USAID
FROM THE AMERICAN PEOPLE

VERSION 1 | JUNE 2013

TECHNICAL NOTE

Conducting Mixed-Method Evaluations

Monitoring and Evaluation Series

This document provides guidance on using mixed-methods for evaluations.

Technical Notes are published by the Bureau of Policy, Planning and Learning and provide key concepts and approaches to USAID staff and partners related to the Program Cycle.

INTRODUCTION

Over the past three decades, evaluators and others in the development field have increasingly recognized that incorporating multiple methods into a single evaluation often results in a stronger, more complete evaluation than conventional evaluation approaches relying on only one method. This trend has led to a rapidly growing interest in mixed-method evaluations among both practitioners and evaluators. At least two journals dedicated to mixed-method evaluations have been launched, and the number of books on the subject is growing steadily. Notably, USAID's *Evaluation Policy* strongly endorses mixed-method evaluation approaches: "Given the nature of development activities, both qualitative and quantitative methods yield valuable findings, and a combination of both often is optimal." (p. 4) This Technical Note provides guidance to USAID staff and partners on how mixed-method evaluations are conducted and important considerations when managing a mixed-method evaluation.

DEFINITION

A mixed-method evaluation systematically integrates two or more evaluation methods, potentially at every stage of the evaluation process, usually drawing on both quantitative and qualitative data. Mixed-method evaluations may use multiple designs, for example incorporating both randomized control trial experiments and case studies. They also may include different data collection techniques such as structured observations, key informant interviews, household surveys, and reviews of existing secondary data. In short, a mixed-method evaluation involves the systematic integration of different kinds of data, usually drawn from different designs. As a result, mixed-method evaluations require advanced planning and careful management at each stage of the evaluation process.

RATIONALE

The three main cases in which mixed-method designs help to strengthen an evaluation are:

(1) When different evaluation questions require different methods, or when a single evaluation question requires more than one method to answer all components.

(2) When different methods are used to answer the same elements of a single question, increasing confidence in the validity and reliability of the evaluation results.
(3) When the results from one method are used to help design future phases of the evaluation using other methods.

In addition to these three main reasons, there are other benefits that can be realized by using mixed-method designs or data collection strategies. For example, mixed-methods approaches:

- Are more likely to reveal unanticipated results.
- Can provide a deeper understanding of why change is or is not occurring as planned.
- Often capture a wider range of perspectives than might be captured by a single method.

(1) USING DIFFERENT METHODS TO ANSWER DIFFERENT QUESTIONS OR TO ANSWER DIFFERENT PARTS OF THE SAME QUESTION

In many cases, one evaluation method will be insufficient to answer all of the questions included in an evaluation statement of work (SOW). For example, suppose an SOW involves an evaluation of a project that includes a new teaching technique, and includes two questions: #1. "Was there a statistically significant difference between female and male students' academic achievement test scores?" and #2. "How did students' parents perceive the effects of the project?" A single method will likely be insufficient to adequately answer both of these questions.

To answer the first question, the evaluator might choose a quasi-experimental design that uses existing test scores from before project implementation and new test scores from after completion to compare the performance of male and female students. This approach would address the question of whether the program resulted in differences in test scores between females and males. But these methods would not help to answer question #2. To understand parent perceptions, the evaluator likely would use individual or focus group interviews of a sample of parents, and perhaps conduct an evaluative case study in order to more deeply understand how parents view the program.

Sometimes, one evaluation question may contain multiple parts, and it may be necessary to use different methods to address each part. For example, question #1 could be changed slightly to ask, "Was there a statistically significant difference between female and male students' scores? And, what explains possible gender differences in test scores?" In this case, simply relying on test scores in a quasi-experimental design would be insufficient. To understand the mechanism behind differences in female and male scores would require a different method, such as key informant interviews with teachers or focus group discussions with students.

A NOTE ON TERMINOLOGY

In the literature on evaluation, "method" is sometimes used to refer to a data collection technique (interviews, surveys, observations), and other times to an evaluation design or approach (experimental, quasi-experimental, non-experimental). Though the definition is not completely settled in the literature, this Technical Note treats evaluations that combine methods in either sense as mixed-method evaluations.

(2) USING DIFFERENT METHODS TO ANSWER THE SAME QUESTION: TRIANGULATION

Even if an evaluation question can be answered using only one method, often it is preferable to combine multiple methods to answer the same question in order to gain a more complete understanding of the issue and more confidence in the findings. By approaching the same question from more than one perspective or by using more than one technique, evaluators can then compare and contrast the results from these different methods. This process is known as triangulation. If the findings from the different methods are similar, or reinforce one another, then users can have greater confidence in the findings than if they are based on only one method.

If the findings from different methods vary significantly, the user and evaluator must carefully consider what might have happened to produce these divergent findings. One possible explanation could be bias in one set of data. Triangulation can help to minimize bias in cases like these, with data from one method acting as a check or balance against data from another method. For example, evaluators may use secondary data from the Ministry of Economy to measure changes in exports related to a trade facilitation project. But they may also suspect that firms are underreporting their exports to the government in order to pay less in taxes. To help mitigate the risk of bias caused by this underreporting in the government data, the evaluation team may distribute a survey to supported firms and also conduct in-depth interviews with key informants from a sub-sample of firms in order to obtain a more accurate picture of how the project has influenced exports. Evaluators looking to answer one question with multiple methods often combine them using the parallel process described below. It is important for the evaluation manager to understand how this process of triangulation will work, because it has implications for the resources needed to carry out such an evaluation.

(3) USING ONE METHOD TO INFORM THE DESIGN OF ANOTHER METHOD

In some cases, one method can be used to help guide the use of another method, or to explain the findings from another method. In the first case, imagine an SOW for the evaluation of a youth vocational training project including the evaluation question: "Why do youth choose to participate in project activities?" The evaluator may wish to conduct a survey of participants, but be unsure how to word the questions, or what answer choices to include. By first conducting individual and focus group interviews with participants and non-participants, the evaluator may be able to identify some common reasons for participation among the target population, and then use these data to construct the survey. In this way, the qualitative methods (individual and focus group interviews), conducted first, can inform the quantitative method (survey), that comes afterward. Because this use of mixed-method evaluation requires each method to be sequenced, one after the other, these methods are often incorporated into mixed-method evaluations using sequential processes. Again, this design choice has time and resource implications, as discussed below.

QUICK ADVICE FOR THE USAID EVALUATION MANAGER

Mixed method evaluation involves the systematic integration of different kinds of data, usually drawn from different evaluation designs. As a result, mixed-method evaluations require advanced planning and affect the evaluation budget.

Mixed method evaluations yield valuable findings, and a combination of both is often optimal given the nature of development activities. Mixed methods are more likely to reveal unanticipated results (a key advantage of evaluation over performance monitoring), which can provide a deeper understanding of why change is or isn't taking place as planned. It often captures a wider range of perspectives than might be captured by a single method.

At the planning stage, the evaluation manager must decide which methods to use and how to combine them. These decisions will be based primarily on the purpose of the evaluation and the key evaluation questions, but evaluation managers must also take into account factors such as time and cost. An evaluation manager should seek advice from a colleague with evaluation design expertise, as needed.

Lastly, using mixed methods requires mixed skills on the evaluation team. It is important to consider during the planning stage what kind of skills will be needed in order to conduct each aspect of the evaluation successfully, and then to select team members accordingly.

HOW TO MIX METHODS

As mentioned above, evaluators must consider carefully how they will integrate the different methods used into a coherent, thoughtful evaluation design. This section outlines three of the most common ways in which

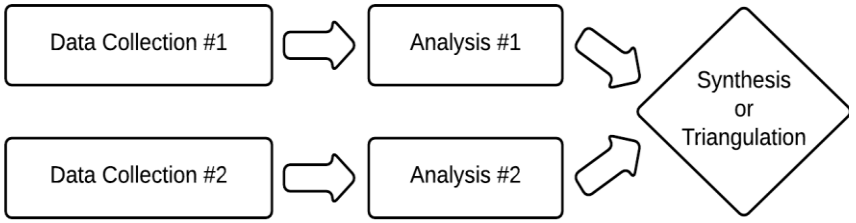
methods can be combined to accomplish the purposes described above: parallel combinations, sequential combinations, and multi-level combinations. (Table 2, at the end of this document, lists two additional techniques: conversion and data synthesis; for purposes of this note, however, the three techniques discussed in the text are sufficient to provide an overview of the issues.) A single evaluation might use more than one, or even all, of these combination patterns at different points during the evaluation process.

Each of these variations involves important management decisions on issues such as which data sources to rely on and how to gain access to them, what sites should be included in the evaluation (all project sites or a sample, and if just a sample, how the sample should be drawn), how much depth of explanation is needed, which stakeholders' views are most relevant, and so on. These kinds of questions normally need to be addressed in the SOW so that the evaluators can develop designs that answer the questions of interest. And while the evaluators may help to clarify questions and provide guidance on practical implications of design choices, it is the responsibility of the USAID managers to have made the key decisions, such as what questions need to be answered and what kinds of evidence are needed, when developing the SOW. These choices have important implications for time and resources. Mixed-method evaluations must be carefully planned with a thorough understanding of why and how each method is to be used to answer the questions. One must carefully consider the integrity of the design of the evaluation. Adding more methods to a design with the idea that "more is better" can lead to unnecessary complications and cost.

PARALLEL COMBINATIONS

In parallel combinations, methods are used separately and the findings are integrated after the data are analyzed. The same evaluation team might be involved in implementing multiple evaluation methods, and the actual data collection and analysis can happen over the same period of time, or at different times. **The key point is that in parallel combinations, each method is conducted in its entirety, separately from the other methods,** as shown in Figure 1.

FIGURE 1: PARALLEL COMBINATIONS



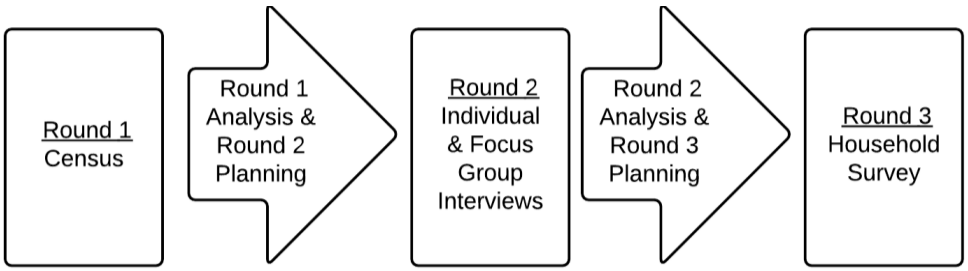
For example, a team evaluating an alternative development project could collect data from government statistics on the number of acres of land converted from illegal to legal crops, and then analyze this quantitative data to estimate the impact of the project. At the same time, the same team also might conduct individual and focus group interviews with farmers to better understand their choice of what to grow. These two methods could take place simultaneously, and the data analyzed separately. Then, if the data from the two methods were intended to answer the same question, the findings could be triangulated. If they were intended to answer different questions, then the results would be combined, or synthesized, in the evaluation report.

SEQUENTIAL COMBINATION

Multiple methods also can be used at different times and in a specific order. **With sequential combinations, methods are employed one after the other, with the findings from methods used earlier in the evaluation informing the design and implementation of methods used later in the evaluation.**

Round 1 data collection: Based on the evaluation design, the evaluator constructs data collection instruments, conducts a first round of data collection, and analyzes that data. For example, as illustrated in Figure 2, if the evaluation is intended to determine whether a water, sanitation and hygiene (WASH) project is leading to higher rates of hand-washing in a particular community, the evaluator may first collect and analyze quantitative data from the most recent census to gather information about the population of the community and any relevant demographic characteristics, such as age, socio-economic status, and ethnicity.

FIGURE 2: SEQUENTIAL COMBINATIONS



Use of Round 1 findings to inform Round 2 data collection: The evaluator then can use the findings from the first round of data collection to inform the second round. In this example, the census data would help the evaluators identify the demographic characteristics that should be represented among the second-round interviewees. These interviews would help bring to light the common reasons why the affected population chooses to wash their hands, or not, and what obstacles may exist that prevent them from doing so.

Use of Round 1 and 2 findings to inform Round 3 data collection: Based on qualitative information collected during the interviews, the evaluator would design the content of a household survey. This survey would help to answer the central question of what proportion of villagers has changed their hand-washing behavior, as well as why their behavior changed. The household survey would be conducted in the community, possibly following a sampling strategy based on Round 1 quantitative data. The data collected during the Round 3 household survey would directly address the original evaluation questions. While it might be possible to do only a survey, the mixed methods approach would have a number of advantages: thanks to the Round 1 analysis the sample would be more representative of the total population of the village, and the survey questions would be more appropriately tailored to the local context and to the diversity of the community thanks to the Round 2 analysis.

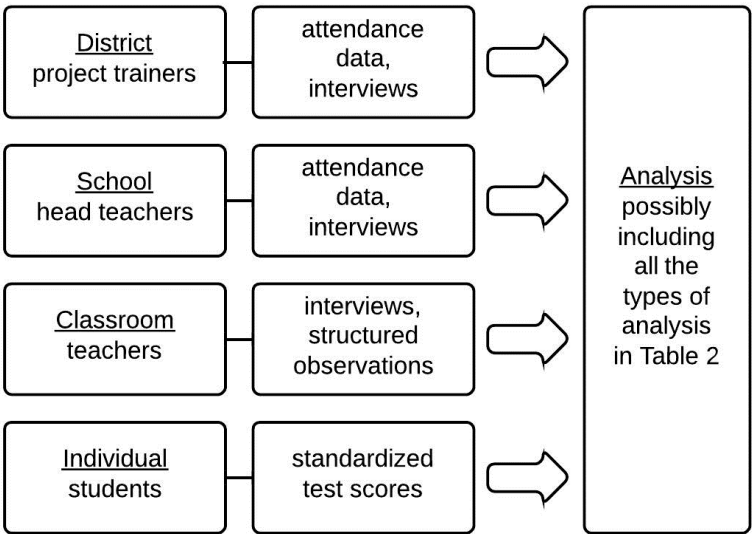
MULTILEVEL COMBINATIONS

Many projects involve systems with multiple levels, and the evaluators of these projects often must collect data and draw conclusions about each of these levels in order to have a clear understanding of the overall performance of the project. Not surprisingly, the evaluator may conclude that different kinds of methods are best suited to collecting and analyzing information from different levels.

For an example of this type of evaluation, consider an education project which is intended to raise student literacy by introducing more effective teaching strategies in a set of project schools. The project is designed so that the head teacher from each project school is trained by project trainers at the district level, and these teachers then return to their own schools to train their fellow teachers.

To understand whether the project has been effective, the evaluator will need to collect data on student literacy, likely based on the scores on a standardized test. But to understand *why* the project has been effective or not will also require data from additional levels. As shown in Figure 3, the evaluator may also want to collect attendance data from the trainings that occurred at the district level and conduct in-depth and focus group interviews with the project trainers responsible for that training. At the school level, the same type of data could be collected from both the teachers who received the district training, and from those other teachers who received the trainings in their own schools. The evaluator may also conduct structured observations of classrooms in project schools in order to see firsthand whether teachers are using the literacy techniques promoted by the trainings. Incorporating different types of data at these different levels provides the evaluator with a more complete, holistic understanding of how the project operates and how it achieved, or did not achieve, its goals.

FIGURE 3: MULTILEVEL COMBINATIONS



Multi-level mixed-method evaluations can be combined with either parallel or sequential processes, or a combination of the two, depending on the levels of the project and the purpose of the evaluation. In the example above, there is a parallel process embedded in the multi-level design. However, integrating different types of data at different levels like this does make the overall evaluation design more complex, and requires additional planning, coordination, and management of the evaluation to make sure that all the data collected are analyzed and incorporated into the final report. While much of this responsibility necessarily falls to the evaluators, the implication for managers is that they must provide careful oversight of this complex evaluation process to ensure that the efforts are timely and well-coordinated, and carried out within time and budget constraints. Thus, from the manager's point of view there is a trade-off between the quality of these mixed-method evaluations in terms of accuracy and completeness, and the time and resource costs necessary to carry them out.

INTEGRATING MULTIPLE METHODS INTO THE EVALUATION

In order to get the most out of a mixed-method evaluation, the evaluation manager and the evaluator must consider carefully what purpose each method is intended to fulfill, and how they will be combined, at each stage of the evaluation, to most efficiently and effectively accomplish these purposes.

PLANNING THE EVALUATION

In order to conduct a successful mixed-method evaluation, the evaluator must start at the planning stage. At this point, the evaluator or evaluation manager must decide which methods to use and how to combine them. These decisions will be based primarily on the purpose of the evaluation and the key evaluation questions, but evaluation managers must also take into account factors such as time and cost.

When drafting the evaluation SOW (Scope of Work), the evaluation manager must first decide which questions need to be answered and how rigorous the evaluation needs to be. If the evaluation is intended to estimate impact that can be attributed to a specific intervention, then the evaluation will include some kind of experimental or quasi-experimental design which typically includes a control or comparison group. This

decision must be made at the planning stage of the project so the project design and management will allow for this kind of evaluation design. If the evaluation is intended to answers questions more oriented toward project management, then non-experimental designs likely will be sufficient. In cases in which an evaluation needs to answer both attribution-related and other types of questions, the evaluation manager may choose an evaluation that incorporates aspects of both experimental (or quasi-experimental) and non-experimental designs. If at the planning stage the managers decide that the evaluation requires mixed-methods, they need to be aware that it will require careful management throughout the evaluation to ensure they accomplish their dual purposes.

To see how combining these kinds of methods, or evaluation designs, can strengthen an evaluation, imagine a project that aims to increase farmers' income by training them in improved farming techniques. The project design stipulates treatment and control groups in order to allow an evaluation to measure the contribution of the project to increased farmer income. In this case, the evaluation manager might decide on a quasi-experimental design with a treatment group, who receive the training, and a control group, who do not. At the same time, the evaluation also may aim to determine whether the project was implemented according to its original design, and this part of the evaluation may rely on a non-experimental design, focusing on a description of project operations, to fulfill this aim. The quasi-experimental component of the evaluation may reveal no significant difference in income between farmers in the treatment group and those in the control group, which would indicate that the project had no impact. Some stakeholders may also conclude from this finding that training projects of this kind do not increase farmer's incomes. At the same time, the results of the non-experimental component of the evaluation may reveal that only a small number of farmers in the treatment group actually received the intended training, or that they all did receive the training and then shared the improved farming techniques with the farmers in the control group. In either case, this additional information would help to explain the findings of the quasi-experimental part of the evaluation: it is possible that training of this type can lead to increases in farmers' income, but because of faulty implementation, it is impossible to know for sure from this evaluation.

CHOOSING THE RIGHT EVALUATION TEAM

Mixed-method evaluations require specialists with different kinds of skills. Even if an evaluator can construct a statistical regression model, this does not mean that the same person will be able to conduct a structured observation or effectively facilitate a focus group interview. It is important to consider during the planning stage what kind of skills will be needed in order to conduct each aspect of the evaluation successfully, and then to select team members accordingly.

Once the evaluation team has finalized the overall evaluation design, the next step is to plan for data collection. In some cases, the evaluator may choose to begin using multiple data collection methods even during the planning stage, to help design tools or methodologies that then will be used during the data collection or data analysis stages of the overall evaluation. This use of mixed methods follows the sequential pattern of combining methods described above. A review of relevant literature and secondary data is one method commonly employed at the planning stage. Stakeholder consultations, also a common part of the planning stage, can take the form of individual and focus group interviews. The evaluator also could decide to conduct a mini-survey of some of the stakeholders to quickly get data from a larger group of people than is possible with in-depth interviews. The range of methods selected during the planning stage will depend in part on what data is already available.

For example, if the data the evaluator needs are available for the whole population in secondary data sources, such as government statistics, then there may be no reason to do a mini-survey. Alternatively, it may be beneficial to conduct focus group interviews with certain groups of stakeholders, but there may not be enough time before the start of the evaluation to do so. The evaluation manager must think carefully about the information needed to plan the rest of the evaluation, how these different methods would contribute to this

planning, and how best to spend the resources available for the evaluation. Using a mix of methods at the planning stage can greatly strengthen evaluations, and if appropriate, the evaluation manager should include guidance for their use at this stage in the evaluation SOW.

Whether a mix of methods is applied during the planning stage or not, planning for mixed-method evaluation still requires the evaluation manager to approve the methods to be used, including data collection tools, data sources, sampling strategies, data management, analysis techniques, reporting techniques, and how these different pieces of the different methods will be woven together into a coherent mixed-method design that meets all the requirements of the evaluation. One tool that can be helpful during the planning stage is an evaluation design matrix like the one in Table 1. Such a matrix, completed by the evaluation team as part of the planning process, can assure the evaluation manager that the team has considered what it needs to do to answer the evaluation questions, and has a plan for carrying out those tasks.

SELECTING THE APPROPRIATE MIX OF METHODS: A HYPOTHETICAL EXAMPLE

This section presents an example of how multiple methods can be applied in the context of a mid-term evaluation of a multi-year gender equity project. The project goals are to increase the number of women who graduate from secondary school. The project attempts to achieve these goals by combining scholarships and mentoring provided by the host government's Ministry of Education (MoE) with support and training from the project staff.

The evaluation questions ask:

(1) Has the project been successful in meeting its targets for number of scholarships distributed and number of mentors connected with scholarship recipients?
(2) In the targeted secondary schools, did the project increase the number of female graduates?
(3) What changes could be made to the project to increase its effectiveness?

TABLE 1: SAMPLE MIXED-METHOD EVALUATION DESIGN MATRIX

Q#	Data collection method	Data collection instrument	Sample questions on the instrument	Data source
1	Desk review	Annotated bibliography	NA	Project managers
1	Data quality review	Checklists, project reports	NA	Project staff
2	Review of School statistics	School data form	Enrollment records by year, grade, and gender	Project staff, government officials, school administrators
2	Semi-structured interview	Interview protocol	What are the reasons you stayed in or left school?	Students, teachers, mentors
2	Focus group interview	Focus group protocol	How do you decide whom to send to school?	Parents
3	Semi-structured interview	Interview protocol	What worked? What did not work?	Project staff, government staff, school administrators, students, teachers, mentors
3	Focus group interview	Focus group protocol	What worked? What did not work?	Students, teachers, parents, mentors
3	Survey	Survey instrument	Did the project receive enough money?	Project staff, government staff, school administrators, Students, teachers, mentors

The first question deals with contractual targets, while the second aims to test the primary theory of the project-namely that scholarships and mentoring for girls leads to increases in female secondary student graduation rates. The third question asks evaluators to determine where there are weaknesses in the project and suggest ways to strengthen them. Answering these three questions will require multiple research methods. By first examining each question individually, the evaluator will then be able to choose methods to collect all the data necessary to answer all of the questions.

To answer the first question, a document review of quarterly and annual reports could establish the up-to-date numbers of scholarships granted and of mentors trained. These could be cross-checked through examinations of project records at multiple sites, including at government offices, at the project head office and field offices, and at target schools. The quality of the management information system and monitoring processes is a central concern when answering this type of question.

To answer the second question, the evaluation manager must decide how rigorous the research should be, given time, resources, and conditions in the field. If available, multi-year enrollment records for a sample of the targeted schools from before and during the project could be collected using student records. These data could then be compared to relevant regional and national trends, or to a more systematically identified comparison group in an impact evaluation design to more confidently address attribution of outcome changes to the project. Evaluators also could conduct interviews with students, teachers, and parents to triangulate the findings from the project reports, and also to see if any other factors influenced decisions about schooling for girls. A deeper level of analysis would examine the retention and graduation data for the entire school population—scholarship and non-scholarship students—and compare them. Non-scholarship recipients and their families could also be interviewed.

To answer the third question, evaluators must determine among all the stakeholders involved with the project who should be consulted. Evaluators likely would conduct semi-structured interviews with key project staff, teachers, mentors, students, and their parents. A survey tool could be developed as well and administered to select groups of beneficiaries to rate multiple aspects of the project.

TAKING INTO ACCOUNT PRACTICAL CONSTRAINTS

In many cases, the ideal mix of methods will not be possible, either due to constraints on the evaluation itself, such as limited time or funding, or due to contextual factors, such as a challenging geographic or political environment. All evaluations must deal with these types of constraints, but they are especially relevant in deciding among different methods and ways to integrate them into a mixed-method design. In some cases a mixed-method approach can help overcome some of these obstacles. For example, if the evaluation ideally would include a large-scale household survey, but this method is too costly or time-consuming, an analysis of existing census data could be used to select a small purposive sample of informants to interview that included members of all the relevant social groups. Or if a certain portion of the population of interest is inaccessible due to security concerns, statistical matching techniques based on secondary data can be used to identify a similar, alternative group to include in the evaluation sample.

After examining the questions, it is apparent that either a parallel or multi-level approach can work. The questions could be investigated separately but simultaneously, and then the findings compared in the analysis stage. All three questions, however, share some common populations and sites, and so it is likely more efficient in this case to use a multi-level design. The mix of different tools that would be used in this evaluation model—project statistics, school data forms, semi-structured interviews for four different populations, focus group protocols, and a survey—would be influenced by the location and population of the different research sites, in this case sites at the federal and state government, schools, and households.

DATA COLLECTION

Having designed the evaluation and identified data needs and sources, the next step is to carry out the data collection strategy. Mixed-method designs usually require multiple data collection methods to accommodate the different kinds of data needed to carry them out. This can be complex, time-consuming, and costly. However, with careful planning, it may be possible to design data collection instruments and sampling methodologies so that different kinds of data can be collected with a relatively small additional investment of time and other resources.

For example, if an evaluation of an education project requires the evaluator to administer a short, standardized test to a sample of students in project schools, it might be possible to add a small number of closed-ended or open-ended survey questions to the end of the test, asking, for example, what the student had for breakfast that morning, so the evaluation also could examine how student nutrition affected educational outcomes. The evaluator could also coordinate with school officials to schedule a number of classroom observations, a focus group interview with teachers from the school, and an in-depth interview with the principal, all on the same day. In this way, several different types of data can be collected in only one visit, minimizing the disruption to the students and teachers, as well as travel time and transportation costs. Another example is a household survey which, in addition to asking questions of family members, also instructs enumerators to observe and note characteristics such as the building materials the house is made of, which often correlates with family wealth. With some forethought and creativity, data collection instruments can be systematically integrated to help capture a broad range of data effectively and efficiently.

© Photo by Amer Sawalha

DATA ANALYSIS

It is important to design the data analysis strategy before the actual data collection begins because the way that the data will be analyzed and used can affect both the choice and the order of data collection methods. Analyzing data collected from a mixture of methods is often more complicated than analyzing the data derived from one method, as the evaluator must integrate multiple data analysis methods in order to determine and understand key findings. There are several different general techniques to analyze data from mixed-method approaches, including parallel analysis, sequential analysis, and multilevel analysis, corresponding to the patterns of combining methods described above, as well as conversion analysis and data synthesis. Table 2 briefly describes these different analysis techniques and the situations in which each method is best applied. More specific data analysis techniques, such as key word coding or theme analysis for qualitative data, and cross-tabulations or regression modeling for quantitative data, can also be used within the framework of any of these more general data analysis techniques. As with all evaluations, the choice of analytical techniques depends on the purpose of the evaluation and the type of data involved, as well as time and resources available.

When Findings Do Not Converge

In cases where mixed-method evaluations employ triangulation, it is not unusual that findings from the separate analysis of each data set do not converge to support the same conclusions. If this occurs, the evaluator must try to resolve the conflict among divergent findings. This is not a disaster. Often this kind of situation can present an opportunity to generate more nuanced explanations and important additional findings that are of great value.

One method to use when findings from different methods diverge is to carefully re-examine the raw qualitative data through a second and more in-depth content analysis. This step is taken to determine if there were any factors or issues that were missed when these data were first being organized for analysis. The results of this third layer of analysis can produce a deeper understanding of the data, and can then be used to generate new interpretations. In some cases, other factors external to the project might be discovered through contextual analysis of economic, social, or political conditions or an analysis of operations and interventions across project sites.

Another approach is to reanalyze all of the disaggregated data in each data set separately, by characteristics of the respondents as appropriate to the study, such as age, gender, educational background, socio-economic status, or locale of respondents. The results of this analysis may yield other information that can help to resolve the divergence of findings. This further analysis will provide additional explanations for the variances in findings. While most evaluators build this type of disaggregation into the analysis of the data during the design phase of the evaluation, it is worth reexamining patterns from disaggregated data.

Sometimes data quality issues, such as the validity of secondary data sources or possible errors in survey data from incomplete recording or incorrect coding of responses can cause dissonance in results. If the evaluators are still at the project site, it is possible to resolve data quality issues with limited follow-up data collection by, for example, conducting in-depth interviews with key informants.

The data analysis approaches noted above underscore the need for USAID evaluation managers to plan adequate time and resources for data analysis in evaluation SOWs, as it is a critical component of any evaluation, and perhaps even more so in one using a mixed methods approach.

Dealing with Divergent Findings: An Example of Triangulation

Imagine an evaluation to assess the impact of a school feeding project on school attendance rates. School records showed that daily enrolment had increased by ten to fifteen percent after the school feeding project began. A review of records from other local schools without the feeding project did not find any similar increase. A household survey before and after the project detected only a much smaller increase. Interviews with key informants provided different opinions. NGOs involved with the project, or who ran other school feeding projects, reported the project had been successful, whereas several other informants said that the school might deliberately over-report the increase to convince sponsors to continue the feeding project. How would triangulation be used to obtain the best estimate on the basis of this conflicting information?

First, the evaluators responsible for managing each kind of data collection would meet to explore possible explanations for the differences and to understand any possible sources of bias in the data. How reliable are attendance records? Does someone check names each day while the students are there, or does the teacher try to recall at some later point in the day after the students have gone home? Is there any reason why schools would intentionally inflate the number of children attending? When is attendance recorded? If it is at the start of the day, might some children just come for the breakfast and then leave – but be reported as attending? Are there any reasons that parents might misreport the number of their children attending school? Might key informants have any reason to over- or under-estimate the effects of the school feeding project?

Second, if this discussion does not fully explain the differences, as it probably would not in this case, the evaluators might agree on the additional kinds of data that would help clarify the situation and which would be feasible to collect. The most obvious approach would be to examine in more detail the school attendance records to address the questions mentioned above.

Third, a return to the field would be organized (where feasible). When revisiting the schools, the evaluators also would compare the attendance records with the number of children in the classroom after the breakfast had ended. They also might identify some of the children who were not in school but whose parents had reported in the survey that they attended regularly. Some of the families might then be revisited to check on the consistency between parents' reports of attendance and actual school attendance. They would then conduct informal interviews to try to understand reasons for the discrepancies.

These follow-up returns to the field are extremely valuable, but they can only take place if time and resources have been budgeted in advance. Consequently, it is recommended that, the evaluation budget should include time and money for this purpose as these kinds of inconsistencies are very commonly detected during data analysis.

REPORTING

Through the use of mixed-method evaluations, findings and conclusions can be enriched and strengthened. Yet there is a tendency to underuse, or even not to use, all the data collected for the evaluation. Evaluators can rely too heavily on one particular data source if it generates easily digestible and understandable information for a project manager. For example, in too many cases quantitative data that can be presented easily in graphs or tables are emphasized, whereas possibly more important but harder to present data generated from qualitative methods are insufficiently analyzed and reported. Evaluation reports should strive for interesting graphical presentations of findings from qualitative data as well.

One way to prevent underutilization of findings is to write a statement of work that provides the evaluator sufficient time to analyze the data sets from each method employed, and hence to develop valid findings, explanations, and strong conclusions that a project manager can use with confidence. Additionally, statements of work for evaluation should require evidence of, and reporting on, the analysis of data sets from ***each*** method that was used to collect data, or methodological justification for not having included analyses from any data sources used.

TABLE 2: TECHNIQUES FOR ANALYZING MIXED-METHOD DATA[1]

Type	Analysis Technique	Analytical Output
Parallel	Two or more data sets collected using a mix of methods (quantitative and qualitative) are analyzed separately. The findings are then combined or integrated.	Triangulating designs to look for convergence of findings when the strength of the findings and conclusions is critical, or to use analysis of qualitative data to yield deeper explanations of findings from quantitative data analysis.
Conversion	Two types of data are generated from one data source beginning with the form (quantitative or qualitative) of the original data source that was collected. Then the data are converted into either numerical or narrative data. A common example is the transformation of qualitative narrative data into numerical data for statistical analysis (e.g., on the simplest level, frequency counts of certain responses).	Extending the findings of one data set, say, quantitative, to generate additional findings and/or to compare and potentially strengthen the findings generated from a complimentary set of, say, qualitative data.
Sequential	A chronological analysis of two or more data sets (quantitative and qualitative) where the results of the analysis from the first data set are used to inform the analysis of the second data set. The type of analysis conducted on the second data set is dependent on the outcome of the first data set.	Testing hypotheses generated from the analysis of the first data set.
Multilevel	Qualitative and quantitative techniques are used at different levels of aggregation within a study from at least two data sources to answer interrelated evaluation questions. One type of analysis (qualitative) is used at one level (e.g., patient) and another type of analysis (quantitative) is used in at least one other level (e.g., nurse).	Evaluating cases where organizational units for study are nested (e.g., patient, nurse, doctor, hospital, hospital administrator in an evaluation to understand the quality of patient treatment).
Data Synthesis	A multi-step analytical process in which: 1) a rating of project effectiveness using the analysis of each data set is conducted (e.g., large positive effect, small positive effect, no discernible effect, small negative effect, large negative effect; 2) quality of evidence assessments are conducted for each data set using "criteria of worth" to rate the quality and validity of each data set gathered; 3) using the ratings collected under the first two steps, develop an aggregated equation for each outcome under consideration to assess the overall strength and validity of each finding; and 4) average outcome-wise effectiveness estimates to produce one overall project-wise effectiveness index.	Providing a bottom-line measure in cases where the evaluation purpose is to provide a summative project-wise conclusion when findings from mixed-method evaluations using a triangulation strategy do not converge and appear to be irresolvable, yet a defensible conclusion is needed to make a firm project decision. Note: there may still be some divergence in the evaluation findings from mixed data sets that the evaluator can still attempt to resolve and/or explore to further enrich the analysis and findings.

[1] See Teddlie and Tashakkori (2009) and Mark, Feller and Button (1997) for examples and further explanations of parallel data analysis. See Teddlie and Tashakkori (2009) on conversion, sequential, multilevel, and fully integrated mixed methods data analysis; and McConney, Rudd and Ayers (2002) for a further discussion of data synthesis analysis.

ADDITIONAL RESOURCES

The following resources can be used as samples or templates, or to provide more information on the topics reports and on evaluation in general. Some other resources exist but are out-of-date with current USAID guidance. Where information differs, the USAID Evaluation Policy and the USAID ADS (Automated Directives System) 200 series take precedence over other resources.

Bamberger, Michael. *Integrating Quantitative and Qualitative Methods in Development Projects,* World Bank Publications, 2000.

Greene, Jennifer C. 2007. *Mixed Methods in Social Inquiry*. Sage Publications.

Journal of Mixed Methods Research. Sage Publications.
http://mmr.sagepub.com.

Mark, Melvin M, Irwin Feller and Scott C. Button. "Integrating Qualitative Methods in a Predominantly Quantitative Evaluation: A Case Study and Some Reflections." *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms.* Greene and Caracelli eds. *New Directions for Evaluation*. Jossey-Boss Publishers, No. 74 Summer1997. pp. 47-59.

McConney, Andrew, Andy Rudd and Ayres Robert,. "Getting to the Bottom Line: a Method for Synthesizing Findings Within Mixed-Method Program Evaluations." *American Journal of Evaluation*. Vol. 3. No. 2. 2002. pp. 121-140.

USAID. Automated Directives System (ADS), Chapter 203, Assessing and Learning.