



## PERFORMANCE MONITORING & EVALUATION

# TIPS

## CONDUCTING MIXED METHOD EVALUATIONS

### ABOUT THIS STYLE GUIDE

These TIPS provide practical advice and suggestions to USAID managers on issues related to performance monitoring and evaluation. This publication is a supplemental reference to the Automated Directive System (ADS) Chapter 203.

#### INTRODUCTION

This TIPS provides guidance on using mixed methods for evaluation research. Frequently program evaluation statements of work specify that mixed methods should be used to answer key evaluation questions. The rationale for using a mixed method evaluation design, guidance for selection of methods, and examples of techniques for analyzing data derived from mixed method data collection practices are provided. This TIPS also shows how methods are selected in an example that calls for evaluating the effectiveness of a training program, and how these data can be analyzed using a common technique referred to as parallel analysis.

#### MIXED METHOD EVALUATIONS DEFINED

A mixed method evaluation is used to describe any evaluation study that uses two or more data collection techniques or methods to collect information. This includes a combination of different data collection methods ranging from structured observations and key informant interviews to pre and post-test surveys and reviews of government statistics. Implicit in this definition is the collection and use of both quantitative and qualitative data to analyze, identify findings and develop conclusions in response to the evaluation questions.

#### RATIONALE FOR USING A MIXED METHOD EVALUATION DESIGN

There are several different reasons why mixed methods are used in an evaluation design:

- When a mix of different methods is used to collect data from different sources to provide independent estimates of key indicators, it increases the validity of conclusions related to an evaluation question. This is referred to as triangulation (see TIPS 5 Rapid Appraisal, and Bamberger, Rugh and Mabry 2006, for further explanation and descriptions of triangulation strategies used in evaluations).

## Key Steps in Developing a Mixed-Method Evaluation Design and Analysis Strategy

1. In order to determine the methods that will be employed, carefully review the purpose of the evaluation and the primary evaluation question/questions. Then select the methods that will be the most useful and cost-effective to answer that question/questions in the time period allotted for the evaluation. Sometimes it is apparent that there is one primary method that can be used to answer most, but not all aspects of the primary evaluation question.
2. Select complementary methods to cover different aspects of the evaluation question (for example, the how and why issues) that the first method selected cannot alone answer, and/or to enrich and strengthen data analysis and interpretation of findings.
3. In situations when the strength of findings and conclusions to a key question is absolutely essential, employ a triangulation strategy. What additional data sources and methods can be used to obtain information to answer the same question in order to increase the validity of findings from the first method selected?
4. Re-examine the purpose of the evaluation and the methods initially selected to ensure that all aspects of the primary evaluation question will be covered thoroughly. This is the basis of the evaluation design. Design data collection instruments accordingly.
5. Design a data analysis strategy to analyze the data that will be generated from the selection of methods chosen for the evaluation.
6. Ensure that the evaluation team composition includes members that are well-versed and experienced in applying each type of data collection method and subsequent analysis.
7. Ensure that there is sufficient time in the evaluation statement of work for evaluators to fully analyze data generated from each method employed and to realize the benefits of conducting a mixed method evaluation.

- When reliance on one method alone may not be sufficient to answer all aspects of evaluation questions.
- A mixed method design can be used in a complimentary way by using the analysis of data collected from one method to help interpret findings from the analysis of data collected from another method. For example, in-depth interviews or focus groups can help interpret statistical patterns from data collected from a quantitative method such as a survey using a quasi-experimental design. This yields a richer analysis and

can also provide a better understanding of the context in which a program operates.

There are a number of additional benefits derived from using a mix of methods in any given evaluation.

- Using mixed methods can more readily yield examples of unanticipated changes or responses.
- Mixed method evaluations have the potential of surfacing other key issues and providing a deeper understanding of program context that should be considered in the analysis of

data for the development of findings and conclusions.

- Mixed method evaluations often yield a wider range of points of view that might otherwise be missed.

### DETERMINING WHAT METHODS TO USE

In a mixed method evaluation, the evaluator may use a combination of a survey using comparison groups in a quasi-experimental or experimental design, key documents, government statistics, in-depth

interviews with key informants, focus groups, and structured observations. The selection of methods, or mix, depends on the nature of the evaluation objective and the key questions to be addressed.

## SELECTION OF DATA COLLECTION METHODS – AN EXAMPLE

***The selection of which methods to use in an evaluation is driven by the key evaluation questions to be addressed. Frequently, one primary evaluation method is apparent.***

For example, suppose an organization wants to know about the effectiveness of a training program conducted for 50 individuals to set up their own small business after the completion of the training.

The evaluator should then ask what methods are most useful and cost-effective to assess the question of the effectiveness of that training program within the given time frame allotted for the evaluation. The answer to this question must be based on the **stated outcome** expected from the training program. In this example, let us say that the organization's expectations were that within one year, 70% of the individuals will have started a small business utilizing the skills and knowledge that they obtained through the training program.

What is the best method to determine whether this outcome has been achieved? The most direct means of answering this question is to survey 100% of the individuals who graduated from the training program using a close-ended questionnaire. It follows that a survey instrument should be designed to determine if these individuals have actually succeeded in starting up a new business.

While this sounds relatively straightforward, organizations are often interested in related issues. If less than 70% of the individuals started a new business one year after completion of the training, the organization generally wants to know why some graduates from the program were successful while others were not. Did the training these individuals receive actually help them to start up a small business? Were there topics that should have been covered to more thoroughly prepare them for the realities of setting up a business? Were there other topics that should have been addressed? In summary, this organization wants not only to learn if at least 70% of the individuals trained have started up a business, but how effectively the training equipped them to do so, what were both the strengths and shortcomings of the training that can be identified to improve future training programs.

The organization also wants to understand what some of the key factors were -outside of the actual intervention- that had a bearing on success or failure. For

example, did some individuals find employment instead? Was access to finance a problem? Did they conduct an adequate market analysis? Did some individuals start with prior business skills? There could be numerous reasons.

***The selection of additional methods to be employed, is again, based on the nature of each aspect of the issue or set of related questions that the organization wants to probe.***

To continue with this example, the evaluator might expand the number of survey questions to address issues related to the effectiveness of the training and external factors such as access to finance. These additional questions can be designed to yield additional quantitative data to probe questions such as level of satisfaction with the training program, degree of utility of the training program in establishing the individual's business, whether the training graduate was approved or turned down for a small business start-up loan, if the size of the loan they received was sufficient, whether the individual used their own capital to start up their business, whether they are still in the process of starting up their business, or whether the individual stopped trying or instead found employment. Intake data from the training program on characteristics of each trainee can also be examined to see if there are any particular characteristics, such as sex or ethnic background, which

can be correlated with the survey findings.

But to help explain the statistical findings derived from survey responses, probe the strengths and shortcomings of the training program, further understand issues related to access to finance, and to determine if there were any unanticipated factors that had a bearing on success in starting a business, it is important to draw on additional methods. In this case, the evaluation design would include open-ended topic guides to conduct in-depth interviews with a random sample of individuals who graduated from the training program to further address these questions.

In this example assume that the organization has learned, through informal feedback, that access to finance is likely a key external factor in determining success in business start-up in addition to the training program itself. Depending on the evaluation findings, the organization may design a finance program that increases access to loans for small business start-ups. To determine the validity of this assumption, the evaluation design relies on a triangulation approach to assess whether and how access to finance for business start-ups provides further explanations regarding success or failure outcomes. The design includes a plan to collect data from two other sources using a separate data collection method for each source. The evaluation designers determine that the second data source will

be comprised of managers of local banks and credit unions that survey respondents indicated were approached for start-up loans. In-depth interviews will be conducted to record and understand policies for loaning to entrepreneurs trying to establish small businesses, the application of those policies, and other business practices with respect to prospective clients. The third data source is comprised of bank loan statistics for entrepreneurs who have applied to start up small businesses. Now there are three independent data sources using different data collection methods to assess whether access to finance is an additional key factor in determining small business start-up success.

Together, the survey of all 50 training graduates, use of training intake data on characteristics of trainees, in-depth interviews with a sample of graduates, interviews with institutions and organizations that provide loans for start-up businesses, and an examination of loan data constitute the total mix of methods the evaluator would use in this example. The use of mixed methods was necessary because the organization in this case not only wanted to know how effective the training course was based on its own measure of program success, but also, whether access to finance contributed to either success or failure in starting up a new business.

The last step in the process of designing a mixed method

evaluation is to determine how the data derived from using mixed methods will be analyzed to produce findings and to determine the key conclusions.

## **ANALYZING DATA FROM A MIXED METHODS EVALUATION – DESIGNING A DATA ANALYSIS STRATEGY**

It is important to design the data analysis strategy before the actual data collection begins. Having done so, the evaluator can begin thinking about trends in findings from different data sets they are collecting to see if there is either a convergence or divergence in findings. Analyzing data collected from a mixture of methods is admittedly more complicated than analyzing the data derived from one data source and the use of a single data collection. This involves a process whereby quantitative and qualitative data analysis strategies are eventually connected to determine and understand key findings. There are several different analytical techniques that are used to analyze data from mixed methods, including parallel analysis, conversion analysis, sequential analysis, multilevel analysis, and data synthesis. The choice of analytical techniques should be matched with the purpose of the evaluation using mixed methods. Table 1 briefly describes the different techniques

that are used for the analysis of data collected using mixed methods and in what situations the method is best applied. In complex evaluations with multiple issues to address, skilled evaluators may use more than one of these techniques to analyze the data.

## EXAMPLE OF APPLICATION

Here we present an example of parallel mixed data analysis because is the most widely used analytical technique employed in mixed method evaluations. This is followed by examples of how to resolve situations where divergent findings arise from the analysis of data collected through a triangulation process.

### PARALLEL MIXED DATA ANALYSIS

The most common strategy used to analyze data from mixed methods is called parallel mixed data analysis, and it is comprised of two major steps.

**Step 1:** This involves two or more analytical processes. The data collected from *each* method employed, usually a mix of quantitative and qualitative methods, must be analyzed separately. For example, a statistical analysis of quantitative data derived from surveys, height/weight measures, or government statistics is conducted. Then, a separate and independent analysis is conducted of qualitative data derived from, for example, in-

depth interviews, case studies, focus groups, or structured observations to determine emergent themes, broad patterns, and contextual factors. ***The main point is that the analysis of data collected from each method must be conducted independently.***

**Step 2:** Once the analysis is completed from the data generated by each data collection method, the evaluator focuses on how the analysis and findings from each data set can inform, explain, and/or strengthen findings from the other data set. There are two primary analytical methods used in the second step of parallel analysis, and sometimes both methods are used in the same evaluation. Again, the method used depends on the *purpose* of the evaluation.

- In cases where more than one method is used specifically to strengthen and validate findings to the same question through a triangulation design, the evaluator compares the findings from the independent analysis conducted on each data set generated by each data collection method to determine if there is a convergence of findings. This method is used when it is critical to produce defensible conclusions that can be used to make major program decisions (e.g., end or extend a program).
- To interpret or explain findings from quantitative analysis, evaluators use findings from the analysis of qualitative data.

This method can provide a richer analysis and set of explanations affecting program outcomes that enhances the utility of the evaluation for program managers. Conversely, patterns and associations surfaced from the analysis of quantitative data can inform additional patterns to look for in analyzing qualitative data. The analysis of qualitative data can also enhance the understanding of important program context data. This method is often used when program managers are looking for not only whether or not a program is achieving its intended results, but also, why or why not.

### WHEN FINDINGS DO NOT CONVERGE

In cases where mixed-method evaluations employ triangulation to corroborate and strengthen the validity of findings, it is not unusual that findings from the separate analysis of each data set do not automatically converge. If this occurs, divergent findings must try to be resolved. 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 evaluators use when findings from different methods do not corroborate each other is to more carefully re-examine the qualitative data through a second and more in-depth content analysis. The results of this third layer of analysis can produce a

## REPORTING ON MIXED-METHOD EVALUATIONS

deeper understanding of the data, and can then be used to generate new interpretations to resolve what first appeared as divergent findings. In some cases, other factors external to the program might be discovered through further analysis such as different economic or political conditions found at different program site locations or variations on how the program intervention was implemented at different program sites. Another approach is to reanalyze all the disaggregated data in each data set, separately, by characteristics of the respondents as appropriate to the study, such as age, gender, position, educational background, economic strata, etc., and/or by geography/locale of respondents. The results of this analysis may yield other information on additional factors to help resolve the divergence of findings. In this case, the evaluator should attempt to rank order these factors in terms of frequency of occurrence to both strengthen and provide further explanations to the findings. While most professionals 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.

Evaluators should also check for data quality and data validity

issues such as problems arising from the validity of secondary data sources, possible errors in survey data from incomplete recording or incorrect coding of responses to determine if inconsistencies can be resolved. Sometimes additional, limited follow-up data collection may be warranted if the evaluators are still at the program site (if time and budget permit), such as conducting in-depth interviews with key informants, for example.

Another analytical tool that is used to resolve divergent findings in cases where an overall summative program conclusion is required is the data synthesis method (described briefly in Table 2) whereby ratings of the strength of findings generated from the analysis of each data set collected is assigned based on 1) size of program effects on outcome and 2) the quality and validity of each data set. An overall rating is assigned for each data set, but different weights can then be assigned to different data sets if the evaluator knows that certain data sources or methods for collecting data are the agreed upon standard. Lastly, an index is created based on the average of those ratings to synthesize overall program effect on the outcome. See McConney, Rudd and Ayres (2002) to learn more about this method.

Mixed-method evaluations generate a great deal of data and to profit from the use of those methods, evaluators must use and analyze all data sets. By using mixed-method evaluations, findings and their explanations can be enriched and strengthened. Yet there is a tendency to underuse, or even not use, the data collected from all methods employed, with a concurrent over-reliance on one particular data source if it generates easily digestible and understandable information for a program manager. For example, frequently data generated from qualitative methods are insufficiently analyzed. In some cases only findings from one source are reported.

One way to prevent this from happening 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 program manager can use with confidence. Additionally, scopes 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.

## REFERENCES

- Bamberger, Michael, Jim Rugh and Linda Mabry. *Real World Evaluation: Working Under Budget, Time, Data and Political Constraints*, Chapter 13, "Mixed-Method Evaluation," pp. 303-322, Sage Publications Inc., Thousand Oaks, CA, 2006.
- Greene, Jennifer C. and Valerie J. Caracelli. "Defining and Describing the Paradigm Issue in Mixed Methods Evaluation," in *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms*, Green and Caracelli eds. *New Directions for Evaluation*. Josey-Bass Publishers, No. 74, Summer 1997, pp 5-17.
- Mark, Melvin M., Irwin Feller and Scott B. Button. "Integrating Qualitative Methods in a Predominantly Quantitative Evaluation: A Case Study and Some Reflections," in *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms*, Green and Caracelli eds. *New Directions for Evaluation*. Josey-Bass Publishers, No. 74, Summer 1997, pp 47-59.
- McConney, Andrew, Andy Rudd, and Robert Ayres. "Getting to the Bottom Line: A Method for Synthesizing Findings Within Mixed-method Program Evaluations," in *American Journal of Evaluation*, Vol. 3, No. 2, 2002, pp. 121-140.
- Teddlie, Charles and Abbas Tashakkori, *Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Behavioral Science*, Sage Publications, Inc., Los Angeles, 2009.

**TABLE 1 – METHODS FOR ANALYZING MIXED METHODS DATA<sup>1</sup>**

Analytical Method	Brief Description	Best for...
Parallel	Two or more data sets collected using a mix of methods (quantitative and qualitative) are analyzed <i>independently</i> . The findings are then combined or integrated.	Triangulation designs to look for convergence of findings when the strength of the findings and conclusions is critical, <i>or</i> to use analysis of qualitative data to yield deeper explanations of findings from quantitative data analysis.
Conversion	Two types of data are generated from <i>one data source</i> beginning with the form (quantitative or qualitative) of the original data source that was collected. Then the data are <i>converted</i> 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 to <i>generate additional findings and/or to compare and potentially strengthen the findings</i> generated from a complimentary set of data (e.g., QUAN) from a different data source using a different data collection method from the method used to collect data for the set (QUAL) that was converted.
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 <i>inform</i> the analysis of the second data set. The type of analysis conducted on the second data set is <i>dependent</i> 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 <i>two</i> data sources to answer <i>interrelated</i> 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).	Evaluations where organizational units for study are <i>nested</i> (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 <i>rating</i> of program effectiveness using the analysis of each data set is conducted (e.g., large positive effect, small positive effect, no discernable 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 <i>quality and validity</i> 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 <i>and</i> validity of <i>each</i> finding; and 4) <i>average</i> outcome-wise effectiveness estimates to produce <i>one</i> overall program-wise effectiveness index.	Providing a bottom-line measure in cases where the evaluation purpose is to provide a <i>summative program-wise conclusion</i> when findings from mixed method evaluations using a triangulation strategy <i>do not converge and appear to be irresolvable</i> , yet a defensible conclusion is needed to make a firm program 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.

<sup>1</sup> 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 explanation of data synthesis analysis.

**For more information:**

TIPS publications are available online at [insert website].

**Acknowledgements:**

Our thanks to those whose experience and insights helped shape this publication including USAID's Office of Management Policy, Budget and Performance (MPBP). This publication was authored by Pat Vondal of Management Systems International.

Comments regarding this publication can be directed to:

Gerald Britan, Ph.D.

Tel: (202) 712-1158

[gbritan@usaid.gov](mailto:gbritan@usaid.gov)

Contracted under RAN-M-00-04-00049-A-FY0S-84

Integrated Managing for Results II