/or learning standards. To take full advantage of national assessments (i.e., to influence national education policy), the results must be disseminated promptly to all relevant stakeholders. Stakeholders are then in a position to use the results of the assessment to make key decisions—both on the national and individual school levels—designed to improve student learning outcomes. Decisions may also be made about resource allocation, teacher training needs, and establishing benchmarks. For national assessments to continue over time, key stakeholders in the host country (such as the MOE) must see the utility of conducting them.

The projects reviewed suggest that getting government buy-in and commitment for a national assessment program is crucial to the successful implementation of the assessment-related activities, and ultimately, the continuation of the effort after the project ends. Because of the expense and resources required to maintain a national assessment program, it is important for the government to be committed toward maintaining such a large initiative. Without the government commitment to continue with the national assessment effort, the program could not continue to thrive or to sustain the support of local educators.

Through our interviews, we found that the more engaged the government was in the process, the more committed they were to ensuring the success of the program. For example, in Namibia, when the NMOE decided that they wanted to build a national assessment program, they reached out to BES III in 2008 and expressed a need for technical assistance in developing a national assessment system. This request came following an extensive review by NMOE to examine the different assessment models employed around the region and the decision that the BES III model best fit the needs of the Namibian context. To ensure a successful national assessment endeavor, the NMOE set aside financial resources, in addition to the USAID funding for technical assistance, to ensure a dedicated team of staff at the DNEA was available to support and learn about the assessment development and administration process, and cover the cost of developing and administering the tests nationally to all students in grades 5 and 7. NMOE and DNEA have been, and continue to be, active partners in securing additional funding for technical assistance and championing the expansion of its national assessment program to other grades.

Similarly, in Honduras, there was a lot of government commitment and buy-in for establishing a comprehensive assessment system from the outset when the RFA was being drafted. The Ministry of Education in Honduras

(HMOE) was heavily involved in designing the education reform strategy, which included the implementation of a national assessment system. Furthermore, HMOE recognized that a successful national assessment program was a means for demonstrating achievement on its EFA goals. Because HMOE was so heavily invested in the success of the program, they were open to new ideas and flexible about adding components that were not part of the original design outlined in the RFA, such as learning standards and pacing guides. HMOE continues to be a strong partner in the national assessment effort and is working with the implementing partner to explore the possibility of administering the Spanish and math tests on a census basis to students in grades 1–6 through school-administered end-of-year assessments.

As in Honduras, the assessment results obtained from BECAS in Ghana were intended to be used as an indicator of progress towards the country's EFA goals. However, the interest and ownership of the assessments and student results expressed by the Ministry of Education of Ghana (GMOE) gradually waned over time, beginning with the departure of a key champion of the effort from GMOE. With little GMOE support, the assessments administered under BECAS began to be regarded largely as initiated by the donor community for accountability purposes. Although GMOE had assigned the responsibilities of test development and administration to Ghana Examination Services (GES), there was no real concerted effort by the staff to support or learn about the assessment process. The failure of the GMOE's continued support created some challenges around buy-in at the school and local levels. Thus, when funding for BECAS ended, the national assessment responsibilities were transferred over to another USAID-led education initiative, Education Quality for All (EQUALL).

Although the assessment efforts in Egypt began with tremendous support from the Minister of Education, it lost momentum with the appointment of a new Minister in 2010. The former education Minister not only understood the importance of a national student assessment system but also supported the efforts to adopt ERP's student assessments. Under the former Minister's leadership, the NCEEE was tasked with collaborating with assessment experts from ERP supporting work at the national level. However, when the current Minister came to office the assessment work came to a screeching halt and the momentum gained up until that point was lost. Making matters worse, the NCEEE staff trained in the technical aspects of test development and statistical analysis were appointed to other administrative positions at the EMOE and replaced with staff who lacked proper psychometric experience.

The experiences of BECAS and ERP underscore the importance of establishing a transition plan with local Ministries of Education in the event of leadership change, so that huge investments that are made to develop an assessment program are not lost. However, not all changes in government leadership lead to the dismantling of national assessment programs. Despite the coup d'état in the summer of 2009 to oust the president of Honduras, the national assessment efforts in Honduras survived and continued to thrive after the situation normalized. The resilience of the assessment efforts in Honduras may be due to the careful planning of the implementing partner to engage local support for the assessment system, which is discussed in detail in the next lesson. The degree of implementation of the assessment programs in Honduras, Ghana, and Egypt were very different and may explain why the efforts in Honduras continue to gain ground.

4. PRIORITIZE WINNING LOCAL SUPPORT (BEYOND THAT WHICH THE MINISTRY OF EDUCATION PROVIDES) FOR THE NATIONAL ASSESSMENTS AMONG STUDENT, TEACHERS, PARENTS, HEAD TEACHERS, AND DISTRICT AND PROVINCIAL ADMINISTRATORS EARLY IN THE PROCESS. SHOW ALL EDUCATION STAKEHOLDERS HOW ASSESSMENTS CAN PROVIDE USEFUL INFORMATION ABOUT STUDENTS' LEARNING AND FACILITATE DECISION MAKING

One of the key components to building a successful national assessment program is ensuring that the results and conclusions drawn from the assessments are shared with the education community at large. As assessment results are shared with the education community, it creates an opportunity for educators to finally realize the pay-off for the hard work that goes into the developing and administering technically sound assessments. It becomes evident that sound assessment results can inform and enrich the curriculum and instruction provided and have the power to uncover strengths and weaknesses in student outcomes and pockets of uneven outcomes among clearly identified sub-populations (e.g. by gender, geographic background). The utility of the results creates widespread support and buy-in for assessments among the education community, establishing and reinforcing an ethos in using assessment in concert with curriculum and instruction. Assessments may then become part of the normative framework for education practice, creating sustainable patterns of reliance on assessment as an integral part of an effective overall education system. As this perspective takes hold in the education community, it will no longer be necessary for an

assessment initiative to be pushed forward by government agencies, but will instead become an organic part of standard education practice.

A plausible reason why the national assessment program in Honduras continued to thrive after a break in implementation activities following the 2009 presidential coup d'état was the widespread educator support for the program throughout the country. At the beginning of the initiative, there was a lot of educator resistance to establishing a national assessment program in Honduras. In fact, the teacher unions voiced the loudest opposition. Undeterred, MIDEH put a lot of effort into changing teacher perceptions towards the assessments by making sure that presentations introducing them were conceptually sound and included value-added, high-quality materials that teachers could take back into their classrooms.

The project was also mindful of including teachers from various regions of the country on as many assessment development activities as possible. Furthermore, after the tests were administered, the project provided training to help educators understand the data reports and degree of remediation needed to improve student learning. When teachers understood the importance of the assessments and what the project sought to do, many more teachers expressed interest in participating in the assessments program. With the increased support and buy-in by local educators in Honduras who recognized the value added to classroom instruction (especially from the formative assessments developed by MIDEH alongside the end-of-grade assessments) any subsequent discussion of discontinuing the national assessment program would have met with resistance from local educators.

The efforts to involve teachers in assessment-related activities in Egypt was, perhaps, not as strategically planned as those in Honduras in large part because the assessments were initially created for project monitoring as opposed to national assessments. As a result, it was not as important for ERP to gain stakeholder buy-in by involving key education leaders or teachers in test development, and attempts to train teachers in how to use the assessment results in the classroom were sporadic and inconsistent. Consequently, there was little support for the national assessments at the school level, which probably made it easier for the new Minister to dismantle the program and re-assign the NCEEE staff to other positions with very little resistance from the education community.

Likewise, the information gathered about BECAS suggests that the project staff faced numerous challenges with implementation because of the limited buy-in at all levels of education for the national assessment efforts in Ghana. At the local level, principals, as well as district and regional education officials

were suspicious of the BECAS assessment efforts primarily because they were concerned about the potential sanctions that would be placed on them if student performance turned out poorly. And at the national level, there was some difficulty in getting the GMOE to acknowledge the importance of the student results. GMOE officials questioned the considerable financial investment into a national assessment program that did not yield information that they were not already aware of, i.e., that a majority of Ghanaian students were not performing well academically.

With the lack of GMOE support, the results from the BECAS assessments were not widely shared across the education community. Educators at the school, district, or regional levels were not provided with comprehensive training regarding data interpretation and use of data for planning and decision making. The inability to share the data and explain the benefits of using assessment data for informing instructional planning and teacher training further impeded efforts to establish buy-in at the local levels. The implementing partner recognized this limitation and developed a strategy for disseminating the assessment results to the schools. Although this plan was discussed with USAID, it was never funded. Contrary to the reactions expressed by GMOE and the school community, donor groups were very pleased to obtain concrete data to substantiate the impact of the interventions they were funding. It is likely that without pressure from the donor groups on GMOE to initiate the national assessment efforts these assessments may never have been administered in the first place given the lack of support from local stakeholders.

Contrary to events in Ghana, the BES III project was successful in winning local educator support in the regions where it was implemented. Educators participating in BES III were trained on how to use the assessment data to improve teaching and in-service training. At the time, using the assessment data to improve teaching was a new concept for educators in Namibia and perceived as a value-added activity that helped teachers and regional education specialists to make meaningful decisions that improved student learning in the classroom. As a result, the application of assessment data for informed decision making received widespread support from participating teachers, head teachers, and regional subject specialists, which eventually caught the attention of officials at the NMOE and led them to dedicate funds to scale up this model nationally through NSAT.

Although NSAT is still in the midst of rolling out its assessment program, early indications are that teachers and principals support the decision to administer a national assessment. The assessments in grade 5 were administered for the first time in 2009 and NMOE set performance

standards for these assessments in 2010. Shortly after the assessment reports were sent to each elementary school in Namibia, DNEA conducted training workshops in all 13 regions to introduce principals to these reports to show them how to use the results with their teachers in grades 4, 5, and 6. Principals were instructed to share the 5th grade test results with 4th grade teachers to help them understand what students need to master at the end of grade 4 to be ready for learning in grade 5. The 5th grade test results were shared with 6th grade teachers so that they would understand the level of knowledge and skills 5th grade students mastered upon entering 6th grade. This training model will be repeated in 2011 with the grade 7 NSAT tests. Currently, there is an independent evaluation taking place to find out if the principal training efforts were successful and whether principals are using the NSAT data with their teachers to improve instruction in Namibian schools. It is unknown at this time the degree of success NMOE has achieved in obtaining local support for the NSAT.

5. ENSURE THAT THE GOVERNMENT IS CLEAR ABOUT THE ASSESSMENT DEVELOPMENT PROCESS AND PROVIDES THE NECESSARY RESOURCES TO SUPPORT THE WORK. IDENTIFY CLEAR ROLES AND RESPONSIBILITIES FOR DEVELOPING AND IMPLEMENTING THE ASSESSMENTS

As mentioned before, the process of developing and validating an assessment involves a series of activities that include identifying the purpose of the assessments; identifying and operationally defining the construct to be measured; writing items to measure the behaviors, knowledge, and skills that are associated with the construct; pilot testing the assessment items to ensure that the test-takers understand the test directions and questions; and refining the assessment based on analysis from the pilot test data. Thus, given all the complexities associated with test development, it is important that the cooperating Ministries of Education understand the development process and what the work will require in terms of staff, time, and resources. Mainly, ministries need to appreciate that they not only need to assign sufficient numbers of dedicated staff to the project, but that these individuals need to have the right qualifications and background to carry out assessment-related activities. Ministries will also need to commit to accomplishing specific test activities in certain timeframes as test development is fairly linear and deadline-oriented. Failing to complete agreed upon activities will jeopardize the completion of other activities.

The assessment development process in Honduras and Namibia were completed with strong host government support, and, for the most part, in Egypt as well. Conversely, the development efforts were more challenging in Ghana. The work in Ghana was largely hampered by a misunderstanding on the Ghana Examination Service's part regarding the distribution of USAID funding. GES was under the impression that the award would go directly to them, and they would, in turn, use the money to pay the implementing partner. But when the funding was not allocated according to GES expectations, it soured the relationship between the implementing partner and the GES, and made any effort to obtain government buy-in or cooperation difficult from that point forward. This strained relationship may explain why a group of full-time GES staff was not assigned to support the BECAS assessment development.

Although GES did assign five staff to BECAS, their responsibilities on other tasks were not reduced. Hence, to these five individuals, BECAS was regarded as added work without a reduction in existing responsibilities or an increase in salary. Given the heavier workload, lack of compensation, and minimal GES buy-in, this group of five staff members was not inclined to make BECAS a priority or committed to making it a success. In fact, the GES staff attended training sessions only when per diems were provided. It is unclear how the lack of commitment by GES staff to support the national assessment program limited day-to-day activities, but over the long haul, it did hamper efforts to build local capacity to sustain the work.

Unlike in Ghana, government institutions in Egypt and Namibia were heavily involved in the development and administration of the national assessment. In both countries, it was the staff provided by the respective MOE who managed and coordinated all the assessment activities, with some support from technical experts provided through USAID funding. Specifically, in Namibia the NMOE created a new division within DNEA and hired new staff to support NSAT. The NSAT staff coordinate all the test development, preparation, and training activities for five subject tests across two grades. Similarly, in Egypt a subset of NCEEE staff were dedicated towards all assessment efforts for the national tests.

6. BE REALISTIC ABOUT THE FEASIBILITY OF BUILDING CAPACITY OF THE LOCAL STAFF TO SUSTAIN THE ASSESSMENT EFFORTS AFTER PROJECT FUNDING ENDS. CAPACITY BUILDING MAY NOT BE POSSIBLE IF THE STAFF DEDICATED TO THIS WORK DO NOT HAVE THE RIGHT QUALIFICATIONS OR BACKGROUND

As USAID increasingly adopts the strategy of building capacity among local educators to sustain the investments it makes, it needs to be realistic about the feasibility of implementing this model as it relates to national assessments in specific countries. Developing and implementing a national assessment program requires local staff to acquire technical knowledge for item development, including content understanding and data analysis skills, as well as logistical and operational knowledge for test administration. Although it is easier to impart knowledge to local staff about the logistics and operational steps of test administration, teaching the same group about the technical aspects of item writing and data analysis takes longer and assumes that the individuals assigned are equipped with a specific set of prerequisite skills (e.g., familiarity with age-appropriate content for test subjects and rudimentary statistical knowledge and skills).

The effort to build a national assessment in Egypt would have been considered fairly successful if not for the reassignment of the assessment technical staff and statisticians to other positions in the EMOE. The initial team at NCEEE not only supported the effort to develop the tests for student evaluation, but went on to apply their skills and knowledge, with minimal help from external assessment specialists, to develop the Teacher Cadre assessments in Egypt. One interviewee commented that the successful development and implementation of the Teacher Cadre assessments was a testament to the successful efforts to build the assessment capacity of the local staff in Egypt.

There are several reasons why capacity building of the NCEEE staff was a successful endeavor: (a) NCEEE oversaw several cycles of test administration for the Trends in International Math and Science Study (TIMSS) and, therefore, had several years of experience with large scale test administrations and was familiar with basic psychometric concepts; (b) some of the personnel at NCEEE held advanced degrees in social science and statistics and at the very least had basic knowledge of quantitative and qualitative research methods, data collection, and statistical analysis; and (c) prior to collaborating with ERP, NCEEE worked with other donor groups to complete evaluation projects. The combined experience, knowledge, skills, and motivation of the NCEEE staff most probably facilitated the acquisition

of assessment concepts after a single development cycle of the national assessments and development of the Teacher Cadre assessments.

The situation in Namibia is somewhat similar to Egypt. The primary responsibility of DNEA prior to NSAT was to oversee the development and large-scale administration of the national school examinations across the country.³ Because of this familiarity with the national examinations, the DNEA staff had the right background and experience that allowed them to grasp psychometric concepts and processes relatively easily and play a critical role on NSAT development. The staff at DNEA works closely with a technical consultant for assessment who visits Namibia on a regular basis to refine and streamline the procedures for assessment development and administration, and set performance standards for grades 5 and 7.

Interviewees reported that the capacity of MOE staff in Ghana was low despite attending a training program about assessments at the UNESCO International Institute for Educational Planning (IIEP). Interviewees concluded that it was probably the lack of interest and commitment to the work that affected the way staff applied their knowledge and skills to the BECAS activities. As a result, the implementing partner had to rely heavily on the local university personnel who were trained in assessment at universities in the United States to assist with test development and data analysis, while individuals who were involved in previous assessment efforts in Ghana were recruited to help with logistics, test administration, and scoring. The implementing partner recommended to USAID that the responsibility of sustaining BECAS be transferred to a consortium of local universities with the technical expertise and interest to continue with the assessment work; however, that plan did not materialize as GES wanted to retain authority and control over the assessments. Although BECAS continues to be administered in Ghana, the cost associated with maintaining the program continues to be borne by the donor community, in part, because the government sees little incentive to fund a program in which students continue to perform poorly.

3 In 2005, DNEA was responsible for the administration of approximately 72,393 national school examinations (i.e., the Junior Secondary Certificate, JSC, International General Certificate of Secondary Education, IGCSE, and the Higher International General Certificate of Secondary Education, HIGCSE) across the Namibia in 629 examination centers. These enrollment figures were obtained from http://mafrii.com/m_dir_viewdirectorare.php?id=13&directorare=Directorate%20of%20National%20Examinations%20And%20Assessments on March 29, 2011.

7. ENSURE ADEQUATE FUNDING TO BUILD AND INSTITUTIONALIZE THE NATIONAL ASSESSMENT PROGRAM, WHICH IS A COSTLY AND SLOW PROCESS. INVEST IN HIRING AN ASSESSMENT EXPERT WITH DEMONSTRATED CAPACITY IN IMPLEMENTING NATIONAL ASSESSMENT PROGRAMS TO PROVIDE CONTINUOUS IN-COUNTRY SUPPORT.

When probed about whether they thought there was adequate funding to support the assessment activities, the responses of the interviewees depended largely on the type of assessment system they were trying to build. Generally, ERP and BES III felt that funding was adequate to develop the assessment tools needed for project monitoring and evaluation. However, NSAT, MIDEH, and BECAS felt that the funding provided to build a national assessment program was insufficient.

Of all the projects reviewed, BECAS was the most underfunded. Given the purpose of the BECAS project, i.e., to build a formative and national summative assessment system, the initial funding provided to the project to achieve their project goals was grossly inadequate (less than \$1 million over three years). With the funding awarded, BECAS was expected to develop items for formative and summative assessments; pilot the items in schools; produce copies of the test booklets and administer the test to a national sample of students; and conduct analysis on the data to determine student learning trends across the country. As a result of all the competing expenses, the project was unable to afford a manager in Ghana to oversee the project whenever the consulting assessment expert left the country. Hence, the project work would stagnate and fail to be completed in a timely manner each time the assessment expert was not on the ground to manage the work.

Eventually, the implementing partner was able to procure more funds from USAID to hire a project manager to supervise the work while the assessment expert was out of country. While it made a significant difference to have a full-time project manager in Ghana, this individual had never been involved in or managed a national assessment program before and was, thus, learning about the process during project implementation. The limited funding for BECAS affected activities in other ways as well. For example, the project was not able to adapt much or respond accordingly to the realities on the ground. All BECAS test administrators (approximately 500) had to be trained in one session, which severely hindered opportunities to answer questions or clarify misunderstandings. In fact, according to some interviewees, project funding was insufficient to accomplish the objectives set forth by USAID.

In the event that sufficient funds cannot be set aside to establish an assessment program, USAID should explore a cost-sharing structure with the local government similar to the existing agreement with NMOE. Apart from the cost of hiring assessment experts, NMOE has borne the cost for developing and administering the assessments in Namibia – including trainings for item writers; test administrators; standard setting; producing test booklets for pilot testing; creating an item bank; printing subject-specific test booklets for every 5th and 7th grade student in the country; processing student responses for analysis and reporting; printing score reports for every school; and training head teachers. If the local government cannot or does not want to contribute funds towards building a national assessment program, then USAID should consider funding the country to participate in an international assessment instead. This funding should include in-country guidance and assistance with hands-on support and capacity building for the successful implementation of large-scale test administration, this experience could then be used to prepare local staff for future local test administrations. Funds could be dedicated funds to contract the services of experienced educational research analysts to conduct an extensive and comprehensive secondary data analysis on data obtained from an international test (funded by another donor) to shed light on specific systemic weaknesses and education issues both nationally and regionally.

8. ENSURE ADEQUATE TIME TO BUILD AND INSTITUTIONALIZE THE NATIONAL ASSESSMENT PROGRAM.

Arguably, the ultimate goals of any national assessment program are to 1) provide educators information about what students are successfully mastering in the classroom and areas in which they need additional support, and 2) establish the systematic practice of using data to make program decisions that impact education nationally or regionally, and 3) target decisions to improve instruction within the classroom. Although providing data to educators (the first goal) is relatively easy to achieve, educating individuals to understand and use assessment results to change entrenched practices in education (the second goal)

Formative Assessments

Formative assessment is a planned process in which assessment-elicited evidence of students' status is used by teachers to adjust their ongoing instructional procedures or by students to adjust their current learning tactics.

Popham (2008)

Summative Assessments

Summative assessments are assessments of student learning. They are given at a point in time to measure the students' achievement in relation to a clearly defined set of standards. These assessments are given after learning is supposed to occur.

Iowa Department of Education, 2011

is a vastly more challenging activity that requires more time for strategic development and careful implementation. Without adequate time to properly apply the second goal, any national assessment effort is, at best, a superficial one that will fail to bring about the desired change and eventually be written off by local governments as a non-value-added activity, which was what happened on BECAS.

BECAS was expected to develop a national assessment program in three years, when in reality, it takes approximately three years to develop and validate the assessments and work out any problems with the implementation procedures. BECAS did not have the time it needed to train local educators about the importance of assessment data or how to use them. Furthermore, the time allocated to project implementation neither allowed for GES to develop a relationship with the implementing partner nor for proper local capacity building. So, the assessments were not well understood or appreciated by GMOE which, in turn, led to the subsequent refusal by GMOE to allocate additional education funds. Because of the inadequate time to develop and implement the unrealistic goals of the national assessment program, it is safe to conclude that BECAS was not set up from the outset to succeed.

9. DEVELOP A PLAN FOR SUSTAINING THE NATIONAL ASSESSMENT PROGRAM WITHIN THE FIRST TWO YEARS OF THE PROJECT. ENSURE THAT THE LOCAL GOVERNMENT IS HEAVILY INVOLVED IN THE DEVELOPMENT OF AND HAS FULLY BOUGHT INTO THE TRANSITION AND SUSTAINABILITY PLAN.

Typically plans for sustaining activities after the project closes are considered closer toward the final year of the project, which is often too late to ensure a successful transition to the local government. In fact, we recommend that planning for program sustainability begin alongside the development of activities at the beginning of the project. Doing so would ensure that proper consideration is given to implementing activities in a way that is manageable for the local government that will be assuming the responsibilities for implementing all the activities after the project closes.

The implementation model utilized on NSAT whereby DNEA assumes all responsibility for and leads all development, administration, training, and reporting activities with periodic support from an assessment expert is one that most probably promotes the highest level of sustainability. Under this model, NMOE is not only engaged in the planning but also actively involved in making key implementation decisions that it thinks are appropriate and

beneficial for its students, while the assessment expert plays an advisory role to help the NMOE (a) design a sound national assessment program, (b) anticipate the consequences of different decisions, and (c) build staff capacity. With the NMOE at the helm of the Namibian national assessment program, transitioning the project entirely to the government would be hassle-free and seamless. However, it is rare to have the host government as involved in the development of the national assessments as NMOE has been.

In Honduras, the MIDEH project has been developed and led entirely by the implementing partner with support from HMOE, which is typical of most USAID projects. To sustain MIDEH's efforts after it closes in July 2011, HMOE is working to create a semi-autonomous institution for assessment and evaluation to oversee and manage the entire assessment production, administration, and data analysis process. Under ordinary circumstances, planning for a transition a few months prior to closing a project would be considered too late, however, USAID is extending the MIDEH activities under a cooperative agreement. As part of this new cooperative agreement the implementing partner must help the government of Honduras identify a strategy to develop a solid and realistic financial plan that will enable them to absorb and sustain the MIDEH activities in their ongoing education budgets from the outset of the project.⁴ Building this requirement as an objective to the upcoming MIDEH project ensures that the implementing partner begins to identify a sustainability plan during the early phases.

⁴ Mejorando el Impacto al Desempeño Estudiantil de Honduras (Improving the Impact of Student Development in Honduras), Request for Application No: RFA-522-11-000003. USAID, Honduras

The development of a technically sound, comprehensive, effective, and sustainable assessment program is an important priority for national education policy initiatives. Indeed, such an assessment program is a fundamental and integral element of any effort toward systematic educational improvement. However, the development of a sound assessment system also requires long-term planning, investment, and patience before an education community will fully realize the promise of this work. An education community engaged in establishing a sound assessment system must make a long-term commitment to this initiative at every level, from national government officials to the local teachers and administrators in the nation's classrooms. Just as government officials must be committed to setting long-term policy goals and benchmarks and securing the necessary resources to support assessment initiatives, so too must local educators be committed to being serious and conscientious in their critical roles—supporting the successful execution of day-to-day assessment development and administration processes, and making use of the promise of assessment results in the ongoing evaluation and improvement of curriculum and instruction.

If any part of this complex network of commitment by multiple parties to various elements of the assessment system is not in place, along with the patience and long-term perspective necessary to sustain this commitment, then there will be a risk that an assessment initiative may not bear the fruit that it would otherwise be capable of producing. Careful attention to each of these elements, anticipating the necessary assessment expertise, educator training, assessment system design considerations, and funding sources—along with acknowledging and addressing any potential risk within the policy environment that could potentially compromise the success of assessment implementation—can secure the critical conditions for a sustainable and effective assessment system and its related improvements in student learning.

Knowing all of this, USAID must examine each potential funding opportunity by evaluating whether or not these critical conditions are readily apparent within the social fabric of the potential beneficiary. MOE officials who demonstrate a clear commitment to the policy goals and long-term perspective that will be required for success must be in place. Local education communities must be prepared to receive, understand, and apply

the training and support that will empower them to fulfill their critical roles. All parties must be committed to long-term sustained efforts along the path of educational improvement, and must be ready and willing to employ assessment as an invaluable tool that will help them get where they want to be for the sake of their nation's schoolchildren. When these conditions are not present USAID can still make a difference by funding work to ready conditions for the successful development and implementation of a sound assessment system in the future. Conversely, when these conditions are present, USAID must be willing to step forward and provide the informed, targeted, long-term support that will allow the promise of a sound assessment system within such a society to be realized – in an exciting and rewarding enrichment of the education opportunities that the nation will provide to its schoolchildren.

REFERENCES

Iowa Department of Education. (2011). "Assessment of Learning." Viewed at http://educateiowa.gov/index.php?option=com_content&task=view&id=1044&Itemid=1172

Popoham, W.J. (2008). *Transformative Assessment*. Arlington, VA: Association for Supervision and Curriculum Development.

USAID/Ghana. (2003). *Request for Application: Basic Education Comprehensive Assessment System*.

USAID/Honduras. (2011). *Request for Application: Mejorando el Impacto al Desempeño Estudiantil de Honduras (Improving the Impact of Student Development in Honduras)*.

The EQUIP2 Lessons Learned in Education Series: Guides to Education Project Design, Implementation, and Evaluation Based on Project Reviews of USAID-funded EQUIP2 Associate Awards. Other topics in this series include:

- Decentralization
- Education Management Information Systems (EMIS)
- Policy Dialogue
- Secondary Education
- Teacher Professional Development

For more information, please contact:

USAID

Patrick Collins
EGAT/ED/BE, USAID
Washington
1300 Pennsylvania Ave., NW
Washington, DC 20532
Tel: 202-712-4151

FHI 360

Audrey-marie Schuh Moore
EQUIP2 Project Director
1825 Connecticut Ave., NW
Washington, DC 20009
Tel: 202-884-8187
Email: aumoore@fhi360.org
Web: www.equip123.net