

INTRODUCTION TO MIXED METHODS IN IMPACT EVALUATION

Michael Bamberger



This is the third guidance note in a four-part series of notes related to impact evaluation developed by InterAction with financial support from the Rockefeller Foundation. The other notes in this series are: Introduction to Impact Evaluation; Linking Monitoring & Evaluation to Impact Evaluation; and Use of Impact Evaluation Results. The complete series can be found on InterAction's website at: <http://www.interaction.org/impact-evaluation-notes>.

Photo: Alissa Everett

CONTENTS

Introduction	1
Part I. Why Mixed Methods?	3
1.1. What is a mixed methods impact evaluation design?	3
1.2. The limitations of an exclusive reliance on QUANT or QUAL evaluation approaches	3
1.3. The benefits of a mixed methods approach	4
Part II. The Mixed Methods Approach	9
2.1. Four decisions for designing a mixed methods evaluation	9
2.2. Applying MM approaches at each stage of the evaluation	13
Part III. Applications of Mixed Methods Designs	19
3.1. Sampling strategies for QUANT and QUAL oriented MM evaluations	19
3.2. Using mixed methods to evaluate complex interventions	21
3.3. Assessing processes of behavioral change	24
Part IV. Managing Mixed Methods Evaluations	27
4.1. Mixed methods designs require a special management approach	27
4.2. Tips for resource constrained NGOs to mobilize the expertise and resources required to conduct mixed methods evaluations	30
Part V. Case Studies Illustrating Different Applications of Mixed Methods Designs	32
References	37

TABLES

Table 1. Mixed methods are used differently for evaluation designs with a dominant QUANT or QUAL orientation	11
Table 2. Widely used QUANT and QUAL data collection methods	16
Table 3. Different types of triangulation used in mixed method evaluations	17
Table 4. Examples of mixed method data analysis	18
Table 5. Additional areas of research expertise that may be required for QUANT and QUAL oriented teams to conduct MM evaluations	29
Table 6. Examples of MM evaluations summarized in Annex 10	36

BOXES

Box 1. How “impacts” and “impact evaluation” are used in the guidance notes	3
Box 2. Operational benefits from the use of mixed methods	5
Box 3. Attribution, contribution and substitution analysis	22
Box 4. Program outcomes are affected by how beneficiaries respond to services and how staff react to beneficiary responses. Studying behavioral change in a school feeding program in Nicaragua.	25
Box 5. Planning and budgeting additional time and resources that may be required for a mixed method evaluation	28
Box 6. Conducting mixed methods impact evaluations on a shoestring: tips for NGOs with limited resources and evaluation expertise	31

FIGURES

Figure 1. Using a national household sample frame to ensure the representativeness and credibility of a QUAL case study evaluation	6
Figure 2. Using QUAL ethnographic village and household studies to help interpret the findings of a national QUANT sample survey	8
Figure 3. Sequential mixed method design with a dominant quantitative approach: Studying interhousehold transfers as a survival strategy for low-income households in Cartagena, Colombia	12
Figure 4. Sequential mixed method design with a dominant qualitative approach: Evaluating the adoption of new seed varieties by different types of farmers	12
Figure 5. Multilevel mixed methods design: Evaluating the effects of a school feeding program on attendance and performance	14
Figure 6. Using mixed method designs for evaluating complex interventions	23

ANNEXES (available at <http://www.interaction.org/impact-evaluation-notes>)

Annex 1. Strengths and weaknesses of quantitative evaluation designs
Annex 2. Strengths and weaknesses of qualitative evaluation designs
Annex 3. Examples of evaluation designs at each point on the QUANT - QUAL continuum
Annex 4. Characteristics of QUANT and QUAL approaches at different stages of the evaluation
Annex 5. How QUANT and QUAL approaches complement each other at different stages of an evaluation
Annex 6. Comparing random and purposive sampling methods
Annex 7. A range of quantitative, qualitative and theory-based approaches for defining the counterfactual
Annex 8. Strategies for reducing the costs of data collection and analysis
Annex 9. Example of triangulation: comparing estimates of household income and poverty from different sources
Annex 10. Case studies of MM evaluation designs with predominant QUANT, QUAL and balanced orientations
Annex 11. How mixed methods can strengthen QUANT evaluation designs

Introduction

Mixed methods (MM) evaluations seek to integrate social science disciplines with predominantly quantitative (QUANT) and predominantly qualitative (QUAL) approaches to theory, data collection, data analysis and interpretation. The purpose is to strengthen the reliability of data, validity of the findings and recommendations, and to broaden and deepen our understanding of the processes through which program outcomes and impacts are achieved, and how these are affected by the context within which the program is implemented. While mixed methods are now widely used in program evaluation, and evaluation RFPs frequently require their use, many evaluators do not utilize the full potential of the MM approach.

This guidance note explains the essential elements of the MM approach and how it can be used in an impact evaluation (IE), while highlighting potential applications and benefits for NGOs. Part I addresses the question, “Why mixed methods?” We discuss what an MM impact evaluation design is, what distinguishes it from a QUANT or QUAL impact evaluation design and why the approach is helpful for understanding development evaluations and the complexities of the real world within which they are implemented (section 1.1). The increasing popularity of MM comes from the recognition of the limitations of an exclusive reliance on either QUANT and QUAL methods (section 1.2), and the

potential benefits that can be achieved when both approaches are appropriately combined (section 1.3). While MM can be used as part of a large and well-funded impact evaluation, the methods have the flexibility to be equally useful for the many NGOs that require credible evaluations of their programs, but whose resources and expertise for conducting impact evaluations are limited.

Having laid out the case for MM, Part II then describes four key decisions that have to be made when designing an MM evaluation (section 2.1): at which stages of the evaluation will MM be used; is the MM design sequential or concurrent; will the

design be predominantly QUANT or QUAL, or will a balanced design be used that gives equal weight to both approaches; and will the design be implemented on a single level or will it be a multilevel evaluation? Section 2.2 explains how MM can be used at each stage of the design, implementation and analysis of an evaluation—not just as a way to diversify data collection methods, as many evaluators assume.

Part III reviews some applications of MM designs. Section 3.1 discusses MM sampling strategies when using predominantly QUANT or QUAL designs and shows how MM can strengthen both kinds of sample design. Section 3.2 discusses the use of MM for evaluating complex development interventions, and section 3.3 how MM designs can help evaluate programs that involve processes of behavioral change.

Part IV addresses issues in the management of MM evaluations. Section 4.1 explains why a special management approach is required, and section 4.2

discusses how NGOs can effectively utilize MM for evaluations that are conducted “on a shoestring” (i.e., with budget and time constraints and with limited research expertise).

Section V presents three case studies illustrating how MM are used in predominantly QUANT and QUAL evaluations, and in a balanced evaluation giving equal weight to both QUANT and QUAL approaches. [Annex 10](#) presents 17 examples of MM evaluations illustrating a wide range of approaches and including both large, well-funded evaluations and evaluations conducted under budget, time and data constraints.

A challenge in preparing this guidance note (GN) was the fact that there is a very extensive literature on MM, some of it quite technical or specialized. We have tried to keep the text accessible to a wide and non-specialist audience while providing a set of annexes (available at <http://www.interaction.org/impact-evaluation-notes>) that go into more detail.

Part I. Why Mixed Methods?

1.1. What is a mixed methods impact evaluation design?

Different evaluators use the terms “impact” and “impact evaluation” in different ways. So, in order to ensure that we are all on the same page, Box 1 summarizes how the terms are used in these guidance notes. Mixed methods have the same objectives, ask many of the same questions and draw on all of the impact evaluation tools and techniques described in guidance note 1 of this series (GN1), and build on project monitoring and evaluation systems in the ways described in GN2. Like other impact evaluation designs, MM can be applied at any level, from the evaluation of a project operating in a single village to a multicomponent national development initiative involving many different international and national agencies.

There is rarely a single evaluation methodology that can fully capture all of the complexities of how programs operate in the real world. Consequently, evaluators must find creative ways to combine different evaluation frameworks, tools and techniques¹—hence the growing interest in MM approaches. The unique feature of mixed methods approaches is that they seek to integrate social science disciplines with predominantly QUANT and

¹ An important related topic concerns the choice of the appropriate evaluation design. Given the many different kinds of programs that are evaluated, the varied contexts in which they operate and the diversity of evaluation questions of interest to stakeholders—there is no single “best” evaluation design that will work in all situations. The choice of evaluation design requires a careful analysis of the nature of the program, the purpose and context of the evaluation, and the environment within which it operates. See Bamberger, Rugh and Mabry (2012) Chapters 2 and 11 for a discussion of strategies for selecting the most appropriate evaluation design. See also Stern et al (2012) for a detailed discussion of choice of methods.

predominantly QUAL approaches to theory, data collection and data analysis and interpretation. Although many evaluators now routinely use a variety of methods, “What distinguishes mixed-method evaluation is the intentional or planned use of diverse methods for particular mixed-method purposes using particular mixed-method designs” (Greene 2005:255). Most commonly, methods of data collection are combined to make an evaluation MM, but it is also possible to combine conceptual frameworks, hypothesis development, data analysis, or frameworks for the interpretation of the evaluation findings.

BOX 1. HOW “IMPACTS” AND “IMPACT EVALUATION” ARE USED IN THE GUIDANCE NOTES

Guidance Note 1, “Introduction to Impact Evaluation” (page 2), defines impacts as:

the positive and negative, intended and unintended, direct and indirect, primary and secondary effects produced by an intervention. (OECD Development Assistance Committee definition).

Impacts are usually understood to occur later than, and as a result of, intermediate outcomes. The distinction between outcomes and impacts can be relative, and depends on the stated objectives of an intervention. According to our definition, an impact evaluation includes:

any evaluation that investigates systematically and empirically the impacts produced by (that can attributed to) an intervention.

1.2. The limitations of an exclusive reliance on QUANT or QUAL evaluation approaches

When used in isolation, both QUANT and QUAL evaluation methods have strengths and weaknesses. The purpose of MM is to draw on the

strengths of both QUANT and QUAL approaches and integrate them to overcome their weaknesses. Despite the many powerful benefits of QUANT data collection and analysis methods, they also have a number of inherent limitations ([Annex 1](#)). Many of the criticisms concern the reduction of narrative data into numbers, and inflexible designs and data collection methods procedures that are difficult to adapt to changing circumstances. The standardized categories in questionnaires and data coding often fail to capture nuances within the groups or communities studied, and the analysis often lacks the depth and detail of QUAL methods. QUANT evaluation risks becoming *decontextualized*, ignoring how programs are affected by the economic, political, institutional and socio-cultural characteristics of the populations studied. Another frequent criticism of many QUANT evaluations is that they assume that programs operate as planned and that everyone receives the same package of services (both in terms of quantity and quality). This is often referred to as the “black box” approach, as the evaluation does not look inside the project “black box.”

QUAL methods are also powerful tools for data collection and analysis. However, when used on their own, QUAL evaluation designs also have a number of potential weaknesses ([Annex 2](#)). QUAL evaluations often focus on individual subjects and situations and it is more difficult to generalize from the findings. Many QUAL evaluators also believe that each evaluation is context-specific and it is not possible or appropriate to generalize. Many, but certainly not all, QUAL evaluations use a holistic approach, making individual elements and factors harder to isolate and making it more difficult to understand the specific contribution of different components or approaches of the program. Some clients also feel uncomfortable that there may seem to be too much reliance on the opinion and

perspective of the evaluator, with no way for the reader to easily review the large amounts of written and recorded data that the evaluator has drawn on. A final point is that many QUAL evaluations do not provide the kinds of detailed documentation on the methodology that are usually presented in QUANT evaluation reports, making it difficult to check on the validity of the data collection and analysis procedures.²

1.3. The benefits of a mixed methods approach³

There are five main reasons for using mixed-method designs (Greene 2005:255–56):

- *Triangulation of evaluation findings*: enhancing the validity or credibility of evaluation findings by comparing information obtained from different methods of data collection (for example comparing responses to survey questions with what the interviewer observes directly). When estimates from different sources converge and agree this increases the validity and credibility of findings or interpretation. When different estimates are not consistent, the researcher explores further to understand the reason for the inconsistencies (see [Annex 9](#)).
- *Development*: using results of one method to help develop the sample or instrumentation for another.
- *Complementarity*: extending the comprehensiveness of evaluation findings through results from different methods that broaden and deepen the understanding reached.

² For example, many QUAL evaluations do not include detailed documentation on how focus group members were selected, and few can provide transcripts of interviews (for considerations of cost and time), so it is normally not possible for the reader to independently review the data and to assess how the findings and conclusions were reached.

³ For a recent review of the benefits of mixed methods approaches see Adato (2012).

- *Initiation*: generating new insights into evaluation findings through results from the different methods that diverge and thus call for reconciliation through further analysis, reframing or a shift in perspective.
- *Value diversity*: incorporating a wider diversity of values through the use of different methods that themselves advance difference values. This encourages greater consciousness about the value dimensions of the evaluation.

An additional benefit is that an MM approach is more likely to ensure buy-in from both QUANT- and QUAL-oriented evaluators and users.

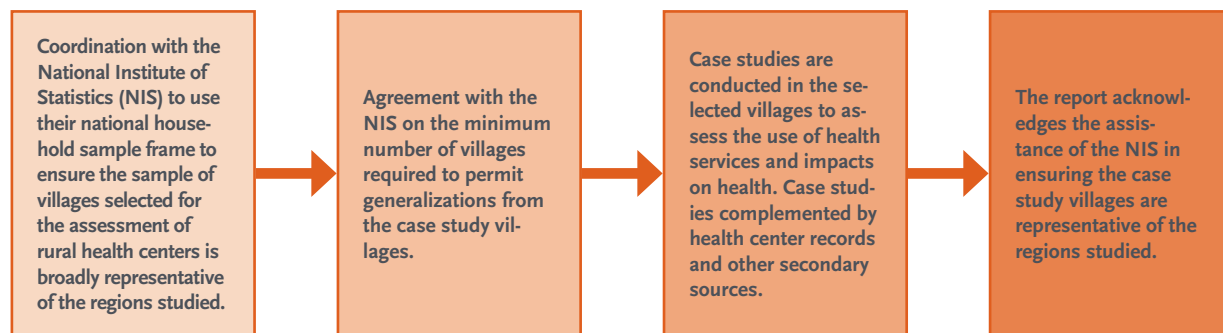
Box 2 summarizes some of the operational benefits of using an MM approach.

To illustrate some of these benefits, let us take the example of an evaluation of the impact of rural health centers on the health of the rural population, especially women and children. A particular concern is the accessibility of health services to poor and vulnerable sectors of the population. In this first example, the evaluation adopts a predominantly **QUAL approach**. The evaluation is intended to influence national health policies by identifying some of the reasons why poor and minority families do not use the health centers. The evaluation team is aware that the Ministry of Health has criticized earlier QUAL evaluations for having intentionally focused on communities known to have particular problems, meaning that the findings are not representative of the whole country and will present the Ministry of Health in a poor light. The evaluators are also aware that the

BOX 2. OPERATIONAL BENEFITS FROM THE USE OF MIXED METHODS

- Understanding how local contextual factors help explain variations in program implementation and outcomes.
- Reconstructing baseline data for QUANT evaluations when it was not possible to conduct a baseline survey. Many evaluations are commissioned toward the end of the program and do not have very reliable information on the conditions of the project and comparison groups at the time the program began. This makes it difficult to determine whether observed differences at the end of the project can be attributed to the effects of the program or whether these differences might be due, at least in part, to preexisting differences between the two groups. For example, women who apply for small business loans may come from families that are more supportive of women owning a small business than most families, or they may already have more business experience than women who do not apply for loans. If these preexisting differences are not identified, there is a risk of overestimating the effects of the loan program. It is often possible to use such QUAL techniques as in-depth interviews, key informant interviews or focus groups to obtain information of the characteristics of program beneficiaries and nonbeneficiaries at the time the program began. This kind of information, which is often quite simple and economical to collect, can greatly enhance the validity of exclusively QUANT survey data.
- Strengthening the representativeness of in-depth QUAL studies (for example, by linking case study selection to the QUANT sampling frame) can make it easier to compare findings with QUANT survey data.
- Providing a good sense about validity and value of different kinds of QUANT and QUAL data.
- Promoting greater understanding of stakeholder perspectives on the nature of the intervention or how it is expected to achieve its objectives. This promotes a more participatory approach and greater alignment between stakeholders and evaluators.

Figure 1. Using a national household sample frame to ensure the representativeness and credibility of a QUAL case study evaluation



ministry has sometimes used the lack of statistical representativeness as a convenient excuse for ignoring valid criticisms, and they want to ensure that their study will not be dismissed on these grounds. Consequently, the evaluation team coordinates with the National Institute of Statistics and uses their national household sample frame to ensure that the sample of communities they select is broadly representative of the whole country (or the region where the study is conducted). The evaluation uses the same QUAL methods, but it is now possible to indicate that the sample of communities is broadly representative of all communities in the regions studied (see Figure 1).

Let us now assume that the same evaluation is to be conducted by a different team planning to use a **QUANT approach** based on a nationally representative household sample survey. While a well-designed survey can obtain reasonably reliable estimates of the proportion of the population using the health centers (though even then there is a potential for misreporting), the evaluators are fully aware that the survey will not provide a good understanding of the reasons why households do, or do not, use the health centers. Consequently, they invite an ethnographer to join their team and conduct in-depth studies in a small number of communities. The

ethnographic studies will explore the health-related attitudes and beliefs of different ethnographic groups and the factors influencing their decision to use the health centers or not. The studies will also examine the economic, political, organizational, cultural and ecological factors affecting the operation of the health centers in different communities. The first part of the analysis will address broad cultural differences likely to affect all health centers, and the latter part (the contextual analysis) will help to explain factors affecting the performance of different centers (Figure 2). The evaluation director is aware that mixed method designs work well only when there is respect and understanding and a feeling of equality among team members from different professions, so the ethnographer was invited to join the team from the time of the first planning meeting. The following are some of the ways in which the QUANT and QUAL approaches can be integrated into this evaluation:

- Rapid ethnographic studies (QUAL) are conducted in selected communities to understand the issues that must be addressed in the survey and to help phrase the questions.
- A QUANT household survey using a nationally representative sample is then conducted.
- The analysis of the household survey can

produce a typology of households according to their level of use of the health centers, or their reasons for not using. A sample from each type will be selected to prepare case studies. The case studies will often reveal that reported use or reasons for not using are not correct. For example, women will usually not mention sexual harassment in response to the survey and may instead give a reason such as the opening hours are not convenient.

- Triangulation will be used to obtain independent QUANT and QUAL estimates for key variables (such as use of health facilities and attitudes toward these facilities). A key feature of triangulation is that procedures are built in to identify any inconsistencies in different estimates and to follow up to understand the reason for the differences. For example, observation of how people entering the clinic are received, or spending time with households and discussing informally whether and when they use the clinics, will be compared with survey responses.
- Separate draft QUANT and QUAL reports will be prepared, and the teams will then meet to identify any areas on which there are apparent differences of facts or interpretation. In the example of inconsistencies between survey response on utilization of health centers and data from observation and in-depth interviews, the QUANT and QUAL researchers will meet to discuss the reasons for the inconsistencies. They may agree that one or other sources of information is more reliable. For example, in-depth interviews with women when no other household members are present may be considered more reliable. However, if it is not clear which source is more reliable, then researchers might return to the field to collect more data or other sources of information might be sought—for example, a review

of health center records on patient visits, or consultations with key informants, such as community leaders, nurses, etc.

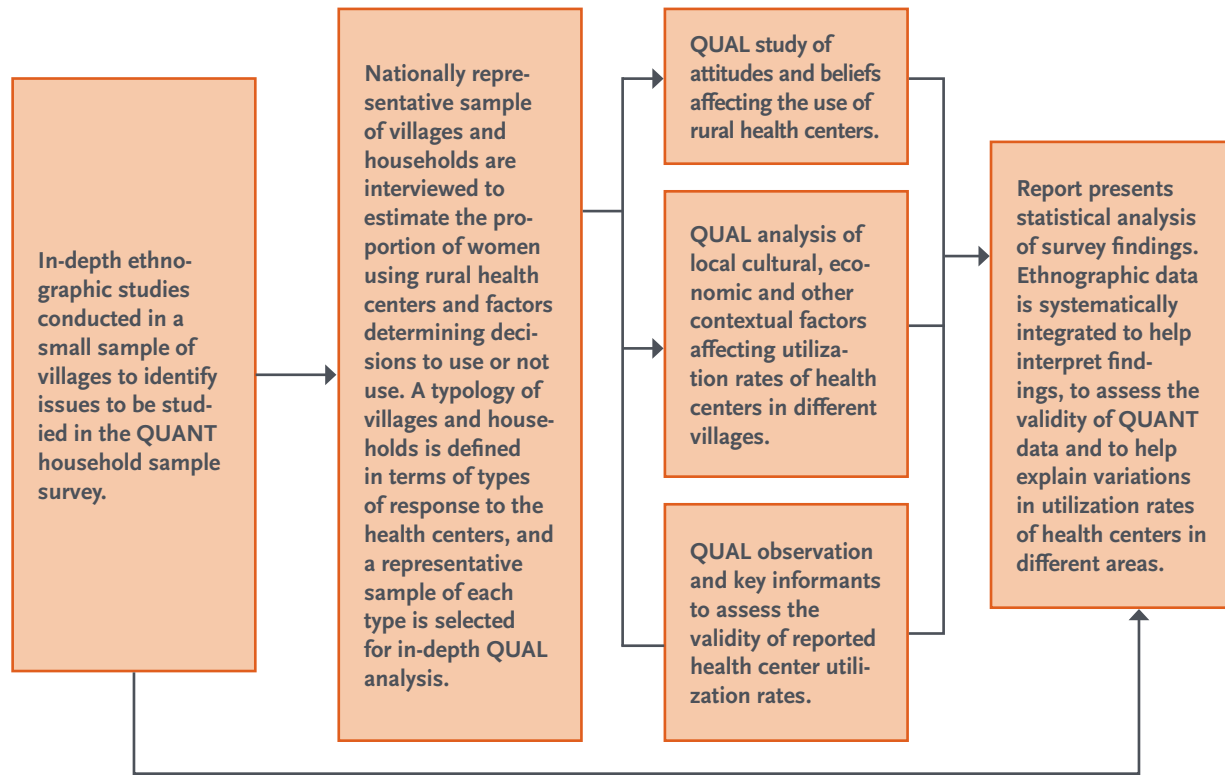
In both of these cases, MM can strengthen the evaluation. However, the focus is quite different when MM are used to strengthen a predominantly QUAL designs (Figure 1) than to strengthen a predominantly QUANT design (Figure 2).

Generally speaking, an MM approach is particularly helpful for:

- Examining the interactions among the complex and changing contextual factors that can influence program implementation and impacts.
- Defining and measuring indicators of the cultural, historical, political, legal, environmental and psycho-social factors that affect implementation. Different methodologies are required to measure these.
- Capturing complex processes of organizational and behavioral change (sections 3.2 and 3.3).
- Taking into account how programs change in response to how they are perceived and used by different sectors of the target population. The experience of early users and the feedback they give to neighbors can dramatically affect how a program evolves.⁴
- Many processes and outcomes are difficult to observe, or in some cases even to know they exist. This is particularly important for evaluating the situation of vulnerable groups and for programs that affect illegal or socially disapproved activities, such as drug use, sex work or illegal immigration. All of these challenges are multiplied for post-conflict, humanitarian and other kinds of emergency relief programs.

⁴ *Realist evaluation* (Pawson 2006) provides a useful framework for the analysis of behavioral change and for the analysis of how programs actually operate in the field.

Figure 2. Using QUAL ethnographic village and household studies to help interpret the findings of a national QUANT sample survey



Part II. The Mixed Methods Approach

2.1. Four decisions for designing a mixed methods evaluation

When planning an MM evaluation, four decisions are required:

1. At which stage or stages of the evaluation will MM be used?
2. Will QUANT and QUAL methods be used sequentially or concurrently?
3. Will QUANT and QUAL methods be given relatively equal weight, or will one methodology be dominant?
4. Will the design be single- or multilevel?

Decision 1: At which stages of the evaluation will mixed methods be used?

Most MM evaluations only combine QUANT and QUAL methods in one or perhaps two stages of the evaluation—most frequently data collection. However, an MM design is much stronger if QUANT and QUAL approaches are integrated into several (or ideally all) stages of the evaluation. Section 2.2 explains how MM can be used to strengthen each stage of an evaluation, [Annex 4](#) compares QUANT and QUAL approaches at each stage of an evaluation, and [Annex 5](#) give examples of how QUANT and QUAL approaches complement each other to strengthen each stage of the evaluation.

Decision 2: Is the MM design sequential or concurrent?

Sequential Mixed-Method Designs

In sequential designs, QUANT and QUAL methods are used in phases. For example, the evaluation may begin with a QUAL exploratory study

to help understand the key issues and how these are perceived by the affected populations. This helps design a QUANT survey, which is then administered to a randomly selected sample. The data could then be analyzed using QUANT and/or QUAL analysis methods. In another example, a rapid QUANT survey could be used to identify and quantify the main kinds of farms and farming activities. This information would then be used to select a representative sample of farms for the preparation of in-depth QUAL case studies. The case studies would probably be analyzed using QUAL methods and the sample survey would be analyzed using QUANT techniques. Figure 3 is an example of a sequential design used to assess interhousehold transfers as a survival strategy of poor families. This evaluation began with an ethnographic (QUAL) study to understand the characteristics of the communities, followed by a QUANT household survey and econometric analysis of the findings.

Concurrent Designs

In concurrent designs, the QUANT and QUAL approaches are used at the same time. An example of a concurrent design is where QUANT and QUAL data are collected simultaneously, using triangulation to compare information on outcomes, impacts and other key indicators from different independent sources. Another example is when QUAL methods are used to conduct a contextual analysis of a project site (or the surrounding areas) at the same time that a QUANT sample survey of households or individuals is being carried out. This provides the opportunity for a very rich but more complicated analysis in which the interactions

between the setting (context) and the project implementation process are analyzed.

Operational considerations in deciding between sequential and concurrent designs

An advantage of sequential designs is that the logistics are often easier to organize. Data collection using structured questionnaires often requires a large team of interviewers in the field following a precisely defined schedule of household selection and number of interviews to be conducted each day. The field supervisors need to know where every enumerator is working, because quality control often involves follow-up visits to a subsample of households. The supervisor must also be on hand to answer questions from the enumerators. In contrast, ethnographic and many other kinds of QUAL methods have a much more flexible schedule in terms of duration and where the researchers will be at any given time. For this and other reasons, concurrent MM designs can often be more difficult to manage, particularly for evaluation teams with only a few experienced supervisors. Concurrent designs can be a particular problem in areas where logistical planning (e.g., travel to sites, places to stay, security) can become difficult to coordinate, and they also make it more difficult to handle feedback, as adjustments would have to be made more

quickly than for sequential designs. On the other hand, concurrent designs have the advantage that data collection and analysis can be completed more quickly.

Decision 3: Will the MM design be predominantly QUANT or QUAL or will a balanced design be used?

It is useful to think of evaluation designs as representing a continuum of approaches ranging from exclusively QUANT approaches through approaches that give equal weight to both QUANT and QUAL methods to exclusively QUAL approaches (Bamberger et al 2012 pp. 324–34; Greene and Caracelli 2003). This is important as different evaluators—who likely began their careers with either a predominantly QUANT or QUAL orientation—may have quite different expectations as to what an MM evaluation will involve. It is also important because, due to the professional orientation of the evaluators, a QUANT or QUAL (rather than a balanced) approach is dominant in most MM evaluations.⁵

Table 1 illustrates how MM are used in evaluations where the dominant approach is QUANT or QUAL and [Annex 3](#) gives examples of evaluation designs at different points on this continuum.

⁵ A new generation of evaluators is emerging who have been trained in MM as an integrated evaluation approach and some studies are starting to appear with a more balanced approach without a dominant orientation, but these are still in a minority. The Journal of Mixed Method Research is a good source for examples of balanced designs.

Table 1. Mixed methods are used differently for evaluation designs with a dominant QUANT or QUAL orientation

Which approach is dominant?	How the dominant approach works	How the other orientation is used to strengthen the design
QUANT	The evaluation typically administers a structured questionnaire to a randomly selected sample of individuals, households, groups, institutions or communities and the analysis mainly relies on econometric or other quantitative methods.	In-depth interviews, observation and group interviews are used to help design the questionnaire. Small samples of cases selected from the main sample can also provide deeper understanding of statistical relationships found in the QUANT analysis. Cases can be representative of each main category identified in the analysis, or used to study outliers or other groups selected purposively.*
Equal weight is given to QUANT and QUAL approaches	QUANT surveys are combined with a range of different QUAL techniques. Sometimes the latter focus on the process and contextual analysis, in other cases the focus is on the same unit of analysis as the surveys (e.g., individuals, households, communities, organizations) but different data collection methods are used.	
QUAL	Case studies, in-depth interviews and other QUAL techniques are applied to relatively small samples of individuals, households, communities or groups.	A rapid QUANT survey is used either to identify the issues or groups to be covered in the in-depth QUAL studies or to show that the QUAL sample is reasonably representative of the total population.
* See Annex 6 and Bamberger et al 2012 pp. 360–61 for a description of purposive sampling.		

Figure 3 describes a sequential design with a dominant QUANT approach. This is a study of interhousehold transfers of money and goods as a survival strategy of poor urban households in Colombia (Wansbrough, Jones and Kappaz 2000). The purpose of the study was to describe the patterns of transfers and to estimate whether they were sufficiently large to act as an informal social safety net providing help to the poorest sectors of the community in times of need. These interhousehold transfers are difficult to identify and measure, so an anthropologist lived in the community for a month to study the patterns of transfers and

to help design the questionnaire for the QUANT survey which was then administered to several hundred households. The data were analyzed using QUANT econometric analysis.

Figure 4 illustrates a sequential design with a dominant QUAL approach. This describes a hypothetical evaluation to assess the adoption of new varieties of seed by different types of rural families. The principal data collection methods are qualitative: interviews, focus groups, observation, and case studies of individual households and small farming communities. The

Figure 3. Sequential mixed method design with a dominant quantitative approach: Studying interhousehold transfers as a survival strategy for low-income households in Cartagena, Colombia

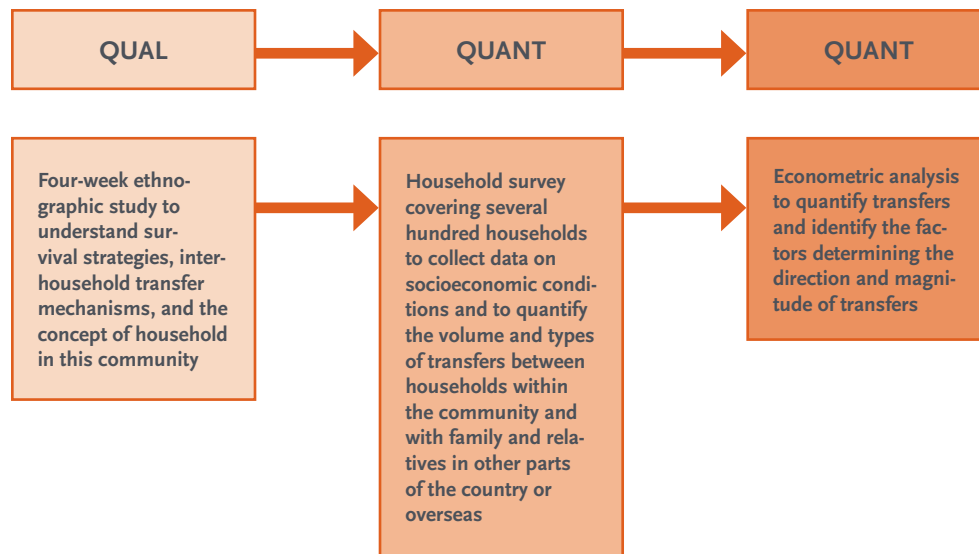
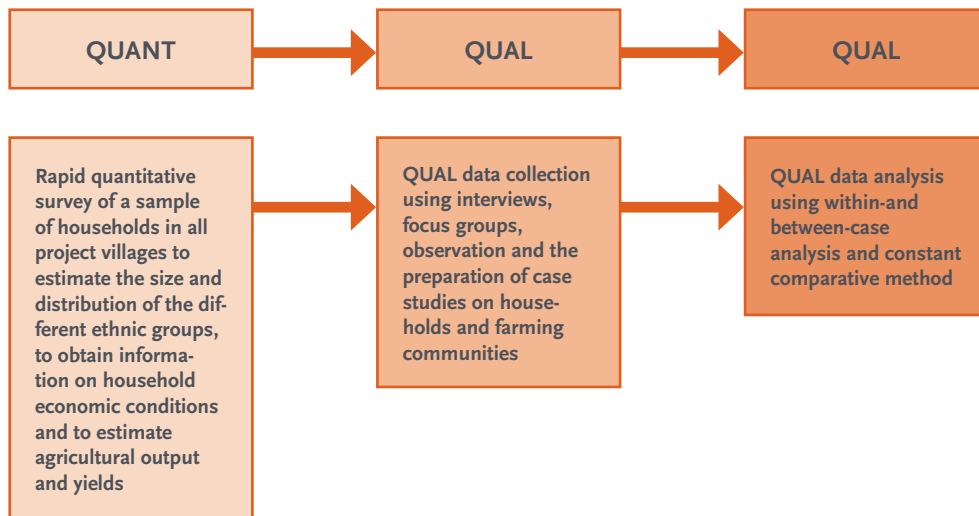


Figure 4. Sequential mixed method design with a dominant qualitative approach: evaluating the adoption of new seed varieties by different types of farmers



principal methods of analysis are also qualitative: within- and cross-case analysis and the constant comparative method. However, to obtain information on the ethnic distribution of households, household economic conditions and agricultural production, the evaluation begins with

a rapid QUANT household survey covering a sample of households in all the villages covered by the agricultural extension project. The findings of this study were used to help identify the types of households to be studied in more depth through the QUAL data collection methods, and

to ensure that the selected cases were broadly representative of the total survey population.

Either of the previous evaluation designs could have been modified to give equal weight to both QUANT and QUAL approaches. In the case of the interhousehold transfer study, the household survey could have been complemented with QUAL case studies on families or informal transfer networks. These could then have been integrated into the analysis to compare the description and interpretation of the functions and operation of the transfer networks obtained from the QUAL studies with the findings of the econometric analysis. In the second example, a QUAL or QUANT study of marketing outlets could have been conducted to estimate the changes in sales of agricultural produce from the project areas and, possibly, the changes in the purchase of consumer goods by project area families.

Example 3 in Part V describes a balanced (integrated) MM evaluation design used to evaluate a large community development program in India. The design gives equal weight to QUANT and QUAL approaches in all stages of the evaluation.

Decision 4: Will the MM design be on a single level or will a multilevel design be used?

The designs we have discussed so far operate on a single level, such as the farm or household. However, MM also provides a powerful tool for the evaluation of service delivery systems (e.g., district education departments, state-level health services, a national program to strengthen municipal governments) that require description and analysis of links between different levels. These evaluations can become very complex and expensive. Mixed method designs that combine QUANT and QUAL data at each level can often provide valid and credible

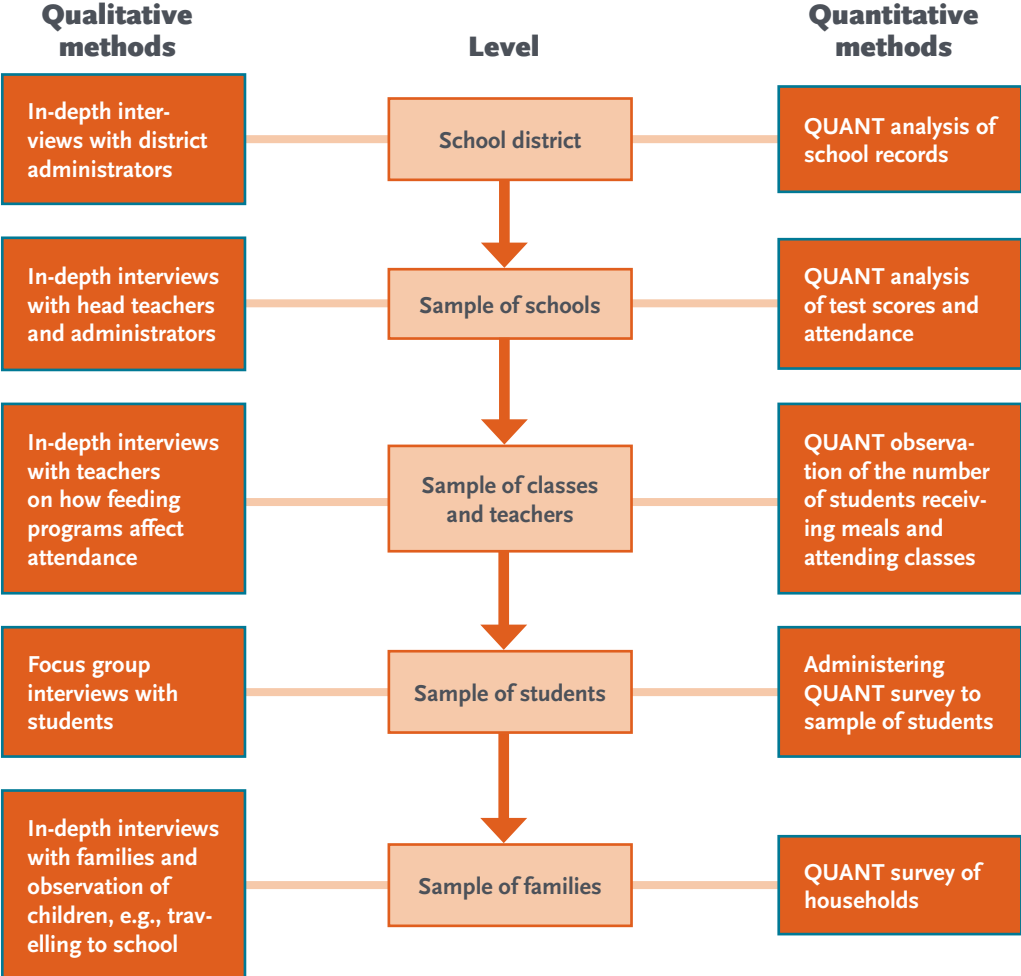
findings on the basis of smaller and more economical samples.

Figure 5 illustrates a multilevel mixed method design to evaluate the effects of a school feeding program on enrolment and attendance. The evaluation must collect data at the level of the school district, a sample of schools, a sample of classes and teachers within each school, and a sample of students and families. At each level, both quantitative and qualitative data are collected and compared. QUAL methods—such as observation, focus groups and key informant interviews—can also help examine linkages between the different levels (e.g., interactions between the district officials and school administrators and teachers).

2.2. Applying MM approaches at each stage of the evaluation

This section explains the different ways that QUANT and QUAL approaches are typically applied at each stage of an evaluation, and how the two approaches can be combined in an MM design. [Annex 4](#) and [Annex 5](#) provide more details. [Annex 11](#) provides examples of how MM can help address common issues arising during evaluation design, data collection and analysis, as well as help promote the utilization of evaluation findings and recommendations. While reading this section it should be understood that for a large, well-funded MM evaluation, the team might include one or more members who are contracted because of their specific QUANT or QUAL expertise (for example, statistical sampling for a large-scale study on malnutrition, or in-depth QUAL interviewing on such sensitive topics as domestic violence). However, many (perhaps most) evaluations do not have this luxury and team members will be required to apply both QUANT and QUAL approaches as required. Of course, even for large evaluations, it is obviously desirable for

Figure 5. Multilevel mixed methods design: Evaluating the effects of a school feeding program on attendance and performance



all team members to have some familiarity with all of the evaluation methods that will be used. The following paragraphs illustrate how QUANT and QUAL methods can be integrated at different stages of the evaluation.

- *Formulation of hypotheses.* QUANT evaluations usually derive hypotheses *deductively* from existing theories or literature reviews, while QUAL evaluations develop hypotheses *inductively* as the study evolves. MM combines both approaches. For example, a hy-

pothesis developed deductively using a QUANT approach can be explored and refined through QUAL approaches, such as interviews or observation. In contrast, the initial stages of QUAL data collection may describe processes and issues that a QUANT approach can test through data collected in a sample survey.

- *Sampling.* QUAL evaluations normally use a relatively small number of subjects selected *purposively* (theoretical sampling) to ensure that all important groups are covered. In

contrast, QUANT evaluations normally use a relatively large, randomly selected sample permitting generalization to larger populations and the statistical comparison of different groups (e.g., the project and comparison groups). MM sampling uses the same sampling frame to generate both a large QUANT survey sample and to select a small but representative sample for in-depth QUAL analysis. *Ensuring that the QUAL samples are reasonably representative of the total sample population is one of the most important contributions of MM designs.* The example of the health center evaluation illustrates how the QUANT and QUAL approaches to sampling can complement each other. [Annex 6](#) summarizes the differences between QUANT and QUAL sampling strategies.

- *Evaluation design.* Most QUANT evaluations use one of a small number of randomized or quasi-experimental designs. Where possible, representative samples of the project and comparison groups are interviewed at two or more points during the life of the project to compare changes in outcome or impact indicators. In contrast, QUAL evaluations try to describe ongoing processes of change that are often affected by many different factors and that affect different individuals or groups in different ways. Some QUAL evaluations seek to understand the program through analysis of relationships among many different elements of the community or other setting in which the program operates, while others adopt a more in-depth focus on individual subjects without

necessarily focusing on the broader context. In some well-funded evaluations the evaluator may live in the community or visit frequently over a period of time, but in most cases this is not possible and reliance is placed on focus groups, in-depth interviews, key informants, etc. Normally QUAL evaluations do not seek to establish a direct cause and effect relationship between project interventions and outcomes. One of the many ways in which the two approaches can be combined is to use QUAL methods to study the project implementation process and the influence of contextual variables on project performance in some of the communities where a QUANT survey of project participants is being conducted.⁶

- *Data collection and recording methods.* Table 2 lists some of the most widely-used QUANT and QUAL data collection techniques. Whereas QUANT evaluations collect standardized numerical data, QUAL evaluations often use less structured data collection methods that provide greater flexibility and that seek to understand the complexities of a situation. While the strength of QUANT data collection methods is that they produce standardized data that measure changes over time or between groups, these methods are not well suited to capture information on sensitive topics or to interview difficult to reach groups. MM data collection builds on the strengths of QUANT data while digging deeper, capturing sensitive data, studying processes and behavioral change.

⁶ There are a number of techniques for transforming QUAL descriptions of contextual factors into QUANT variables (for example Dummy variables) that can be incorporated into regression analysis (see Table 4).

Table 2. Widely used QUANT and QUAL data collection methods

QUANT	QUAL
<ul style="list-style-type: none"> • Structured surveys of households, farms, users of public services etc. • Structured observation guides • Anthropometric measures of height and weight • Anemia and HIS tests using blood sample collection and tests • Automatic counters (e.g., people entering a building) • Sociometric analysis ‡ ** • GIS (generation and analysis of GPS maps)** • Program MIS on inputs and outputs data • Review of institution data—clinic records, school records, etc.** 	<ul style="list-style-type: none"> • In-depth interviews • Key informants • Participant observation • Non-participant observation** • Case studies • Client exit interviews** • Simulated patient studies • Video or audio recording** • Photography • Document analysis** • Artifacts • Group interviews (e.g., focus groups, community meetings)** • Participatory group techniques (e.g., PRA, Most Significant Change) • Internet surveys
<p>‡ Survey techniques to study group formation, how information spreads, identification of opinion leaders and other patterns of social organization in a community or group.</p> <p>** Indicates that these techniques can be used both quantitatively and qualitatively. They are placed in the column where they are most commonly used.</p>	

- *Triangulation.* A key feature of MM is the systematic use of triangulation ([Annex 9](#)). While both QUAL and QUANT evaluators use triangulation to obtain two or more independent estimates of key outcome variables, MM tend to use triangulation more systematically and as integral part of the evaluation design.

QUANT evaluations use triangulation to build consistency checks into survey instruments or to compare secondary data sources with information provided by survey respondents. QUAL evaluations use triangulation more broadly, but often with the focus on deepening and broadening understanding through multiple perspectives obtained from different sources of information rather than as a consistency check. MM designs triangulate QUANT and QUAL estimates (see Table 3). MM uses information obtained through triangulation

to: enhance the reliability and validity of estimates of key indicators by comparing information from different sources; deepening the understanding of the meaning of statistical relationships identified in the quantitative analysis; and ensuring that the perspectives of all key stakeholders, with particular emphasis on poor and vulnerable groups, are captured and compared. If estimates obtained from different sources are consistent this increases the validity and credibility of the data—particularly of estimates based on small samples—and produces more reliable estimates than if all of the resources had been invested in one particular technique, such as a household survey.

- *Data analysis.* QUAL evaluators use a wide range of data analysis methods to identify broad patterns and relations and to obtain a holistic overview of the complex interactions

Table 3. Different types of triangulation used in mixed method evaluations

Method	Examples
Using different conceptual frameworks	Comparing feminist, human rights, social exclusion or economic (e.g., cost-benefit) analysis frameworks
Different methods of data collection	Comparing structured survey, direct observation, secondary data, artifacts
Different interviewers	Comparing interviewer sex, age, ethnicity, economic status, form of dress, language, etc., on responses
Different times	Comparing responses or observations at different times of day, days of the week, times of year
Different locations and contexts	Comparing responses and observations when interviewers conducted in the home when other people are present, in locations where the respondent may be able to speak more freely, in the street and other public places, at work, in the classroom

between a project and the setting within which it is embedded. The purpose of QUANT analysis, on the other hand, is to describe the statistical characteristics of the key variables, to determine the statistical significance of differences between project and comparison groups, and to identify factors contributing to the magnitude and direction of change. Mixed method data analysis uses QUAL analysis to

help understand the meaning that different subjects or groups give to the statistical associations found in the QUANT analysis and to provide cases and examples to illuminate the findings. On the other hand, QUANT analysis can be used to assess how well the cases included in the QUAL studies represent the total population of interest and which if any sectors have not been covered.

Table 4. Examples of mixed method data analysis

Approach	Description	Example
A. Parallel mixed method data analysis	This involves two separate analysis processes: QUANT data are analyzed using conventional QUANT methods (such as frequency tables, cross-tables, regression analysis, etc.) while a separate analysis of QUAL data is conducted using QUAL methods such as content analysis. The findings of the two sets of analysis are then compared.	In the World Bank 2003 Poverty Assessment in Guatemala, separate teams were responsible for collecting QUAL and QUANT data. QUAL analysis was conducted on 5 pairs of villages, representing the main ethnic groups. QUANT data from the same set of villages was analyzed separately and the two sets of data were only integrated in the final stage of the analysis. The combination of the two independent analysis provided a broader political and historical context for understanding the program operation impacts (Teddlie and Tashakkori (2009) Box 11.5).
B. Conversion mixed method data analysis	<p>a. QUAL data are converted into QUANT indicators (“quantitizing”) using rating, scoring and scaling* so that QUANT analysis techniques can be used</p> <p>b. QUANT data are converted to QUAL indicators (“qualitizing”) so that QUAL analysis procedures can be used</p>	<p>a. Data on the political, economic, social, environmental, legal and administrative context within which a project operates is often presented in a narrative, qualitative form. The indicators can be “quantitized” by conversion to dummy variables. For example: “the economy is growing” = 1, “the economy is not growing” = 0. These dummy variables can then be incorporated into the regression analysis.</p> <p>b. In Figure 4 a quantitative typology of farmers could be “qualitized” by producing narrative descriptions of different attitudes towards the adoption of new seed varieties</p>
C. Sequential mixed method data analysis	<p>a. QUAL data analysis is followed by QUANT analysis</p> <p>b. QUANT data analysis is followed by QUAL analysis</p> <p>c. Iterative MM designs. The analysis includes sequential QUANT and QUAL steps</p>	a. In Figure 3, the study of survival strategies begins with a qualitative analysis of narrative reports on the patterns of interhousehold transfers providing support for vulnerable households. The QUAL analysis helps in the design of the QUANT survey of interhousehold transfers which is then analyzed used econometric techniques.
D. Multilevel mixed method analysis	QUANT and QUAL analysis techniques are used at different levels of a multi-level evaluation design	Figure 5 illustrates using multilevel MM analysis to evaluate the impacts of a school feeding program on attendance and performance. Both QUANT and QUAL analysis were conducted sequentially at the level of the district, the school, the classroom, the student and the family. This permitted an analysis of the interlinkages between the different levels.**

* For example, narrative reports on the attitudes of local political groups to a social development program could be converted into a numerical scale where: 3 = “the political group is favorably disposed to the program; 2 = “the political group is neither favorable nor opposed; and 1 = “the group is opposed to the program.” Rating, scoring and scaling are slightly different ways to make the conversion.

** For example, the analysis of district level records can identify schools with above and below average attendance and/or performance scores. This information can be used to select above and below average schools to be included in the sample. Similarly, in-depth interviews with teachers could be used to select a sample of students with particular characteristics of interest who would be included in the focus groups.

Source: Adapted from C. Teddlie and A. Tashakkori 2009 *Foundations of Mixed Methods Research*. Chapter 11, Sage Publications (with permission). Most of the examples were developed by the present author.

Part III. Applications of Mixed Methods Designs

3.1 Sampling strategies for QUANT and QUAL oriented MM evaluations

Model 1: Using mixed methods to strengthen a mainly QUANT evaluation

With most QUANT evaluation designs, sample surveys are administered before and after the project intervention with a control or comparison group. The required sample size for the QUANT surveys are estimated using effect size and statistical power.⁷ QUAL methods can be used to strengthen the design at one or more of the following points:

- **Exploratory or diagnostic study to understand the context and issues before the survey instruments are developed.** These can involve a rapid qualitative study lasting only a few days or longer studies where an anthropologist or sociologist lives in a community during a period of weeks or months. In a large project operating in different geographical or ecological regions, diagnostic studies might be required in a number of different regions or communities. Sometimes the studies will be conducted by an individual researcher (one researcher per community or region) while in other cases the lead researcher might be assisted by a team of

assistants who conduct rapid surveys or conduct participant or nonparticipant observation studies on, for example, community transport systems, women's time use, or production and marketing systems. Although the studies may last for several days or weeks, the primary sampling unit will usually be a community or group and normally only a few groups or communities will be studied. However, large numbers of individuals may be interviewed using unstructured or semistructured data collection techniques and in some cases rapid sample surveys may also be conducted.

- **Focus groups conducted with different segments of the target population.** These can either be conducted during the preparatory stage of the evaluation or after the quantitative surveys have been analyzed and the principal groups of interest have been identified. Ideally three or four focus groups should be conducted with each economic or demographic group of interest to the evaluation (Teddlie and Tashkkori 2009 Table 8.5), although the numbers of groups will often be smaller when working under budget and time constraints.
- **Specialized semistructured modules can be added to a sample survey and administered to a subsample of respondents.** For example, the main survey may be administered to the household head (who in many cultures is likely male) but in a subsample of households the spouse may be interviewed. Sometimes the

⁷ Effect size refers to the size of the change or impact that is being estimated. The larger the change that is being estimated, the smaller the required sample size. Statistical power refers to the probability that the statistical test will correctly identify that there is a real impact. The higher the required level of confidence, the larger the required sample size. See Bamberger et al 2012 Chapter 15 Section 4 for a discussion of sample size estimation.

same interviewer can administer the special module to the spouse (or other household member), but in many cases it will be necessary to arrange a separate interview, often in a location or at a time when the husband or other household members will not be present. Typically modules are administered to 10–25 percent of the original sample.

- **Preparation of case studies to complement the survey.** It is often useful to prepare case studies on a small sample of respondents covered in the survey to provide a fuller understanding of the issues of interest to the evaluation. For example, in the evaluation of an agricultural program, case studies might be conducted to illustrate different kinds of farming systems. For an education project, the cases might cover higher and lower income families, those who live close to and further from the school, or families from different religious or ethnic groups. Again, the number of cases will normally be quite small, although the duration may be quite long. When case studies are prepared on organizations (such as schools or agricultural cooperatives) or cover whole communities (for example, to illustrate the effects of improved transport systems) the study will be more complicated and often significant numbers of individuals will be interviewed for each case.

Model 2: Using a mixed method design to strengthen a QUAL evaluation

Mixed method designs can also be used to strengthen a QUAL evaluation that uses focus groups, participant observation, nonparticipant observation and the preparation of case studies. A challenge for many of these designs is the danger of bias due to the fact that the samples of individuals or groups are not representative. For

example, often people who attend focus groups are those who have strong feelings for or against a project, those who have the time and resources to attend (they may have to arrange transport), or in some cases (often without the knowledge of the evaluator) some participants may be sent by the local government or other group with particular interest. Consequently, much valuable and insightful information is difficult to incorporate into the evaluation report in a credible way. Similarly, many researchers feel more comfortable talking to some groups than to others, so there may be a bias in the selection of the case studies. Usually the sample is relatively small, but the number of interviews will vary depending on the size of the population studied and the required level of precision of the findings.⁸

Mixed method sampling ensures that QUAL cases and informants are selected to be broadly representative of the total population. This strengthens the validity of the overall findings. Quantitative techniques, such as a rapid sample survey, can be a useful way to compare the socioeconomic characteristics of the individuals or groups covered in the qualitative studies with the characteristics of the total population. Usually the QUANT sample size will be relatively small, as the survey is only used to ensure that the case studies or other

⁸ See Bamberger et al. 2012 Table 15.6 p. 389 for some rules of thumb for estimating sample sizes for different kinds of QUAL data collection methods. However, sample sizes will vary depending on the size and complexity of the program being evaluated and the required level of precision of the estimates. If generalizations are to be made from the case studies, focus groups or other methods it is important to ensure that the cases are selected to ensure they are reasonably representative and also to include enough cases for the findings to be considered credible. So while one or two cases can be valuable for illustrating processes or behavior it is rarely appropriate to use such a small sample to make statements such as “Most farmers felt that ...” or “Most mothers believed that the health centers ...”. It is not possible to use conventional statistical procedures to estimate the appropriate sample size with such small numbers of cases, so inevitably judgment must be combined with consultation with stakeholders as to what sample size would be considered reasonable or credible.

QUAL data are reasonably representative. But where more precise estimates are required, the sample size estimation procedures discussed in the previous section can be used.

Model 3: Using a balanced (integrated) mixed method design

While in most cases the mixed method designs are used to complement a predominantly QUANT or QUAL design, there are cases where an integrated mixed method design that gives equal weight to both approaches might be used. These designs can involve the combination of quantitative and qualitative techniques, as well as specific mixed method techniques, at different stages of the evaluation. Estimating the required sample sizes must combine QUANT procedures for sample surveys and the rules of thumb discussed in the previous section for the QUAL components. Combining the two requires judgment based on experience. Example 3 in Part V describes a balanced (integrated) MM design used to evaluate a large-scale program in India to strengthen the capacity of communities to manage development grants from the state governments. Sample surveys, used to obtain quantitative estimates of program outcomes, were combined with qualitative data collection through regular visits to a sample of project and control communities to observe the processes of change and to interview a small sample of households as well as local officials and key informants. The same sampling frame was used for both QUANT and QUAL samples, and the QUAL samples were selected to ensure their representativeness.

3.2 Using mixed methods to evaluate complex interventions

Complex development interventions⁹ are usually defined as interventions that have some of the following characteristics:

- Country-led planning and evaluation;
- The program evolves in a nonlinear manner;
- Many components or services and multiple objectives;
- There are both quantitative and qualitative outcomes;
- Target populations are difficult to identify or monitor;
- General budget support is provided with no clear definition of the services to be funded
- Multiple donors and agencies;
- Participatory planning and implementation processes are used so that outcomes and impacts are difficult to define in advance and may constantly change;
- The context is complex;
- The situation is constantly changing and the intervention tries to adapt to these changes;
- There are complex processes of behavioral change involving the target population, service providers and other agencies, e.g., the police, military and judiciary (Section 3.3);
- While many complex interventions are large, some are relatively small but may involve complex processes of behavioral change or populations that are difficult to identify or study.

Mixed method designs are well suited to assess complex interventions, as it is possible to combine standardized design, data collection and analysis with tools that can capture the

⁹ For a discussion of complex interventions and their evaluation see Funnell and Rogers (2011), Patton (2011) and Bamberger et al, (2012) Chapter 16.

BOX 3. ATTRIBUTION, CONTRIBUTION AND SUBSTITUTION ANALYSIS

While project evaluation is sometimes able to apply experimental or quasi-experimental designs that can control for alternative explanations of the observed changes (outcomes or impacts) in the project group, it is rarely possible to achieve the same level of rigor in the evaluation of the effects of complex interventions. Consequently, evaluators of the effects of complex interventions must decide which of the following levels of analysis can be applied:

Attribution analysis: The project group is matched to a comparison group so that alternative explanations of the observed changes can be controlled for (eliminated). If there are statistically significant differences between the project and comparison groups, this is taken as evidence that the project intervention contributed to these changes.

Source: Adapted from Bamberger et al 2012 Box 16.3 pp 403–4.

Contribution analysis: The analysis assesses the contribution of a particular development agency to the achievement of the overall changes resulting from the collaborative financial and technical interventions of a number of different development agencies. Sometimes the analysis will also include the programs of national agencies (both government and civil society).

Substitution analysis: An assessment is made of the *net increase in resources* to the target sector or program resulting from the contribution of a particular development agency. This takes into account any diversion of resources to other activities by the national government (reducing the net increase in resources), or any multiplier effect (increasing the increase in resources).

complexities of the program setting, the changing nature of the program and its intended outcomes and the processes of behavioral change that are taking place. It is also possible to combine the perspectives of different stakeholders with “objective” quantitative indicators, and there are a range of QUAL tools for collecting sensitive data from difficult to reach populations. At the same time, MM provide ways to estimate quantitative impacts in the most credible way possible in difficult circumstances and to use triangulation to increase the construct validity of estimated outcomes and of the interpretation of what is actually happening on the ground.

Figure 6 summarizes the wide range of tools and techniques on which a MM design can draw in selecting the appropriate design for assessing a complex intervention (all of the techniques listed in this

figure are described in [Annex 7](#). A central element of the mixed method approach is to broaden the ways in which the counterfactual¹⁰ can be defined and estimated. This is done in two steps.

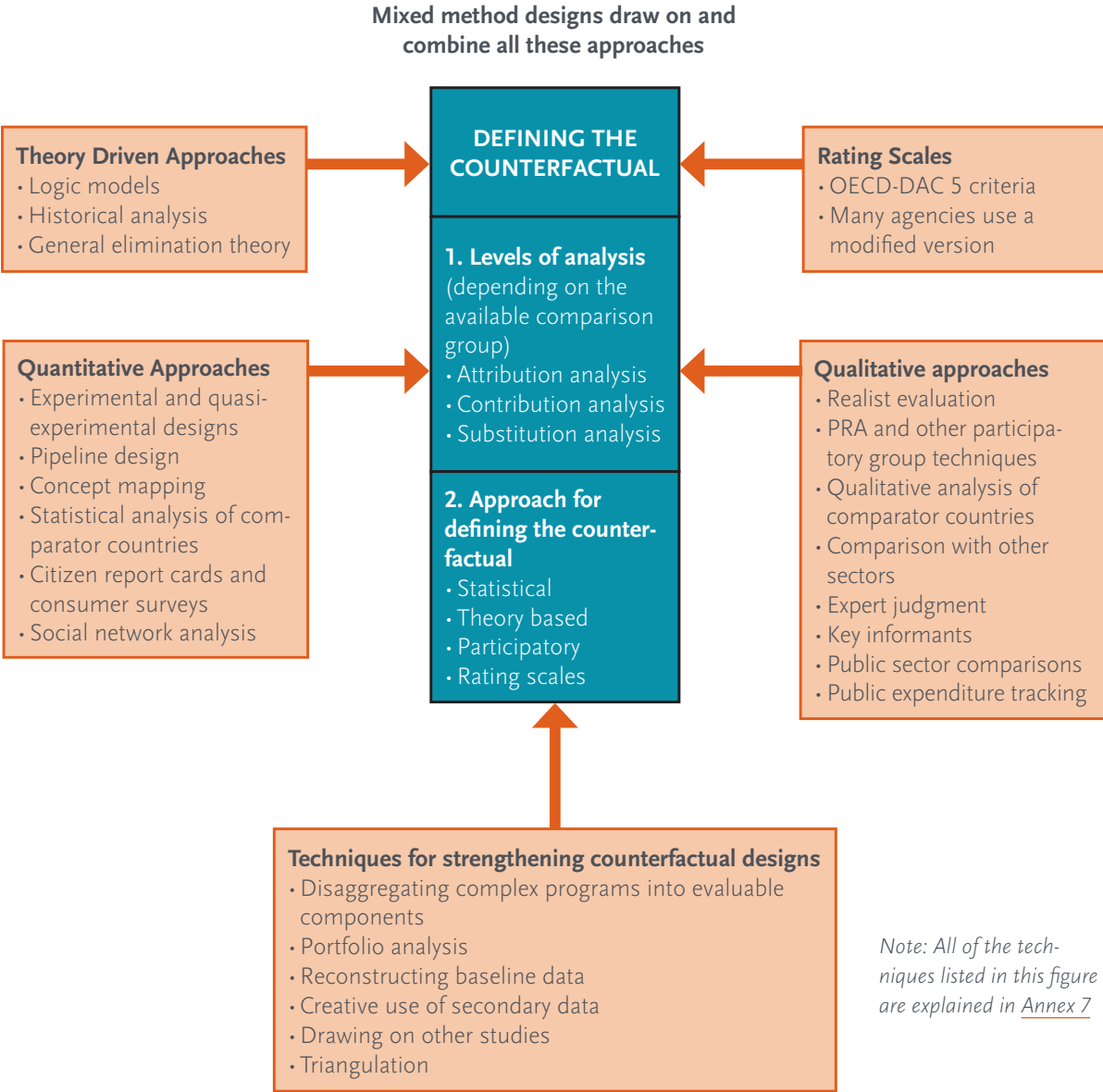
Step 1: Depending on the type of comparison group that is available the evaluation will conduct the analysis at one of the following levels¹¹ (see Box 3):

- Attribution analysis
- Contribution analysis
- Substitution analysis

¹⁰ GN1 (p.1) defines the counterfactual as “an estimate of what would have happened if the intervention had not occurred for example, comparisons with a group who did not receive the intervention.”

¹¹ See GN1 Section 9 for a discussion of a wide range of approaches for assessing attribution and contribution. Bamberger et al (2012) pp 403–405 provides a more extensive discussion including substitution analysis.

Figure 6. Using mixed method designs for evaluating complex interventions



Step 2 then selects the approach, or the combination of approaches that will be used to define and estimate the counterfactual:

- *Statistical comparison group*: using a statistically matched control or comparison group
- *Theory-based*: the implementation process and outcomes defined in the Theory of Change (TOC) are compared with what is observed on the ground. The more closely reality corresponds to the TOC the stronger the case for assuming the intervention contributed to the outcomes. However, it is also important to define and test alternative explanations of the observed changes (rival hypotheses).¹²
- *Participatory-based*: program effects are estimated through assessments by intended beneficiaries and other stakeholders. One example of this approach is Most Significant Change, and another is the use of participatory rural appraisal (PRA) and other group analysis techniques.
- *Rating scales*: experts or stakeholders are asked to rate program performance or the degree of change produced by the intervention on a set of rating scales. The OECD-DAC scales for assessing relevance, efficiency, effectiveness, impact and sustainability are one example of a commonly used rating scale.

Mixed method designs can combine, as appropriate: theory-driven approaches, qualitative and statistical approaches and rating scales. [Annex 7](#) lists the range of options available for each of these

¹² A weakness of many TOCs is that they do not identify alternative explanations of the observed outcomes. If changes are consistent with the TOC this is taken as evidence that the changes can be attributed to the effects of the project—which is not a valid conclusion. A TOC should identify alternative explanations (rival hypotheses) and build into the model ways to test these rival explanations of the causes of the observed changes.

approaches. A number of techniques can then be used to strengthen the counterfactual design (“unpacking” complex interventions, portfolio analysis, reconstructing baseline data, creative use of secondary data and drawing on other studies). The construct validity of the estimates from all of these sources is then strengthened using triangulation.

There is no single best approach to mixed method evaluation, as evaluators must select the set of tools and techniques that are best suited to the budget, time, data and political constraints and the purposes of each evaluation. The case studies listed in Part V illustrate the range of mixed method approaches that have been used for different evaluations. Creativity, familiarity with a wide range of QUANT, QUAL and theory-based approaches, and willingness to draw on different disciplines are essential requirements for mixed method evaluations.

3.3 Assessing processes of behavioral change

There are many projects where the implementation process is much less clear-cut and linear than it first appears. This means that impact evaluation must take into account these processes of behavioral change, as they often result in programs having a number of unintended outcomes and impacts. The following are some of the reasons for this:

- In most cases intended beneficiaries actively decide whether or not to participate in the program, often choosing which services they will use and which not.
- Programs are often modified based on how the initial beneficiaries perceive and respond to the services and the feedback they provide to others. Many program designs are modified as a result of these interactions

BOX 4. PROGRAM OUTCOMES ARE AFFECTED BY HOW BENEFICIARIES RESPOND TO SERVICES AND HOW STAFF REACT TO BENEFICIARY RESPONSES. STUDYING BEHAVIORAL CHANGE IN A SCHOOL FEEDING PROGRAM IN NICARAGUA.

A school breakfast program was organized in a food insecure rural area of Nicaragua to increase school enrolment. In some schools teachers also gave food to younger siblings who came when mothers were bringing students to school. The word was passed to other mothers and as the number of siblings increased some teachers gave all siblings breakfast and others did not. The decision largely depended not only on the attitude of each teacher but also on the willingness of mothers to help prepare the breakfasts and local farmers to donate more food. In some communities these behavioral responses transformed the

school breakfast program into a major nutrition program involving mobilization of food from local farmers, men helping transport large quantities of food, building school kitchens and in some cases installing electric power in the classrooms, while mothers organized the breakfasts and sometimes started school vegetable gardens. The final outcomes in each school and the lessons learned could only be determined by close observation of the behavior of teachers, parents and the communities and seemed to vary depending on individual attitudes, local economic conditions and sometimes ethnicity.

within the target population and between them and service providers (see Box 4).

- Project staff all have the own attitudes toward different groups, beliefs about who will benefit and who will not and their own ways of interacting with the community.¹³ Even when staff receive rigorous training, each person will react in a slightly different way.
- Each project operates in a unique setting where implementation and outcomes are affected by a unique set of social, economic, political, environmental, legal and other factors.¹⁴

13 A review of factors affecting the success of Brazilian state education systems in increasing access to low-income families found that many teachers believed that children from poor households were very likely to drop-out of school and even if they stayed they would get poor grades and would be disruptive. Consequently, many teachers did not feel it was worth making the effort to help or encourage children from poor backgrounds. Bamberger and Segone (2011) argue that many causes of inequality are socially determined as governments and public opinion may not wish to encourage access of different immigrant groups or ethnic minorities to public services.

14 Systems analysis provides useful frameworks for understanding the operation of these contextual actors and also for understanding how history affects the attitudes and expectations of stakeholders and communities to new project interventions. Space does not permit a discussion of systems analysis. For a brief introduction to systems analysis and how it can be used in impact evaluation see Bamberger (2011) How to design and manage equity focused evaluations Section 5.2.C.

The following are some of the ways that mixed method designs can evaluate these programs:

- A rapid diagnostic study conducted at the start of the evaluation can help clarify the setting within which the program is implemented and the contextual factors that are likely to affect implementation. The study can also describe social stratification and identify the marginal and vulnerable groups that might be excluded from access to project benefits. Initial QUAL analysis can be combined with a rapid QUANT survey to estimate the magnitude and distribution of, for example, vulnerable groups.
- The scope of the conventional program monitoring can be broadened to provide more detailed QUANT information on the characteristics of the social groups who do and do not use project services.
- The bottleneck analysis framework developed by UNICEF can provide a more rigorous framework for the analysis of the factors determining which sectors do and do not have access to the project and the supply and demand-side factors determining access.¹⁵

15 See Bamberger and Segone (2011) How to design and manage equity-focused evaluations. Pp. 45–50. UNICEF. Available at http://mymande.org/?q=equity_focused_evaluations_intro.

- A number of QUAL techniques are available to observe what actually happens during project implementation, how different groups respond to the project, and how the original design and implementation plan is modified by the interaction between different sectors of the target population

and project staff. These techniques include, but are not limited to: participant observation; panel studies where a small sample of individuals, households or communities are visited periodically throughout the project; focus groups; Participatory Rural Appraisal (PRA) techniques; and self-reporting.¹⁶

16 Examples of self-reporting techniques include asking a small number of respondents to keep diaries in which they record experiences and activities relevant to the project or asking households to keep a record of their income and expenditure. In some studies respondents are given cell phones and asked to call in to report, for example, their travel patterns, locations in which they feel insecure or tempted to purchase drugs etc.

Part IV. Managing Mixed Methods Evaluations

4.1 Mixed methods designs require a special management approach

Although some applications of mixed-method designs involve only adding additional data collection methods to a dominant QUANT or QUAL design, a fully integrated MM evaluation strategy involves much more than this. To enjoy the full benefit of combining QUANT and QUAL approaches and methods, it is necessary to plan an MM evaluation strategy from day one. This requires a management commitment to guarantee the additional time and resources required to effectively implement the approach—and to defend these from administrative pressures to cut budgets and time. This will often require an energetic campaign by the evaluation manager to ensure buy-in from the senior management and funding agencies. Briefings in management meetings, short reports or trainings on methods, newsletters, short videos, etc., can be useful advocacy tools.

The evaluation team must also be able to produce convincing evidence to show that MM do bring additional benefits. For example, a small number of typical evaluations might be selected with additional time and resources approved to test the MM approach and to compare outcomes with standard evaluation approaches used on similar programs. Is it possible to demonstrate that the extra resources and time required for a MM approach do add value?

In addition to ensuring sufficient time and resources, management of a MM evaluation requires special attention to the following areas (see Box 5):

Composition of the Research Team. Ideally, the research team should include principal researchers from two or more disciplines (e.g., anthropology, medicine, law, sociology, economics). However, for most evaluations, resource constraints will not permit this and ways must be found to ensure that the evaluation team members are able to cover all of the basic QUANT and QUAL tools and techniques. A challenge for the evaluation manager is to help build these basic skills, often with some support from consultants or other agencies. Where team members have different skill sets and different professional orientations, it is important to allow time and opportunities for each researcher to become familiar with the methodology of the others and to develop mutual respect and trust among the members of the team. This is even more important when the evaluation team comprises specialists from different countries as well as from different disciplines. Table 5 indicates some of the additional areas of research expertise that may be required by teams with a QUANT or QUAL orientation.

Managing mixed method approaches during the evaluation design. Management support may be required to ensure that the evaluation framework draws on all the involved disciplines and that the research questions and issues incorporate each of these methodologies. This can be important

BOX 5. PLANNING AND BUDGETING ADDITIONAL TIME AND RESOURCES THAT MAY BE REQUIRED FOR A MIXED METHOD EVALUATION

MM evaluations often require additional time (e.g., for team building) and money. The manager must decide if the extra resources are justified and, if so, ensure they are budgeted.

Composition and integration of the research team: *The evaluation manager has a critical role to play in ensuring that team members from different disciplines work effectively together. This requires extra time and effort.*

- Allow time for researchers to develop an understanding and respect for each other's disciplines and work. Ensure all team members are familiar with the basic literature and current debates in the other field.
- Ensure similar linkages and team building for local researchers.

Integrated approaches during the evaluation design: *The evaluation manager must actively encourage all team members to take full advantage of new frameworks, data collection and analysis tools.*

- Ensure that researchers from different disciplines are brought in at the evaluation design stage or approximately the same time, so that everyone feels they are making a significant contribution to the overall evaluation, and are not just the "icing on the cake." This is especially important where one discipline is dominant.
- Ensure that the evaluation draws on theories and approaches from all the disciplines involved in the evaluation (e.g., anthropology, medicine, law, sociology, economics, demography), with each being used to enrich and broaden the others.
- Ensure that concepts and methods are not taken out of context, but draw on the intellectual debates and approaches within the respective disciplines.

Data collection and the use of triangulation: *Many evaluation proposals refer to triangulation, but it is often not used systematically. Triangulation is a key component of a mixed method approach and it is the responsibility of the manager to ensure that it is fully used.*

- Select QUANT and QUAL data collection methods that complement each other, and specify how they will be combined in the fieldwork and analysis.
- Select at least two independent estimating methods for key indicators and hypotheses.
- Ensure full documentation of all sample selection, data collection, and analysis methods.

Data analysis and possible field follow-up: *The manager should ensure that there is an MM data analysis plan put in place early in the evaluation.*

- Present separate analyses of QUANT and QUAL findings to highlight different interpretations and findings and prepare an integrated report drawing on all of the data.
- Use systematic triangulation procedures to check on inconsistencies or differing interpretations.
- Budget resources and time for follow-up visits to the field.
- Highlight different interpretations and findings from different methods and discuss how these enrich the study. Different, and seemingly contradictory, outcomes should be considered a major strength of the integrated approach rather than an annoyance.
- Present cases and QUAL material to illustrate or test QUANT findings.

Presentation and dissemination of findings: *The manager should encourage the team to broaden the range of presentation and dissemination methods to ensure that the full richness of MM data is captured.*

- Combine conventional written reports and PowerPoint presentations with more participatory presentation methods. Develop more innovative and user-friendly reports and avoid long, technical reports for nontechnical audiences.
- Broaden the range of stakeholders invited to presentations of findings to include community and civil society groups often not be consulted in many QUANT evaluations.

Table 5. Additional areas of research expertise that may be required for QUANT and QUAL oriented teams to conduct MM evaluations

QUANT Oriented Evaluations	QUAL Oriented Evaluations
<ul style="list-style-type: none"> • Knowledge of basic QUAL data collection methods (e.g., in-depth interviews, group interview techniques, observation techniques) • Knowledge of the local culture • Experience with, and evaluation of, community and organizational development programs • Systematic use of triangulation 	<ul style="list-style-type: none"> • Statistical sampling methods • Management of QUANT data collection, particularly the administration of structured questionnaires • Statistical data analysis

for organizations that have traditionally relied on mainly QUANT or QUAL methods and where the evaluation team may have to be encouraged to ensure that equal weight is given to the frameworks and methods of the newly introduced disciplines.¹⁷

Ensuring the use of MM during data collection. The decision whether to use sequential or concurrent data collection is partly a technical issue, but it involves management considerations. Consecutive QUANT and QUAL data collection may save significant amounts of time, which may also save money, but this will require efficient management systems to coordinate QUANT survey and QUAL data collection teams who may have different methods of operation. When data collection is constrained by security and logistical challenges, the use of consecutive data collection may require efficient management of transport to deliver researchers to and collect them from different areas, as well as coordination with police and military in situations where permission has to be obtained to

visit communities and precise itineraries may have to be prepared and approved.

Managing mixed method data analysis. With a dominant QUANT design, data analysis normally does not begin until all, or most, of the data has been collected, entered into a database, and cleaned. However, the analysis of QUAL data may be a continuous process that begins soon after the researchers enter the field. Data management becomes more challenging because it is an interactive rather than a linear process. Initial findings from QUAL data analysis are often used to design questions for the QUANT surveys, and initial QUANT data analysis is often used to select the samples for the in-depth QUAL analysis. This presents a number of special management challenges: (a) data collection must be conducted more rapidly, so as not to delay the start of the next stage of the evaluation design and data collection; (b) data quality procedures are more complex, as they must assess the quality of both QUANT and QUAL data collection; and (c) triangulation procedures must be used systematically to check the consistency of estimates of key outcome indicators obtained from different methods of data collection, and procedures must be in place to follow up and determine the reasons

¹⁷ It will often be the case that professionals from the new disciplines will not be brought in until the evaluation design has already been defined and where they will only be asked to collect data to fit into the already established evaluation designs. Experience shows that QUANT oriented evaluations will often only commission QUAL oriented researchers to conduct a few focus groups or case studies to show the evaluation is using mixed methods, but without really building these into the overall evaluation design.

for inconsistencies between data from different sources. All of these require more complex and efficient data management systems.

4.2 Tips for resource constrained NGOs to mobilize the expertise and resources required to conduct mixed methods evaluations

Mixed methods evaluations offer a great advantage for NGOs that require quality and credible evaluation reports but that must conduct the evaluations “on a shoestring,” with only limited in-house evaluation expertise, and/or few resources to bring in outside experts. For the many NGOs conducting evaluations under these real-world constraints, creativity will frequently be required to obtain the required expertise. This can be done either through additional training for the current research team or by finding ways to obtain the required expertise through collaboration with other agencies. There are no hard and fast rules as to the minimum levels of expertise, time and resources requirements to conduct a methodologically sound MM evaluation, but a minimum level of expertise should ideally be available in all of the areas identified in Table 5. Box 6 offers tips on how to achieve the essential expertise and to mobilize additional resources while working on a limited budget.

All of these constraints affect the credibility of evaluation findings. How credible is an evaluation that had to be conducted on a tight budget, with little time in the field, limited access to data, possibly with security concerns? While there is no “magic bullet” for resolving these challenges, there are a number of ways in which MM can help achieve acceptable methodological rigor and credible findings. The following are some of the useful

ways to achieve credible findings while working within these constraints:

a. Basing the evaluation on a well-articulated theory of change

(Both GN1 and GN2 include extensive discussions of the development and use of the TOC in the impact evaluation design and analysis). A TOC that is developed through participatory consultations with stakeholders can define the steps and processes through which outputs, outcomes and impacts are to be achieved and can identify the critical assumptions to be assessed. The TOC should include agreed-to milestones, so that if an evaluation must be conducted when it is still too early to measure outcomes, the milestones can help increase the credibility of the evidence that the project is on track to achieve its objectives. MM can strengthen the TOC by incorporating both QUANT and QUAL indicators, studying what happens during project implementation and describing important processes of behavioral change.

b. Consultation with stakeholders to ensure that they find the evaluation methods and the key indicators credible, and to be aware of (and to address) the concerns that they may have.

c. Using triangulation to maximize the validity of estimates based on small samples from different sources (see Section 2.2).

d. Using mixed methods sampling to ensure the representativeness of QUAL data from case studies, in-depth interviews and observation (see Section 3.1 Model 2).

BOX 6. CONDUCTING MIXED METHODS IMPACT EVALUATIONS ON A SHOESTRING: TIPS FOR NGOS WITH LIMITED RESOURCES AND EVALUATION EXPERTISE

While mixed methods can be used as part of a rigorous, expensive and complex impact evaluation, most NGOs have limited time and resources to invest in impact evaluations and only limited evaluation expertise. Yet they need valid and credible assessments of the outcomes and impacts of their programs. The good news is that MM tools and techniques are very flexible and can help improve the quality of impact evaluations for NGOs conducting evaluations on a shoestring. The following are some practical tips discussed in this guidance note.

Start gradually and broaden the scope as experience is gained.

- It is usually simpler and cheaper to start by using MM in only a single stage of the evaluation—for example, using at least two independent methods to collect data, or conducting a rapid diagnostic study to help design a questionnaire. Take advantage of data that is already being generated through the program monitoring system. As your experience increases, then consider broadening the scope.
- It is usually easier to start by using *sequential* rather than *consecutive* impact evaluation designs (see Section 2.2).
- While some mixed methods approaches are quite complicated to use, there are many other techniques that are simple and economical and can be easily learned and used by staff with only limited research experience.
- Integrate MM design with the M&E system and make maximum use of data collected through these (see GN2).

Ensure the findings are both methodologically sound and credible to stakeholders.

- To start, focus on the kinds of data that are considered credible by stakeholders. Don't spend all of your resources on large sample surveys if your clients and stakeholders are more concerned with case studies and in-depth description of how the program actually operates.
- At the same time, try to balance stakeholder preferences with a judicious combination of QUANT and QUAL data and with the use of triangulation and other techniques to enhance validity. MM studies collect a wide range of different kinds of

quantitative and qualitative data, so that even from a small evaluation it is possible to select the kinds of evidence that are most convincing to stakeholders.

- MM sampling can ensure that a small number of case studies can be selected to ensure they are broadly representative of the total target population—thus increasing the validity/credibility of the findings (see Section 2.3).
- By creative combining of evidence from different sources, *triangulation* can increase the credibility of estimates from small samples (see Section 2.2 and Annex 9).

Stay within your budget and level of expertise.

- There are many ways to reduce the costs of data collection while still ensuring the collection of reliable data (see Annex 8).
- Many university departments teach mixed methods, and it may be possible to collaborate with one of the faculty or students for free or for token costs for one of their field assignments. Many post-graduate students might consider preparing a case study as one of their course assignments.

Get help.

- There are a lot of free webinars and other internet resources on the design and use of MM (<http://www.interaction.org/resources/training> is a good place to start).
- Many of the large international NGOs have permanent evaluation departments that may be willing to provide some free guidance on the phone or in person if they happen to be visiting your country.

Real-world examples of simple and economical mixed method designs.

- Case study No. 7, Evaluating the UNICEF Education Project in Timor L'Este, and Case study No. 11, Evaluating the Eritrea Community Development Fund, both illustrate how mixed methods can be used where data is difficult to access and resources are limited

Part V. Case Studies Illustrating Different Applications of Mixed Methods Designs

The three evaluations described in this section illustrate the methodologies used in: a predominantly QUANT evaluation, a predominantly QUAL evaluation, and an evaluation giving equal weight to both QUANT and QUAL methods. Together these three examples illustrate the wide range of MM evaluation approaches that can be used. Table 6 (at the end of this section) lists 17 examples of MM evaluation designs that are summarized in [Annex 10](#). The evaluations covered projects supported by NGOs, UN agencies and the World Bank.

Example 1: A QUANT-oriented evaluation: Evaluating a post-conflict reconstruction program in Liberia (DFID and the International Rescue Committee)

The evaluation, completed in December 2008, assessed the impact of a DFID-funded program of community driven reconstruction (CDR) implemented by the International Rescue Committee in post-conflict Northern Liberia in 2006–7. In 42 communities, the CDR program used block grants, elected decision-making institutions, participatory planning and community development projects to improve socio-economic welfare, local governance and community cohesion. The evaluation assessed whether the program achieved its objective.

A mixed methods evaluation design was used. The motivation for using a mixed methods approach was the recognition of the difficulties of assessing,

purely on the basis of QUANT survey data, whether changes in stated attitudes reflected real changes. It was expected that exposure to the project might teach participants what were the expected ways to respond to questions about attitudes to cooperation and democracy, without affecting their actual ability or propensity to engage in community collective action. The evaluation was designed to address these challenges by combining QUANT survey data on changes in attitudes and reported behavior with field experimental data that could measure actual behavior changes. Both the survey and the field experimental methods were based on a randomized control trial—at the outset, communities were randomly assigned to receive CRC treatments (developing community committees and providing support for rapid impact programs), while a roughly equal number of control communities did not receive CRC programs. Other QUANT methods used included a social capital questionnaire to assess both how outcomes were affected by existing social capital and how the program affected social capital, and the collection of survey data on community organization and social cohesion.

Field experimental and QUAL methods included the collection of data on community organization and social cohesion through in-depth interviews and observation with individual behavior in a public goods game. Six months after the CDR program was completed, all treatment and control

communities were given the opportunity to raise funds from a Liberian NGO to implement a public goods project. The only requirements were that the communities choose a project in advance and identify three individuals to handle the money. Communities were also told that the amount of money they would receive (ranging from \$0 to about \$500) would depend on the results of a village-wide public goods game, in which 24 randomly selected individuals could choose privately to retain a sum of money for their own use or contribute it to a community fund (with an additional matching contribution supplied by the researchers) to be used for the public good. The researchers then gathered data on how communities selected projects and representatives and observed patterns of play in the public goods game. It was believed that the games could provide a more accurate estimate of attitudes and behavior as there was a real monetary cost to acting to please outsiders.

**Example 2: A QUAL oriented MM evaluation:
Evaluating a program to prevent violence against
women in El Salvador (Oxfam America)**

The 10 year program to combat violence against women (VAW) was launched in 2008. It operates at the national, municipal and community levels, combining public awareness campaigns with the organization of women and civil society, and promotes more effective coordination between public authorities and organized women. A major challenge concerns the attribution of changes in attitudes and behavior with respect to national and municipal advocacy and campaigning strategies, particularly among public sector agencies. The question of counterfactual is difficult. Another challenge was the limited availability of good QUANT data on levels and changes in VAW. Police, justice and health data were both unreliable and underreported. As always, victims were reluctant to report abuse. Consequently, the

evaluation team decided the most accurate way to measure influence of the campaign was to conduct in depth comparison cases, focused on intermediate objectives, supported wherever possible by available QUANT data.

The MM evaluation design was based on a theory of change. This identified how multiple outcomes promoted in coalition might sequence and combine to produce intended impacts: coordination and exchanges across regional agencies; new policy and norms (legislation, school curricula, municipal prevention policies); changes in social relations and behavior through awareness raising and better application of the law; and wellbeing and agency (increased knowledge; women more confident to take action; women benefit from improved services; women, youth and allies influence decision makers). A set of 11 long-term indicators were defined to assess outcomes and impacts over the 10 year period, and short-term progress “benchmark targets” were defined every three years.

The evaluation was based on a set of process principles (multidisciplinary team and mixed methods approach, regionally based research institution, collaborative design and consultation with all partners, and the use of triangulation and consultation to validate findings). The key elements of the methodological approach included:

- In-depth comparative case studies in two municipalities.
- Major cross-stakeholder document synthesis.
- Interviews on the effectiveness of advocacy choices at the national level.
- Focus on linkages between different levels of intervention (national, municipal and local)
- Different time horizons for assessing different outcomes.

The research pathways and key data sources include:

- Providing context on general trends through collection of national- and municipal-level indicators (government statistics, statistics on violence against women, passage and implementation of laws on VAW, budget commitments).
- Two in-depth comparative case studies focusing on results in particular municipalities, women's perceptions of safety, assessing effectiveness of prevention actions, women's confidence to address VAW. Data sources include in-depth interviews, observation, municipal statistics and trends, and focus groups with women providing accounts of changing conditions.
- Advocacy evaluation: Trace the influence pathway of the campaign through document review and interviews with policy-makers and other actors on key events and decision points, and to ask if progress could have been made without the advocacy campaigns. Assess whether the campaign influenced the final version of legislation, and the effectiveness of civil society mechanisms to influence prevention policies. Use secondary data to compare trends in areas where campaigns did and did not operate.

The analysis and interpretation of findings combined:

- Contribution analysis at the broader program level using project component data, case studies and tracing influence pathways (based on the TOC framework). Causality and attribution can be studied for component projects of the program coalition, but the focus at the program level is to test basic assumptions of a highly complex body of work.

- The TOC was a key element in defining a counterfactual providing a framework for comparing observed changes and processes with the intended process of change.

Example 3: A balanced MM design: Evaluating the impacts of the India Gram Panchayat Community Development Reform Program (World Bank)

The purpose of the program was to devolve power to democratically elected village governments and to promote greater participation of women and scheduled castes. It was also intended to increase the effectiveness of the use of funds through greater community participation. The government provided grants that would be managed by the local community for implementation of economic and social infrastructure projects to be selected by the community. The program being assessed was a two-week training program for citizens on participatory planning and decision-making designed to improve the efficiency with which they would use the grants. The program design involved random assignment of communities to the treatment and control groups (receiving grants but no training).

The evaluation used a balanced mixed methods design involving the following steps:

- Selection of 200 villages (Gram Panchayats) with random assignment to project and control groups (QUANT).
- Exploratory research on land tenure, ownership of public goods, participation and social networks (QUAL).
- Baseline survey prior to the training programs (QUANT).
- In-depth process analysis in five project and five control areas (QUAL) to observe the changes in how communities organize, plan and manage projects, and the effects on

participation and the involvement of women and scheduled castes. This involved one- to two-day visits every week over the two year period. Visits combined observation at the community level, interviews with politicians and community leaders, and in-depth interviews with 20 households in each village. Observation covered: political and social dynamics, corruption, economic change and network affiliation.

- Repeating the baseline survey after two years (QUANT).
- Triangulation was used throughout to compare QUANT and QUAL estimates of change and impacts.

Some of the advantages of the MM approach included:

- Recall of processes and discussions in village meetings are unreliable, so observers attended meetings and transcribed the proceedings and the meeting dynamics.
- Previous studies have found that surveys

produce widely varying estimates of inequality (the variance is too large to be able to use the findings for statistical analysis), so PRA techniques, where participants were selected to be representative of the total population, were used to complement survey data.

- Community development, participation, corruption and other key processes are difficult to capture in surveys, so survey data was compared with observation, key informants and participatory group interviews.
- At the same time it was essential to obtain quantitative estimates of the cost-effectiveness of the training programs, as this was a pilot program to inform decisions on whether the program should be replicated in other states.¹⁸ Consequently, sampling strategies ensured that qualitative case studies and other data could be linked to the QUANT sample survey to ensure that the evaluation findings could be generalized.

¹⁸ The evaluation included a comparison with four other states (not included in the present summary).

Table 6. Examples of MM evaluations summarized in [Annex 10](#)

Title	Agencies supporting the program and conducting or commissioning the evaluation
A. Quantitatively oriented evaluation designs	
1. Post-conflict reconstruction in Liberia.	DFID and the International Rescue Committee
2. Long-term evaluation of the Tostan program to reduce female circumcision in villages in Senegal.	UNICEF
3. Evaluating a conditional cash transfer program in Kazakstan.	Save the Children
4. Impact evaluation of FAO emergency and rehabilitation work in rural DRC.	FAO
B. Qualitatively oriented evaluation designs	
5. Evaluating the impacts of a gender-based violence prevention program in El Salvador.	Oxfam America
6. Life and livelihoods food security program in Bangladesh.	USAID, Save the Children and TANGO International
7. Evaluation of the UNICEF Education Programme in Timor-L'Este.	UNICEF
8. Evaluating the equity-outcomes of the Nepal Education for All Project.	UNICEF, NORAD and other partners
9. Evaluating the equity outcomes of the Cambodia Community-Led Total Sanitation Project.	UNICEF
10. Inter-Agency Real-Time Evaluation of the Humanitarian Response to Pakistan's 2009 Displacement Crisis.	UNICEF and various partners
11. Evaluating the road construction component of the Eritrea Community Development Fund.	World Bank
12. Evaluation of the Egyptian Community Schools Project.	UNICEF
13. Evaluation of the Tanzania Community Justice Facilitation Project.	UNICEF
14. Evaluating UNICEF's Response in the area of Child Protection in Indonesia, to the 2004 Indian Ocean Tsunami.	UNICEF
C. Balanced evaluation design giving equal weight to QUANT and QUAL approaches	
15. Evaluating the Kecamatan Development Project in Indonesia.	World Bank and Government of Indonesia
16. Evaluating the Indian Panchayat Reform Program.	World Bank
D. Meta-analysis (secondary analysis of a number of evaluations to identify general findings)	
17. CARE International. Impact Evaluation Report 2005–2010. Latin America and the Caribbean.	CARE International

References

- Adato, M (2012) “Combining quantitative and qualitative methods for program monitoring and evaluation: Why are mixed methods designs best?” in Lopez-Acevedo, G; Krause, P & Mackay, K (editors) Building better policies: The nuts and bolts of monitoring and evaluation systems. World Bank.
- Bamberger, M. (2010). *Reconstructing baseline data for impact evaluation and results measurement. No. 4. The Nuts and Bolts of M&E Series. Poverty Reduction and Equity Department. The World Bank. Available at <http://siteresources.worldbank.org/INTPOVERTY/Resources/335642-1276521901256/pre-noteME4.pdf>.*
- Bamberger, M., Rao, V., & Woolcock, M. (2010). *Using mixed-methods in monitoring and evaluation: Experiences from international development.* In Tashakkori A. and Teddlie C. Sage Handbook of Mixed- Methods in Social and Behavioral Research. Sage Publications. pp 613–642.
- Bamberger, M., Rugh, J. & Mabry, L. (2012). *RealWorld Evaluation Second Edition.* Sage Publications.
- Chen, H. (2005) *Practical Program Evaluation: Assessing and improving planning, implementation, and effectiveness.* Sage Publications.
- Donaldson, S., Christie, C. & Mark, M. (2009). *What counts as credible evidence in applied research and evaluation practice?* Sage Publications.
- Funnell, S. & Rogers, P. (2011). *Purposive programme theory: effective use of theories of change and logic models.* Jossey-Bass Publications.
- Mayne, J. (2008) “Contribution analysis: An approach to exploring cause and effect.” Rome: Institutional learning and change initiative , ILAC Brief No. 16. May 2008. http://www.cgiar-ilac.org/files/publications/briefs/ILAC_Brief16_Contribution_Analysis.pdf
- Organization for Economic Cooperation and Development. 2010 (Second Edition). *Evaluating Development Cooperation: Summary of Key Norms and Standards.* Paris: OECD. <http://www.oecd.org/dataoecd/12/56/41612905.pdf>
- Patton, M.P. (2011). *Developmental Evaluation: applying complexity concepts to enhance innovation and use.* Guilford Press.
- Proudlock, K & Ramalingam, B (2009) *Improving humanitarian impact assessment: Bridging theory and practice.* ALNAP
- Rieper, O Leeuw, F & Ling, T (2010). *The evidence book: concepts, generation and use of evidence.* Transaction Publishers.
- Salmen, L. (1987). *Listen to the People: Evaluation of Development Projects.* New York. Oxford University Press.
- Stern, E; Stame, N; Mayne, J; Forss, K; Davis, R & Befani, B (2012) *Broadening the range of designs and methods for impact evaluation.* DFID

UNICEF (2009a). Children and the 2004 Indian Ocean Tsunami: Evaluation of UNICEF's response in Indonesia, Sri Lanka and Maldives, 2005–08. Overall synthesis report.

Weiss 2001. Theory based evaluation: Theories of change for poverty reduction programs pp. 103–14 in *Evaluation and Poverty Reduction* edited by O. Feinstein and R. Piccitto. New Brunswick. Transaction

Williams B. (2005). Systems and Systems Thinking in Mathison, S (editor) *Sage Encyclopedia of Evaluation* pp. 405–412.

Williams, B. & Imam I. (Eds.), (2007). *Systems Concepts in Evaluation: An expert anthology*. American Evaluation Association.