Assessing Capacity for Evaluation
A Pilot in Kenya

Introduction

Funders want to know if the programs they support in low-resource countries are succeeding and look to evaluations for the answer. However, if outside experts conduct the evaluations, these countries miss the chance to build their capacity to do this work themselves. Strengthening the capacity of local institutions to implement evaluations is a critical need.

Photo: Riccardo Gnagale/USAID Kenya

The capacity-building approach embraced by MEASURE Evaluation, funded by the United States Agency for International Development (USAID), is "learning by doing." While an evaluation is implemented, a concurrent process develops the knowledge, skills, and competencies of a local organization. One of our objectives is to help organizations in developing countries implement scientifically rigorous evaluations. To achieve this goal, the project often engages a local organization as a research partner and, as the evaluation is implemented, simultaneously conducts training and mentoring to strengthen the local organization's evaluation capacity.

In 2016, with the goal of systematizing this learning-by-doing approach, MEASURE Evaluation prepared a toolkit for evaluation capacity building. It consists of a framework and guidance for facilitators, an assessment tool, and a template for a plan for capacity building. The process begins with an assessment of a research partner's evaluation capacity, followed by a plan to fill identified gaps. The aim of this systematic approach is to introduce efficiencies and conserve a country's limited resources for capacity building.

That same year, MEASURE Evaluation piloted this guidance with a local research partner in Kenya: the African Population and Health Research Center (APHRC), which is a pan-African research institution headquartered in Nairobi. Our goals were to conduct an evaluation and to improve APHRC's evaluation capacity. This experience in Kenya was a valuable test of the toolkit, yielding insights for future applications.

Methods

We began with a review of recent evaluation capacity assessment literature, looking for evaluation-capacity building frameworks, models, and tools. We identified several assessment frameworks that presented evaluation capacity domains consistent with MEASURE Evaluation's result areas:

1. Improved health information systems
2. Improved capacity to manage health information systems
3. Improved tools and approaches for addressing health information challenges
4. Improved capacity for rigorous evaluation

Of these, we chose two to adapt: the Evaluation Capacity Index (ECI) (Nielsen, Lemire, & Skov, 2011) and the Six Dimensions of Evaluation Capacity Framework (Bourgeois & Cousins, 2013).

The ECI was developed to test a capacity-building model with public sector organizations in Denmark. The index divides evaluation capacity into "evaluation supply" and "evaluation demand." Supply and demand are further divided into subdimensions for demand ("objectives," "subjects," and "processes") and supply ("technology" and "human capital"), and each of these five subdimensions has components (Nielsen, et al., 2011).

The Six Dimensions of Evaluation Capacity Framework identifies elements of evaluation capacity in Canadian government organizations. Six dimensions are divided into the capacity to do evaluation and the capacity to use evaluation. Human resources, other organizational resources, and evaluation planning and activities fall under the first capacity and evaluation literacy, integration with organizational decision making, and benefits of learning fall under the second. Each of these has subdimensions (Bourgeois & Cousins, 2013).

Adapting Frameworks for the Pilot Test

The objectives defined for the evaluation activity in Kenya were part of a work plan funded by the United States President's Emergency Plan for AIDS Relief (PEPFAR) to (1) collect and analyze data on program outcomes for orphans and vulnerable children (OVC) as part of PEPFAR's monitoring, evaluation, and reporting requirements for program decision making; and (2) to build the capacity of a local research partner in outcomes monitoring.

Informed by the literature review and the objectives in Kenya, we prepared a hybrid capacity-building framework, drawing from the ECI and the Six Dimensions framework. The ECI framework has a technology component focused on data collection and analysis, which we decided was an important area of capacity building for APHRC and for other settings in which MEASURE Evaluation implements evaluations. We saw the Six Dimensions framework as relevant for improving team capacity and for assessing the ability to do an evaluation.

We customized this draft framework to produce a final version that would work well in Kenya and with APHRC. The resulting framework (Figure 1) has four dimensions; its 12 components are relatively standard for any evaluation that includes primary data collection. The framework also has subcomponents that further describe the 12 components of capacity (not shown).

Figure 1. Framework to build Kenya's capacity to conduct evaluation

Developing Tools for the Pilot Test

To measure progress in evaluation capacity, we established a baseline—using a participatory, self-assessment method—to assess capacity needs and to promote discussion and consensus on the competencies we would seek to improve. The results informed a plan to address capacity gaps. For the end line, we documented progress, using the benchmarks given in that plan. We made this decision chiefly because of the short timeframe of the engagement with APHRC (eight months) and uncertainty at the time of the assessment whether APHRC would be the selected partner for the second round of the OVC study.

During a workshop, APHRC rated itself on the components of our assessment framework. For each component, the team was asked to identify capacity challenges and gaps and to brainstorm strategies to address them. We developed a tool to support this capacity assessment process and a discussion guide.

Results

Participants reported that the assessment exercise and planning process were well received and that the instructions and templates in the guidance being piloted were clear. The APHRC team rated its capacity as "excellent" on eight of the components and "good" on the other two (evaluation design and fieldwork planning) and identified areas for improvement that included better understanding of OVC research, strengthened capacity in sampling and weighting, and standard safety and security practices.

For each area of improvement, the team identified one to three actions to undertake. For example, under best practices in evaluation and research with OVC populations, the team suggested it would add a training session for the team and for

data collectors and would develop guidelines for safeguarding children. The capacity-building plan outlined steps within the planned scope of work and, therefore, required few additional resources other than time. All actions were completed, and all benchmarks were achieved, as planned.

Discussion

The evaluation capacity-building guidance developed by MEASURE Evaluation was successfully applied in Kenya. The approach and the materials proved useful for planning and incorporating capacity building in the collaborative evaluation implementation. The capacity-building plan resulting from the assessment tool was simple and practical and served the intent of this pilot activity: to lay out a clear path for collaboration between APHRC and MEASURE Evaluation, structured around the identified needs of the local partner.

The toolkit was effective for several reasons. The mentoring by MEASURE Evaluation allowed the team to reflect on its experiences with and skills related to the survey. The assessment and resulting plan yielded practical action items that could be implemented within a short timeframe. The MEASURE Evaluation study leader in Kenya felt that the process using the guidance and related tools was useful and added value in team building and for management purposes. APHRC staff found the process useful, even turning it into an opportunity for strategic planning for their future work.

Challenges

This pilot application also revealed several generalizable challenges and benefits to applying the process more broadly.

Although the toolkit was effective in producing a useful assessment and capacity building plan, the process was necessarily subjective. To objectively observe the team's capacity in all evaluation capacity-building dimensions, one would need to observe and document performance during the implementation of the evaluation—from the request for applications to the submission of the final report. However, our approach was to assess and address capacity gaps on evaluation areas during the planning and implementation stages, achieving incremental improvement in the process.

Another subjectivity introduced through this method was that the team bases its projections for future performance on assessments of its own prior experience. This could lead to a missed opportunity for capacity building, if the team were to fail to flag an area where it later encountered challenges. Nevertheless, we felt that prospective assessment was the best option, because it allowed us to use the assessment as the basis of the capacity-building plan.

Even though the tools were well received in Kenya, we were left with some doubt as to whether this was because of the clarity of the tools or APHRC's familiarity with capacity building. APHRC staff immediately understood the intent of the process and the importance of each section of the assessment. They self-organized to complete the assessment and the capacity-building plan. The team members used the assessment process to reflect on its capabilities and to strategically plan and discuss the evaluation. Would a research team less familiar with capacity building have engaged with the tools and process as smoothly?

An additional challenge in this setting was the short formal relationship with the research partner. The duration of the agreement and the time needed for the survey limited the opportunity for significant capacity building, especially using mentoring and the learning-by-doing approach. Because of the short duration, the plan for capacity building was accomplished through the completion of a short plan template. However, our opinion is that the template is sufficiently broad to accommodate planning for capacity-building activities over a longer-term partnership.

Conclusion

Given the experience in Kenya, the toolkit shows promise as a resource for others interested in capacity building with a research partner through collaborative implementation. MEASURE Evaluation will continue to pilot the toolkit in more and varied settings, to further test it for general use. Then we will revise it and publish it for external audiences.

References

Bourgeois, I., & Cousins, J. B. (2013). Understanding dimensions of organizational evaluation capacity. *American Journal of Evaluation*, 34(3), 299−319. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/1098214013477235.

Nielsen, S. B., Lemire, S., & Skov, M. (2011). Measuring evaluation capacity—Results and implications of a Danish study. *American Journal of Evaluation, 32*(3), 324−44. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/1098214010396075.

Selected Evaluation Capacity Assessment Frameworks, Models, and Tools

Bourgeois, I. & Cousins, J. B. (2013). Understanding dimensions of organizational evaluation capacity. *American Journal of Evaluation*, 34(3), 299–319. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/1098214013477235

Cousins, J. B., Goh, S.C., Elliott, C. Aubry, T. & Gilbert, N. (2014). Government and voluntary sector differences in organizational capacity to do and use evaluation. *Evaluation and Program Planning*, 44, 1–13. Retrieved from https://www.sciencedirect.com/science/article/pii/S0149718913001055

Nielsen, S. B., Lemire, S. & Skov, M. (2011). Measuring evaluation capacity—Results and implications of a Danish study. *American Journal of Evaluation*, 32(3), 324–344. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/1098214010396075

Preskill, H. & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, Volume 29, Number 4, 443–459. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/1098214008324182

Preskill, H., & Torres, R. T. (2000). Readiness for organizational learning and evaluation instrument. Retrieved from https://www.fsg.org/tools-and-resources/readiness-organizational-learning-and-evaluation-instrument-role

Taut, S. (2007). Studying self-evaluation capacity building in a large international development organization. *American Journal of Evaluation*, Vol. 28, No.1, 45–49. Retrieved from http://journals.sagepub.com/doi/pdf/10.1177/1098214006296430

Taylor-Ritzler, T., Suarez-Balcazar, Y., Garcia-Iriarte, E., Henry, D. B., & Balcazar, F. E. (2013). Understanding and measuring evaluation capacity: A model and instrument validation study. *American Journal of Evaluation*, 34(2), 190–206. Retrieved from http://journals.sagepub.com/doi/abs/10.1177/1098214012471421

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. FS-18-253

USAID FROM THE AMERICAN PEOPLE

MEASURE Evaluation