



**Assessing the Impact of  
Microenterprise Services (AIMS)**

Management Systems International  
600 Water Street, S.W.  
Washington, D.C. 20024-2488  
Tel: (202) 484-7170 • Fax: (202) 488-0754  
E-mail: aims@msi-inc.com

---

**GUIDELINES FOR MICROFINANCE IMPACT ASSESSMENTS**

**Carolyn Barnes and Jennefer Sebstad**

**Management Systems International**

**Discussion Paper for the CGAP 3 Virtual Meeting  
October 18-29, 1999**

**March 2000**

---

This work was funded by the Microenterprise Impact Project (PCE-0406-C-00-5036-00) of USAID's Office of Microenterprise Development. The Project is conducted through a contract with Management Systems International, in cooperation with the Harvard Institute for International Development, the University of Missouri, and the Small Enterprise Education and Promotion Network.



## ACKNOWLEDGMENTS

The Consultative Group to Assist the Poorest (CGAP) Impact Assessment Methodologies Working Group, chaired by Monique Cohen of USAID, has supported a series of activities over the past three years related to the development of guidelines for conducting microfinance impact assessments. Activities have included background papers, field studies, discussion papers, and virtual meetings. The discussion papers and virtual meetings have been managed by Management Systems International under USAID's Assessing the Impact of Microenterprise Services (AIMS) Project.

Individuals representing a variety of disciplines have prepared the papers related to microfinance impact assessment. For the first virtual meeting held in April 1997, the background papers addressed methodological approaches and practical issues to consider in planning, designing, and carrying out impact assessments. These papers covered transaction costs, income and assets as impact indicators, a participatory action and learning approach, use of control groups, credit expansion, and the perspectives of an NGO practitioner (Appendix 2). The 1997 virtual meeting centered on a discussion paper, "Impact Assessment Methodologies for Microfinance: A Review," by David Hulme. Monique Cohen and Gary Gaile wrote up the highlights of the 1997 meeting (Appendix 1).

As part of this same process, for the 1998 and 1999 virtual meetings CGAP members submitted or commissioned microfinance impact studies or study designs of programs in 11 different countries (Appendix 2). These reflect various methodological approaches. The actual experience and results of these studies have been used to feed into the development of guidelines that are credible, useful, and cost effective. The discussion in the 1998 meeting centered on working towards guidelines for lower cost impact assessments, based on a paper by Jennefer Sebstad that built on the studies submitted by CGAP members. Monique Cohen and Gary Gaile also wrote up the highlights of the 1998 meeting (Appendix 1).

In total, the virtual meetings have involved some 100 donors, researchers, evaluation specialists, and practitioners from around the world. Approximately 25 individuals have participated in all three of the virtual meetings.

This present report builds upon and draws directly from the various background papers, field studies, discussion papers, and exchanges during the virtual meetings. The contributions of those involved in these activities are gratefully acknowledged. To streamline the presentation of the guidelines, we have not provided direct references to all of this work in the text. However, we especially wish to acknowledge the ideas and contributions of David Hulme, Monique Cohen, Gary Gaile, Elizabeth Dunn, Elaine Edgcomb and other members of the AIMS team, which provide the foundation for these guidelines.

A draft of the report was the focal point for the discussion at the third CGAP virtual meeting. Thereafter, the guidelines were modified to reflect the discussion and specific suggestions.

## TABLE OF CONTENTS

I.	Introduction.....	1
A.	Purpose of the Guidelines .....	1
B.	A Role for Standards .....	2
II.	What is an Impact Assessment?.....	4
A.	Key Features.....	4
B.	Comparison of High, Low, and Middle Range Approaches .....	5
C.	Impact Assessments and Impact Monitoring .....	6
III.	Guiding Principles for a Microfinance Impact Assessment.....	8
A.	Credibility .....	8
B.	Usefulness .....	9
C.	Cost Effectiveness .....	9
IV.	Planning for a Microfinance Impact Assessment.....	11
A.	Main Parameters .....	11
B.	Design Framework .....	13
C.	Other Decisions .....	20
V.	Guidelines for Qualitative Approaches to Microfinance Impact Assessments.....	23
A.	Purpose of Qualitative Approaches.....	23
B.	Design Stage.....	24
C.	Implementation.....	30
D.	Analysis.....	31
E.	Dissemination .....	32
VI.	Guidelines for Quantitative Approaches to Microfinance Impact Assessments.....	34
A.	Purpose of Quantitative Approaches.....	34
B.	Design Stage.....	35
C.	Implementation.....	43
D.	Analysis.....	48
E.	Dissemination .....	50
VII.	Checklist of Key Steps for Planning and Implementing a Microfinance Impact Assessment.....	51
	REFERENCES .....	59

## APPENDICES

Appendix 1: List of Discussion Papers Prepared for CGAP Impact Assessment Methodologies Working Group, 1997-1999 .....	59
Appendix 2: List of Background Papers Submitted for CGAP Impact Assessment Methodologies 1998 and 1999 Virtual Meetings .....	60
Appendix 3: Variables with Demonstrated Validity in Previous Microfinance Impact Assessments .....	62
Appendix 4: MFI Performance Indicators and Related Impact Information .....	65

## LIST OF TABLES

Table 1: Basic Differences Between Qualitative and Quantitative Studies .....	13
Table 2: Menu of Selected Impact Assessment Methods .....	14
Table 3: Choosing the Right Method.....	15
Table 4: Trade Offs in Choice of Whom to Include in an Impact Assessment Survey .....	17
Table 5: Trade Offs in Choice of Longitudinal or Retrospective Design .....	20
Table 6: Trade Offs in Involvement of MFI Staff and Managers in an Impact Assessment .....	21
Table 7: Options for Selecting Individuals for Qualitative Studies .....	26
Table 8: Example of Criteria for Selecting Case Studies.....	28
Table 9: Example of a Matrix for Organizing Data on Loan Use.....	32
Table 10: Alternative Sample Designs for Middle-range Microfinance Impact Assessments .....	40
Table 11: Trade Offs in Choice of Sample Size for Quantitative Impact Assessment Surveys .....	41
Table 12: Options for Selecting Individuals for Quantitative Surveys .....	41
Table 13: Client/Non-Client Comparison of the Composite Impact Index and its Component Scores .....	49

## EXECUTIVE SUMMARY

Development organizations undertake impact assessments to determine the difference they are making and to help them improve programs. The information in this document is intended to guide the planning, design and conduct of microfinance impact assessments that are credible, useful, and cost effective. It is the outcome of a series of papers and virtual meetings held on behalf of the Consultative Group to Assist the Poorest (CGAP) Working Group on Impact Assessment Methodologies. The virtual meetings have involved donors, researchers, evaluation specialists and practitioners from around the world.

The guidelines center on impact assessments (IAs) that can be used by microfinance providers, microfinance promoters, donors, and development policy makers to improve microfinance products and services, and justify investments in microfinance programs. In particular, the manual is designed to help managers develop an action plan or scope of work for an IA and assist those leading the assessment to think through the process. It is specifically geared to those who have limited experience with IAs of microfinance programs. As such, it has an educational objective. At the same time, the guidelines provide a checklist of key steps that may be useful to the more experienced managers and researchers.

An impact assessment is conducted to identify changes that have occurred as a result of a program. To do this one needs to be able to make a case that the changes are not due to other factors and would not have occurred without the program. The ability to establish plausible association between the changes identified and program participation strengthens the findings and their interpretation.

The guidelines seek to set forth standards for impact assessments that are credible, useful and cost-effective. In contrast, the standards for rigorous IAs set by the academic community and statisticians are normally beyond the financial reach of organizations seeking to better understand the impact of specific microfinance programs. At the other extreme, IAs have been carried out that are low-cost but the conclusions about impact are not substantiated by the methodology used or its execution. This document centers on a middle-range approach that has evolved through a series of papers and discussions through virtual meetings that address methodological approaches and practical issues.

Rather than advocating a particular impact assessment design, the guidelines lay out a general framework. This framework is intended to be flexible enough to take into account different types of programs and different contexts. It recognizes that there are always tradeoffs and choices to be made, and in varying degrees these have implications for credibility and usefulness. The guidelines are based on there being a role for both qualitative and quantitative methods and the active participation of MFI program managers in the IA process.

Planning for a microfinance impact assessment should involve representatives of the key stakeholders, and this usually means MFI managers and financial supporters. The first step involves determining the IA's objectives and intended use of the results, followed by articulation of the key research questions. Decisions about the methods to be used, including whether or not to use a non-client comparison group, and whether the assessment should be based on a one-time

retrospective design or a longitudinal approach have to be made. At the planning stage, a series of other design questions also need to be addressed at least tentatively in order to calculate the estimated level of effort required, cost and schedule.

Thereafter, these initial decisions are used to guide the design of the IA. The feasibility of the initial plan is determined in the design stage and the decisions made in the planning stage may be revisited. At this stage, those leading the assessment need to think carefully through each step to ensure that the key questions are feasible to answer. They also need to conceptualize how microfinance services lead to changes and what changes are reasonable to expect given the services provided and loan conditions. Attention should also be given to how the information will be gathered and analyzed. Related to the latter, the design needs to pay careful attention to the basis on which claims of impact can be substantiated.

The implementation stage involves finalizing logistical arrangements, selecting and training those who will conduct the interviews or facilitate the qualitative data gathering effort, testing the instruments, and supervising enumerators or facilitators. The latter is essential because unusual and unexpected things arise that require professional direction on how to handle them. The guidelines highlight key items under each of these topics.

Guidance is also provided on data analysis and report writing. Survey data should be analyzed using averages, frequency counts and distributions as well as provide the results of simple statistical tests. Not all of the information should be included in the core of the report; an annex should be provided to substantiate, amplify and expand the discussion in the report. When a mixed method approach is used, the qualitative information should be integrated and not be a separate component of the report.

Since a common weakness with qualitative studies is that the information gathered is described rather than analyzed, the guidelines give attention to analysis of qualitative information. Qualitative information needs to be organized and meaningfully reduced by selecting, focusing and simplifying it. One should highlight patterns and common themes and explain deviations from the patterns. Also, the qualitative component of the IA should provide interesting stories that help illuminate the broader study questions.

Those charged with leading the IA are normally the individuals responsible for writing up the findings, analysis and conclusions. Nevertheless, those engaged at the field level in the MFI have valuable information and insights that may enhance the analysis and conclusions. Along with other key stakeholders, they ought to be brought into the analysis process.

The information for the final report should be organized so that it addresses the key IA study questions. The analysis and conclusions should be substantiated by the data presented and the methodology used.

# I. Introduction

## A. Purpose of the Guidelines

The guidelines presented in this report are intended to guide the planning, design and conduct of microfinance impact assessments that are credible, useful, and cost effective. They center on impact assessments (IAs) that can be used by microfinance providers, microfinance promoters, donors, and development policy makers to *improve microfinance products and services and justify investments in microfinance programs*. The guidelines aim to inform decisions about the type of impact methodology to use in light of different IA objectives, program contexts, and resources (Hulme 1997). They cover both qualitative and quantitative impact assessment methods and highlight the potential role of each.

This manual is intended to help planners develop a scope of work or action plan for an IA, and assist those leading an impact assessment to think through the process. It is specifically geared to *those who have limited experience with IAs of microfinance programs*. For those with experience in impact assessments, the guidelines provide a checklist of key steps. They are based on the premise that the views and concerns of key stakeholders, including clients, program providers, and program donors are important and legitimate.

The guidelines set forth basic elements of impact assessments that can serve as a standard for the microfinance field. They aim for a practical and affordable approach that is within reach of potential funders of IAs and those likely to carry out such assessments. Standards for conducting microfinance impact assessments need not be as high as academic standards. Nonetheless, they should help to insure that microfinance impact assessments meet a minimum standard of methodological rigor so that they are credible.

In setting these forth, it should be noted that there is a role and need for impact assessments that are more rigorous than those addressed in these guidelines. In addition, impact monitoring has a role within microfinance organizations, and may be part of an impact assessment system. Moreover, evaluations that cover program cost-effectiveness can provide vital information, leading to program improvements. An impact assessment is one element of a program evaluation. The guidelines herein for IAs are intended to complement these other approaches to understanding program clients, program effectiveness, and impacts.

Rather than advocating a particular impact assessment design, the guidelines lay out a general framework. This framework is intended to be flexible enough to take into account different types of programs and different contexts. It is applicable irrespective of program objectives and specific impact data needs. In addition, the steps identified in the framework are generally applicable to assessments that cover more than a single program. The guidelines recognize that there always are tradeoffs and choices that must be made in planning and

implementing an impact assessment and that these tradeoffs have implications for credibility and usefulness. Hence, while addressing a range of options, the guidelines underscore basic features for credible and useful impact assessments.

## B. A Role for Standards

Impact assessors often deal with skeptical audiences or stakeholders who seek to discredit findings that are either too critical or too uncritical of a particular program. Depending on the audience, a particular methodology may be questioned as unsound or weak. Ensuing debates about the relative merits of a particular methodology may focus on sample size, control group, the validity of the variables studied, or findings that may be dismissed as anecdotal. In many cases, these debates boil down to different views on ways of knowing (Hulme 1997).

Also, the results of IAs often are subject to criticism when they do not employ an experimental design. This is the most rigorous type of design for attributing changes to participation in a program and hence determining program impacts. An experimental design requires that potential clients be randomly assigned to a “control” or “program treatment” group in order to control for self-selection bias and other factors. Program providers in particular have argued that this type of IA design is too intrusive, and unethical since those who want a service should not have it withheld.

To move beyond these debates, it is important to build consensus within the microfinance field on basic standards for impact assessments. As part of this process, it is important to accept that there is a role for both qualitative and quantitative methodologies in assessing impacts. Both are legitimate ways of knowing. Thus, establishing standards for both qualitative and quantitative methods are important. It also is important to establish standards that allow assessments to make a sound case that participation in a microfinance program has led to the changes identified among clients, and hence to label these changes as program impacts.

The dictionary definition of a *standard* is, “... anything serving as a rule for making judgments or as a basis for comparison.” Standards for microfinance impact assessments thus can be thought of as the basic elements necessary for them to be considered ‘credible’ by a range of actors in the field for the dual purposes of justifying investments and improving programs. Actors may include, for example, microfinance promoters, microfinance providers, and those responsible for making decisions about the allocation of development resources to microfinance.

A benefit of standards is that they can provide an objective (or less subjective) basis upon which to judge or defend the credibility of a particular study. At the same time, they should be ‘calibrated’ at a level that is neither too high nor too low for development practitioners. For microfinance programs and funders with limited resources, the standard is likely to be something less than a complex academic approach, but something more than a simple non-academic approach. It must be practical and affordable. At a minimum, an impact assessment should provide a sound basis for establishing a relationship between participation in a microfinance program and change in people’s lives. Taking these factors into account, these guidelines advocate a middle range IA approach, that is, one that is credible given its objective but affordable in terms of human and financial resources.



The CGAP Impact Assessment Methodologies Working Group has begun the process of building consensus within the field on the *need* for, the *definition* of, and the *appropriate use* of standards for microfinance impact assessments. Towards this end, the following sections discuss guidelines for a range of both quantitative and qualitative methodologies typically used in impact assessments. These include surveys, focus group sessions, case studies, individual interviews, and participatory appraisal and learning techniques.

## II. What is an Impact Assessment?

*"Establishing impact essentially is making a case that the program led to the observed or stated changes. This means that the changes are more likely to occur with program participation than without program participation. It does not imply that the changes always occur from program participation. Rather, it increases the probability that the changes will occur." (Rossi and Freeman 1989)*

### A. Key Features

An impact assessment (IA) is a study to identify changes that result from a program by employing methods to establish plausible association between changes experienced and participation in the program. A simple paradigm for an impact assessment is: X causes Y or a program results in changes. In reality, however, other factors intervene to influence the impacts (e.g., gender, role of enterprise income in the household, location of the enterprise). Also, Y may happen irrespective of X, so it is necessary to pay attention to attribution and rule out plausible rival reasons about why the changes may have occurred. The level and nature of program participation should affect the impacts of the program, so this needs to be taken into account. IAs may link an institutional review of program components and procedures with client level data to determine what is working well and what can be improved.

#### WHAT IS A QUASI-EXPERIMENTAL DESIGN CONTROL GROUP?

A control group under a quasi-experimental research design is a group of microentrepreneurs who have not received program services, but are similar to program participants on key factors, such as gender and geographic location. Changes in this group are treated as changes that would have occurred among program clients irrespective of program participation.

In comparison, a control group under an experimental research design consists of persons randomly assigned to the control rather than the program participant group.

Use of a quasi-experimental design control group is a common approach for ruling out other possible reasons for the changes or for noting why anticipated changes have NOT occurred. The lives of microentrepreneurs, especially the poor are complex and this complexity is compounded by unstable and fluctuating macroeconomic situations and other forces that extend far beyond their control, yet affect their lives and immediate environment. These external factors are taken into account and controlled by including in the assessment a control group of non-client microentrepreneurs within the same geographic area. ***It should be noted that henceforth the term control group refers specifically to a control group selected as part of a quasi-experimental design.***

Impact assessments compare changes in impact variables between two or more points in time. This can be accomplished through a longitudinal study consisting of a baseline and one or more follow-up studies using the same variables and measures. Or, it can be done by a one-time retrospective study that compares the present with a previous point in time in order to assess changes.

## B. Comparison of High, Low, and Middle Range Approaches

Making a case that a particular microfinance program led to an observed or stated change can be done in several ways. Approaches can vary in their level of complexity. Complex approaches, for example, may involve econometric models that require rigorous assumptions about behaviors to obtain control mechanisms and parameter estimates. The use of this approach requires knowledge of production functions, utility, and other econometric concepts that may be unfamiliar and off-putting to many potential users of impact assessments (Gaile and Foster 1996). Or they may involve large scale sample surveys based on a quasi-experimental design that compare the outcomes of an intervention with a simulation of what the outcomes would have been had there been no intervention. A large sample allows the researcher to use more sophisticated controls and analytical techniques, test a larger number of variables, and attribute the changes in the impact variables to the intervention with a high degree of confidence. Complex approaches tend to be expensive, to take more time and be outside the capacity of many microfinance programs to fund.

To date, simple IA approaches have tended to dominate the microfinance field. They are smaller and less expensive. They tend to use a single method, in many cases a onetime survey covering a small sample of clients without a control group. These surveys often are based on recall data. In other instances, a cross-sectional design is erroneously used for causal inferences: for example, the current differences between first year, second year, and third year clients are taken to be a result of program participation.<sup>1</sup> In other cases, the simple approach has involved using a few case studies to illustrate program impacts. The findings generated by simple approaches are less reliable for justifying investments, although they can be useful for understanding impact processes and improving programs.

Somewhere between complex and simple is *a middle range* approach, advocated in these guidelines. A middle range IA may be defined as an inquiry to estimate the value, degree and/or direction of change that can be plausibly associated with an intervention. Compared with complex approaches, they generally are smaller in size, more limited in scope, and use less-complex measures and more simple analytic techniques. Middle range IAs should involve a mix of methods including surveys, case studies, focus group interviews, individual interviews or other qualitative methods. In addition, middle range IAs should normally cover sample groups with and without the intervention to establish “plausible association.” Rather than measure impacts within precise and statistically definable limits of probability, these impact assessments seek to understand intervention processes and to identify and reliably estimate impacts that stand the test of plausible association. The estimation of impacts is often based largely on trend data.

---

<sup>1</sup> This simple approach does not produce valid results. When a cross-sectional survey design is used for causal inferences, the data must be analyzed by structural equation models and related techniques (GAO 1991). When this type of analysis is carried out, it moves from being a simple approach to a middle range approach.

A mix of methods may be used to establish the validity of a relationship between interventions and changes in selected impact variables through triangulation. This involves studying the same hypothesis and impact variable through different methods to determine if the same conclusions can be reached about program impacts. For example, the hypothesis that *participation in a microfinance programs leads to increases in enterprise fixed assets* would be tested by including in the survey, questions about fixed assets acquired for the enterprise over the past year and source of funds used, while the case studies would explore and explain in more depth what fixed assets were acquired, why, how they were used, how they were financed, and the participant’s plans for growing or diversifying the enterprise.

Methods also may be used sequentially. Qualitative approaches can be used to help identify the most likely impact paths and the hypotheses and impact variables to incorporate into a survey questionnaire. They also can aid in designing survey questions and response categories that capture the essence of complex phenomena. Qualitative approaches also may follow a survey to assist in analysis and interpretation of the data. It is possible, although not common, to use a mix of qualitative methods to highlight impacts. This approach might involve, for example, case studies of clients and non-clients combined with information from participatory rapid appraisal methods.

**C. Impact Assessments and Impact Monitoring**

Impact assessments are distinct from impact monitoring in several ways, although they have complementary roles.<sup>2</sup> The foundation of an impact assessment is a focus on *causal linkages* to determine changes that have resulted from program participation. It focuses specifically on the relationship between the program and change (planned and unplanned). As mentioned above, the guidelines advocate that middle range IAs involve a mix of methods and specifically account for external factors that

DISTINGUISHING FEATURES
<p>Simple Impact Monitoring: This involves the collection of information on a regular basis for all clients on a small number of variables using standardized questions. It involves periodic analysis of the data. It does not address whether or not the changes identified are a result of program participation.</p>
<p>Middle Range Impact Assessments: This involves the collection of data according to a set plan and timeframe from a sample of clients and a comparison group in order to make a case for program impact.</p>

influence whether or not change has occurred in order to isolate those changes that can be plausibly associated with the program. The IA guidelines also stress moving beyond this foundation to *explore* in more depth the change process, why specific changes have occurred, and what they mean to the client, and to *describe* whom the program reaches. When the main reason for an impact assessment is program improvement, these IAs are likely to include retrospective and prospective questions dealing with client satisfaction, and seek suggestions on ways to extend and deepen outreach through different products and improved services.

In comparison, impact monitoring normally gathers information on a standardized and limited set of variables from clients on a regular basis. The most common approach is to collect

---

<sup>2</sup> With the current interest in and focus on impact monitoring, the components of the new systems often are very similar to those discussed herein for impact assessment.

and analyze information on those who are in the program to determine a) who the program reaches and b) changes among those who obtain additional loans.<sup>3</sup> In addition, one may analyze outreach data (such as gender of clients, enterprise sector, and number of loans received) with financial data to identify factors such as those related to arrears rates or repeat borrowing. This approach to impact monitoring generates data rapidly and in predictable periods. Impact monitoring, like impact assessments, also may cover client satisfaction and aim at improving products, outreach, and services.

Impact monitoring and impact assessments have complementary and synergistic roles. Which should come first is open to debate: impact monitoring or an impact assessment. Impact assessments can help microfinance programs to identify the impacts that occur most often for all clients or for different subgroups. They also can help in developing simple measures for tracking these impacts. Occasional impact assessments using mixed methods and a non-client comparison group can help to verify links to program interventions and amplify data from monitoring systems. Monitoring systems can serve to inform the design of impact assessments. They provide a sound basis for establishing the sample framework and selecting clients, and make it easier to locate the client sample. In addition, the monitoring data are likely to provide insights in selecting key questions and hypotheses to guide the design of the impact assessment.

Thus, *the main differences between impact monitoring and impact assessment are:*

- the extent to which the findings make a case that the changes are a result of program participation and hence identify program impacts;
- the populations covered;
- the frequency of data collection and analysis;
- the scope of coverage; and
- the methods employed.

They also have different implications for the resources required. Impact assessments require a lump sum of funding for a given period. Impact monitoring requires an up front expenditure to establish the data gathering system, and then ongoing funding to maintain the system and analyze and report on the data. In specific instances, there may be great similarities between impact monitoring and middle range impact assessments depending on how these are designed and established.<sup>4</sup> Or, stated in another way, a comprehensive system would contain a client tracking system, periodic impact assessments, and ad hoc special studies to address key issues (Hyman and Dearden 1998). These guidelines focus on middle-range impact assessments whether or not these are part of a more comprehensive system.

---

<sup>3</sup> E. Hyman and K. Dearden (1997 and 1998) assess the experience and practice of some 30 non-governmental organizations with tracking the impact of microenterprise programs.

<sup>4</sup> When a program performance results monitoring system includes impact indicators, the approach is often a hybrid of impact monitoring and impact assessments. In such systems, data on a few specified impact indicators are collected following a standardized format either annually or bi-annually on a random sample of clients, and may or may not include a control group.

### III.

## Guiding Principles for Microfinance Impact Assessments

---

#### A. Credibility

For a middle range IA to be *credible* it should begin with clearly stated objectives that indicate the types of impacts that will be examined, the intended use of the findings, and the audience. It should have a small set of key hypotheses, some of which have proven valid in previous IAs. Moreover, the measures used to test the hypotheses need to yield reliable and valid data.

The IA should be designed to establish plausible association between changes identified and participation in the microfinance program. Towards this end, quantitative and qualitative studies should be based on a longitudinal design, if possible, to obtain more reliable measures of change. If a longitudinal design is not possible, the quantitative assessment should concentrate on variables for which recall data are easily obtainable and generally reliable. It should employ a comparison group, preferably a quasi-experimental design control group of microentrepreneurs, to provide a basis for associating change with participation in the microfinance program. In one-time studies it is essential that information on changes over a designated timeframe be gathered; it is not valid to compare the current status of clients with non-clients and claim that the better results among clients are due to program participation. Rather, one assesses trends or changes over a period of time between clients and non-clients in order to make a case that the differences identified between the two groups are a result of program participation.

A quantitative assessment should have a sample size that is large enough to ensure effective use of control variables, account for refusals and non-finds, and allow for invalid data issues, but small enough to fit the budget. Here is where trade-offs are required between the number of variables, margin of error, confidence interval, and budget.

Credibility is intricately linked with the way in which the data are used to report on program impacts. This means that claims to impact need to be supported by the sampling design and methods, and by presentation of the data accompanied by the statistical tests of significance.

The IA should use a mix of research methods, for example, small surveys combined with focus groups, case studies, or participatory approaches, and secondary data sources. Information generated by mixed methods can help to establish the validity of the data and the reliability of the measures of change. The credibility of an IA can be improved further by using data-gathering instruments that are well designed and clearly documented. Equally important, the selection of the study participants should not be biased, and adequate time needs to be given to train people who will conduct the study. Supervision and oversight of data collection helps to further enhance credibility. The IA report needs to provide sufficient information to assure the reader that the necessary steps and precautions were taken. When feasible, the IA report should include an analysis of the loan repayment history of the study participants or their credit groups.

## B. Usefulness

Credibility is also a function of usefulness. To be *useful*, an IA must be designed to address the key questions and concerns of the intended users. While the findings may be used to both “justify” and “improve” the program, the primary objective ought to guide the major planning and design decisions for a concise and well-focused IA.

Usefulness is enhanced when those who are expected to use the findings are involved in the planning, design, and analysis stages. At a minimum, these individuals should participate in setting the study objectives and the key research questions. They also should agree on the final research design and provide review comments on a draft of the report. When the key person directing all of the steps belongs to the program being assessed, this individual needs to make sure that the key decision-makers in the organization are brought into the process. Involvement of the key intended users adds to the credibility of the study results when their concerns are reflected in the IA design, execution, and analysis.

If the main objective of an IA is to justify program expenditures and program effectiveness, the results are likely to be useful to the donors and promoters of the organization. They also are likely to be useful to executive officers in accounting to their Board of Directors and generating further support for the organization. The findings may be used to help determine if the program is reaching the intended clientele, and to highlight characteristics of those who remain in the program, those who leave, and those who never join. Impact assessment data also can be used to define strategic objectives, design and deliver appropriate products, and suggest new products. Finally it can be useful for developing strategies to improve portfolio performance by reducing turnover, expanding outreach, and improving portfolio quality. Involving program managers in the impact assessment process is important in insuring that the issues addressed are meaningful and the results are useful. Another element of usefulness is developing a specific dissemination strategy and presenting findings to the intended audience in a timely and comprehensible fashion following completion of the IA.

A key element of usefulness is the timeliness of the data. In order to have the information available when needed, the IA plan and schedule need to be realistic. Also, the design and implementation need to be carried out with an eye on making the data available in a timely manner in order to inform decisions.

## C. Cost Effectiveness

A challenge in designing a middle-range impact assessment is deciding how to get credible and useful results out of a modest budget. An IA can be more *cost-effective* if there is a good “fit” between the objectives, methods, and resources available in terms of money, people, and time. Greater efficiencies can be achieved by building on the lessons of past IAs. Past experience can be especially helpful in identifying meaningful and valid impact hypotheses and variables, developing data-collection strategies for obtaining reliable and valid data, and selecting appropriate analytical techniques. For qualitative IAs or IA components, past experience with applying the same method and techniques can increase cost-effectiveness.

Resources can be allocated more efficiently if the data needed to test the hypotheses are considered at the planning stage and the expertise required at different stages of the process is considered and adequately budgeted for *a priori*. Keeping the sample to a manageable size and using a mix of measures that rely more on the direction or pattern of change than the amount of change can reduce data collection costs. The cost effectiveness of an IA also can be improved if the instrument is pilot tested.

Cost-effectiveness also is linked with the receptivity of the intended users and their organizations to learning. MFIs that constantly strive to be up on the learning curve, so as to achieve their objectives and serve their clients, are likely to be most receptive. Progress tends to occur when learning leads to appropriate actions. Part of establishing a learning environment is building an IA capacity within a MFI program and within the program's local environment so that it can be drawn upon when needed. Donors and other program supporters need to consider IAs and their results within a culture of learning: learning from weakness/failures as well as good practices.



## IV.

# Planning for a Microfinance Impact Assessment

---

A well-planned microfinance impact assessment sets the stage for a credible, useful, and cost effective study. Key questions to guide the planning stage are presented below. The planning process essentially structures the action plan for an impact assessment.

### A. Main Parameters

#### 1. What Are the Objectives and Intended Use of the Impact Assessment?

The credibility of an assessment is enhanced by establishing clear and realistic expectations about the objectives of the assessment (e.g. improve understanding of program impacts on clients), how the information will be used (ranging from justifying investments to improving programs), and by whom (e.g. practitioners, promoters, donors, policy makers, academics). This will help determine the types of impacts that will be examined (social and/or economic), the methodology that will be used (quantitative methods, qualitative methods, or a mix), the types of data that will be collected, the level of reliability that will be required, and the budget (Hulme 1997).

Objectives have implications not only for the audience, but also influence whether the assessment will address questions that are more theoretical or practical and whether the findings can be generalized or are context-specific. Finally, the objectives will have implications for the time scale of the assessment and the degree of confidence expected from the findings (Hulme 1997). If the objective dictates a high degree of confidence and statistical significance in the findings, something beyond a middle range approach should be considered.

The intended use of the findings, bridging improving a program and justifying its expenditures, also guides the type of information to collect beyond who the program reaches and program impacts. The improving objective may lead to the inclusion of other types of questions, for example, those related to client satisfaction with products and services.

#### 2. What Are the Main Features of the Microfinance Program and the Environment in Which It Operates?

Basic information on the program is important for defining research questions and selecting a methodology. What type of program is it? How complex is it? How large is it? What are its objectives? What documentation already exists about the program? What key features characterize the context in which the program operates? These issues are important to address in the planning stage because they influence choices as to types of impacts to be studied and the appropriate methods to use.

The planning stage should involve identifying key program and context information to collect as part of the IA. In part, the choice should consider the availability of existing information. Program information should establish the extent to which the program operates effectively in its environment and is financially sustainable.<sup>5</sup> At a minimum this should include information on the objective of the MFI, its products and services, the date it was established, and its portfolio performance and financial outcomes over the past two to three years. Also, information on the proportion of clients who are repeat borrowers places the sample and findings in a broader context. The information should provide a clear, accurate, and comprehensive description of the operations of the microfinance program, especially its credit activity. The credibility of impact assessments can be enhanced if this program information is linked to the hypotheses and/or used at the analysis stage in interpreting the findings (Snodgrass 1997). Many previous impact assessments provide good examples of the type of program information to collect, and how it can be linked to impact analysis (Edgcomb and Garber 1998; Ouattara, *et al.* 1997).

EXAMPLE OF CONTEXT TOPIC

*Local financial market* includes local formal and informal institutions, as well as government policies and regulations covering the microfinance field. (GRET 1998)

Context influences microfinance program outcomes (outreach, financial sustainability, and impact) by affecting the operation of the program and the economic activities of clients. The physical environment, formal and informal institutions, economic factors, and government policies and regulations all may affect program impacts. In planning for the type of context information to collect in the IA, it is useful to consider factors that may influence microenterprises and microfinance programs. These include, for example, seasonality, natural catastrophes, inflation, economic growth levels and patterns, ethnicity, and local government regulations. The context also includes the local microfinance market, in particular the level and type of competition.

IA planners need to determine what information can be collected formally from primary and secondary sources, and what information can be collected informally during the data interpretation sessions with MFI program managers and staff. In terms of using context information in analyzing the IA data and interpreting the findings, it is useful to observe whether a particular hypothesized influence is present or absent, and whether significant changes in these factors have occurred during the assessment period (Snodgrass 1997).

### 3. What Are the Key Research Questions?

Intertwined with setting forth the objectives of the assessment and the use of the findings is defining the key research questions. ***What does one want to know and why?*** The planner should be able to answer the following: *I need to know \_\_\_\_\_ because I need to decide \_\_\_\_\_.* The key questions should be the overarching ones and limited to the most important that need to be addressed. Thereafter, it is important to determine if the key questions are posed correctly, are answerable, and if so, at what cost (GAO 1991). In deciding this, planners should consider the degree of precision needed.

---

<sup>5</sup> An IA can also assess MFI performance by utilizing MIS data to complement the original data collected for the IA. See Appendix 4.

The key questions should lead to the initial identification of the best methods for addressing them and an estimate of the resources required: time, people and cost. In the design phase, these questions ought to guide the selection of hypotheses, identification of variables, and development of additional questions to be answered by the impact assessment.

EXAMPLES OF KEY RESEARCH QUESTIONS	
1.	Which groups among the poor does UWFT reach?
2.	How do financial services improve client's capacity to manage, control, and build up their asset base to protect against and cope with risks?
3.	How do microfinance services enable clients to use assets to maintain a minimum economic threshold? (Wright, <i>et al.</i> 1999)

**B. Design Framework**

**1. What Method or Mix of Methods Should Be Used?**

Strong impact assessments employ methods that are appropriate to the key questions and the degree and extent of precision needed. The choice of method to use for the impact assessment also depends on the purpose of the assessment and the audience. Common quantitative methods include sample surveys and semi-structured interviews. Common qualitative methods include focus groups, case studies, individual interviews based on key open-ended questions, participatory appraisals, and participatory learning activities. Qualitative approaches can inform quantitative approaches and vice versa. The interaction between qualitative and quantitative approaches clearly can enhance an IA. Tables 1 through 3 list key features and strengths of each of these methods and offer guidance in choosing the right method.

TABLE 1: BASIC DIFFERENCES BETWEEN QUALITATIVE AND QUANTITATIVE STUDIES	
Qualitative	- Approach to address questions related to how and why and what else (unanticipated results) - Results are generalizable to theoretic proposition, NOT to populations or universes
Quantitative surveys	- Approach to address questions related to what, who, where how many, how much, the incidence and prevalence of a phenomena - Results may be generalizable to a larger population, depending on the sampling technique

Adapted from Yin 1994.

TABLE 2: MENU OF SELECTED IMPACT ASSESSMENT METHODS		
Method	Key Features	Strengths
Sample surveys	Collection of quantifiable data. Usually uses a random sample and matched control group to associate changes in the client sample with the project or program. Can be retrospective (based on recall) or administered more than once, the first establishing the base period.	Coverage, representativeness, ease of data standardization; aggregation and synthesis; ability to isolate non-project causes.
Semi structured interviews	Allows for quantitative and qualitative data to be collected from a small number of clients and non-clients. The intended use should determine sampling method and criteria.	Permits follow-up on client responses. Can combine collection quantitative and qualitative information.
Focus groups	Uses a few key questions to generate discussion that yields clients' views and opinions. Works best when the group is relatively homogeneous to allow conclusions to be based on common characteristics.	Use of group dynamic to elicit responses; captures diverse perceptions and negative impacts.
Case studies	Involves detailed studies of a specific unit (group or individual) with open-ended questions based on an interview guide. Histories can trace events from prior to joining the program to the current period.	Ability to capture unexpected impacts and the pattern and reasons for change.
Participatory self learning	Involves a facilitator working with small groups, or sometimes individuals, to stimulate them to discuss and analyze information about their lives and community. Normally tools are used to encourage discussion of topics.	Potential to contribute to capacity building, identify felt needs, client values and perceptions. Information from groups more accurate than individual interviews when validated in an open forum.

Adapted from Hulme 1997.

TABLE 3: CHOOSING THE RIGHT METHOD

Method	When Appropriate	When Not Appropriate
Sample surveys	<ul style="list-style-type: none"> <li>- Project affects a large number of clients</li> <li>- Audience requires accurate estimates of project impacts</li> <li>- Users need statistical comparisons between groups over time or between geographic locations</li> <li>- Project delivery/implementation mechanisms working well over time, thereby justifying investment in assessment</li> <li>- Target and/or client population is heterogeneous and it is important to isolate the influence of factors unrelated to the project</li> </ul>	<ul style="list-style-type: none"> <li>- Project has a small number of people</li> <li>- Audience is concerned primarily with project outcomes</li> <li>- Project implementation is recent and untested and likely to have little impact</li> <li>- The purpose of the assessment is to understand whether the project is meeting the felt needs of the project clientele</li> </ul>
Semi structured interviews or mix of closed and open ended questions	<ul style="list-style-type: none"> <li>- Purpose is to explore or document what changes are due to the MFI program</li> <li>- Purpose is to explore questions and measures in order to develop a survey instrument</li> <li>- Users want timely information on specific issues or questions, such as reasons for arrears</li> <li>- Time limited for development and testing of a survey based largely on close-ended questions</li> <li>- Users require both quantitative data and information that helps to explain the data</li> </ul>	<ul style="list-style-type: none"> <li>- Users want information from a large number of individuals</li> <li>- Users want statistical comparisons between groups or locations</li> <li>- Clients are to be active participants in the process</li> </ul>
Focus groups	<ul style="list-style-type: none"> <li>- Priority is to understand motivations and perceptions of clients</li> <li>- Want to stimulate reflection and discussion on client satisfaction or group dynamics</li> <li>- Need information to interpret quantitative data</li> <li>- Need information quickly to address an issue</li> <li>- Participatory principles are a priority for the MFI</li> </ul>	<ul style="list-style-type: none"> <li>- Need information based on representativeness and to generalize information from participants to a larger population</li> <li>- Want to understand the socioeconomic level of project participants</li> <li>- Local culture prevents individuals from feeling relatively free to express their opinions in a group situation</li> </ul>
Case studies	<ul style="list-style-type: none"> <li>- Need to understand the causal impact process</li> <li>- Want to search out rival explanations for change and unexpected or negative impacts</li> <li>- Need to illuminate and put a human face on quantitative data</li> <li>- Want to understand changes over a long period of time</li> <li>- Need to identify reasons for lack of change in survey impact variables</li> </ul>	<ul style="list-style-type: none"> <li>- Need information about a large number of clients and want to generalize the information gathered to a larger population</li> <li>- When indicators of program impact are clear and easily measurable, and negative impacts unlikely</li> <li>- When information is needed quickly</li> </ul>

Participatory self learning	<ul style="list-style-type: none"> <li>- Program promotes participatory principles</li> <li>- Program's objectives include empowering clients</li> <li>- Want to establish structure for linking participatory learning to program improvement</li> <li>- Attention given to community impacts</li> <li>- Understanding client motivations and perceptions a priority</li> </ul>	<ul style="list-style-type: none"> <li>- Do not have access to highly skilled persons to facilitate discussion</li> <li>- When planners not concerned about participants learning from the assessment process</li> <li>- Sole priority is standardized and statistically representative data for a large and diverse program population</li> <li>- Existing tools inappropriate and do not have time to develop and adequately test new ones</li> </ul>
-----------------------------	--	---

Adapted from Hulme 1997 and Hulme 1999.

While a good IA should span both the ‘justifying’ and ‘improving’ objectives, the *primary* objective of the study should guide the mix of methods to use. If the primary objective is justifying investments, then it is necessary for the IA to present a strong case that the program is related to the change. Either quantitative or qualitative methods can be used to establish this relationship. Quantitative methods, by using a control group, can account for changes that would have occurred without the program and permit generalizing the findings beyond the study sample to the same population. Qualitative methods can rule out rival reasons for the changes. However, they do not allow one to generalize to a larger population. Thus, qualitative data alone have more limited utility in addressing a larger client population. A mix of both quantitative and qualitative methods, advocated here, can generate findings that may be generalized beyond the study respondents and provide a deeper understanding of the relationship between the program and change.

If the primary objective of the IA is improving the program, again, both quantitative and qualitative IA methods are useful. Quantitative surveys can include questions related to improving programs, for example, questions on client satisfaction or drop outs (if the sample design includes clients who have left the program). The advantage here is that a survey can obtain information on a large number of clients. Qualitative methods can probe in detail questions related not only to client satisfaction, but the appropriateness of product design and service delivery mechanisms in relation to client economic goals and opportunities, how they manage household and individual resources, and how they deal with risk in their day to day lives.

**2. Who Should Be Included in the Study?**

A quantitative IA ought to have a comparison group, preferably a quasi-experimental design control group, while a qualitative IA may or may not. A control group lends strength to an IA by comparing microentrepreneurs with and without program services. This is especially important if the primary objective is to justify program expenditures and document program effectiveness. It also is valuable to those who seek to expand and deepen program outreach to microentrepreneurs who are not in the program, and to those interested in understanding the secondary effects of the program.

Table 4: TRADE OFFS IN CHOICE OF WHOM TO INCLUDE IN AN IMPACT ASSESSMENT SURVEY SAMPLE		
	<i>Pros</i>	<i>Cons</i>
1. Quasi-experimental design control group	<ul style="list-style-type: none"> <li>- Enables IA to plausibly associate changes to program participation by comparing those with and without services</li> </ul>	<ul style="list-style-type: none"> <li>- Increases costs</li> <li>- Increases time needed</li> <li>- Requires knowledge of sampling and data analysis techniques</li> </ul>
2. Cross section of current clients by years in the program	<ul style="list-style-type: none"> <li>- Easy to find respondents</li> <li>- Facilitates capturing change based on length in the program and intensity of participation</li> </ul>	<ul style="list-style-type: none"> <li>- By itself does not make a strong case for program impact</li> <li>- By not taking into account program leavers, this approach is likely to over-estimate impacts</li> </ul>
3. Clients who have left the program	<ul style="list-style-type: none"> <li>- Measures the full impact of program</li> <li>- Improves credibility by precluding over- and under- estimation of impacts</li> <li>- Yields useful information for improving products and services</li> <li>- Reveals reasons for departure</li> <li>- Provides useful information for expanding and deepening outreach</li> </ul>	<ul style="list-style-type: none"> <li>- Difficult to locate those who leave, so it is more expensive and time consuming</li> <li>- Program leavers may be reticent to participate</li> </ul>

Low-end middle-range IAs that use a retrospective design often include only a control group and clients who are currently in the program. They do not include individuals who entered the program at the same time as the sample of clients, but have subsequently left the program. In comparison, in most IAs that use a longitudinal design, everyone interviewed in the first round is re-interviewed in the subsequent rounds, so the data on current clients can be analyzed together or separate from data on clients who have subsequently left the program. Including program leavers is likely to take more time than involving only current clients since the former tend to be more difficult to locate. In longitudinal assessments, a 20 percent non-find rate in the second round is common. Nevertheless, the data can be very powerful.

Data on those who leave the program aid managers in understanding what distinguishes the recurrent clients from those who have departed in terms of impacts, socio-economic and other characteristics, and views on program satisfaction. Data on impacts and program satisfaction can help program managers to determine the reasons for departure. Furthermore, information on key characteristics that distinguish leavers from those who remain in the program can help program managers predict who remains and who departs the program based on the products and services received.

Including program leavers is especially important for programs with high drop out rates. When it is not feasible to include them in the IA survey, data on dropouts should be obtained from the MIS system, small quantitative surveys and/or focus groups. At a minimum, reports on program impact ought to a) state the dropout rate by loan cycle, and b) the percentage of those that did not leave voluntarily. Preferably, more information would be provided on reasons for program departure based on systematic data collection among those who have departed.

### 3. What is the Estimated Sample Size for the Study?

After the initial selection of a method or mix of methods and the categories of respondents to include in the study, the next step is to estimate how many individuals to cover under each component of the study. This will drive the timeframe and budget for the study. For survey methods, planners should estimate the sample size, including the quasi-experimental design control group (for more details see section on sample selection under guidelines for quantitative methods). For focus groups and participatory rapid appraisals, planners should estimate the number of groups and sessions. For case studies and individual interviews, planners should estimate the number of individuals to be interviewed (see section on participant selection under guidelines for qualitative methods).

Estimating the sample size during the planning phase is important for developing a realistic budget. The number of people to be covered and the number of focus or participatory rapid appraisal groups will affect the cost of the assessment.

### 4. Where Should the Study Take Place?

If one wants to generalize from the findings to the total portfolio, field sites for impact assessments should be selected to represent the institution's total geographic portfolio. Sites should be in places where the microfinance institution is well established and where staff are likely to be helpful to the impact assessment team, while still representing the program's total portfolio and activities (Gaile, Duursma, and Eturu 1999).

In developing a budget, transport costs for interviewers/facilitators is of utmost importance. When fielding interview teams, in most circumstances it is advisable to make transport available to deliver the interviewers to particular sites each day. This helps to ensure that they get to the site and facilitates supervision. In cases where program staff are used, their usual mode of transport should suffice.

It is best to carry out interviews in a setting where the participant feels comfortable and there is enough privacy. For individual interviews, this often means at the person's house or enterprise. However, if these venues do not allow for privacy, a place nearby can be used. For focus group and participatory rapid appraisal sessions, the location should be convenient for the participants and suitable for holding a discussion. When the program involves group meetings, these can be used to set up appointments with individuals and obtain information on where the participant lives and works. When mutually convenient, some individual interviews may be conducted before and after the meeting in a place that allows for privacy.



5. Should the Study Be Based on a Longitudinal or Retrospective Design?

The primary options are a longitudinal or retrospective design. It should be noted that a design based on a cross-sectional analysis is not discussed since it is not valid to draw conclusions about impact by using information on a single point in time for a cross section of clients and non-clients without the use of sophisticated analytic techniques like structural equation models (GAO 1991).

CHALLENGE OF IMPLEMENTING A  
LONGITUDINAL DESIGN

If there is a relatively high program departure rate, and the majority of the proposed sample resides in urban areas in rented dwellings, then many in the baseline sample may be difficult to locate later (a 20-25 percent non-find rate is considered normal for longitudinal studies). This can be addressed by having no more than one additional data collection round a year after the baseline, and by asking respondents to provide contact information on someone who will know their whereabouts a year later.

In general, collecting data at more than one point in time yields more reliable information than depending on recall over an extended period, irrespective of whether a quantitative or qualitative approach is used. Quantitative data from two or more points in time are important for measuring or estimating change more reliably. This is especially true for measuring change in areas where recall is weak, or if attitudes, opinions and behaviors are likely to change over time. For example, recall data on income, regular expenditures (e.g., on food) or self-esteem are not very reliable, especially when using a long reference period. Other questions lend themselves to more reliable recall, for example, questions on children’s education, key asset purchases, or investments in housing or land. If a longitudinal design is not used, these types of questions should be given preference in the quantitative assessment. If a longitudinal design is used for variables about regular expenditures or income, the recall period should be short, such as last month or last week, to yield more reliable data.

The effects of seasonality must be taken into account, especially in a longitudinal design. This is critical due to the uneven flow of enterprise sales revenue, and household expenditures. Weather patterns associated with agricultural activities, school fee payment schedule, and special holidays are examples of seasonal factors. Conducting the follow up round(s) of interviews in the same months as the baseline is a low-cost way of controlling for seasonality.

A longitudinal design is most appropriate when the program is relatively stable in terms of its organization, management, and methodology. When a program (or the offices/branches to be studied) is in flux, the results of an IA are likely to have little utility when the clientele, products, and outreach process have changed. Under such circumstances, using a longitudinal approach is probably not cost effective.

The decision about using a longitudinal or one-time retrospective design also depends largely on a) funds available and b) the timeliness of the results in relation to their intended use. If the results will be used to inform a major decision within the next 12 months, then a retrospective design should be employed (table 5). However, if the results of the assessment will be used to improve the program or justify expenditures over a period of two or more years, then a longitudinal design is appropriate.

Table 5: TRADE OFFS IN CHOICE OF LONGITUDINAL OR RETROSPECTIVE DESIGN		
	<i>Pros</i>	<i>Cons</i>
Data collection at two or more points in time (longitudinal)	<ul style="list-style-type: none"> <li>- Data more reliable</li> <li>- Sample will include people who have left the program</li> <li>- Makes sense when program is relatively stable</li> </ul>	<ul style="list-style-type: none"> <li>- Impact data not available until second round of data collection</li> <li>- More expensive</li> <li>- Chance of losing a portion of sample</li> </ul>
Data collection at one point in time using recall data (retrospective)	<ul style="list-style-type: none"> <li>- Results available on a timely basis</li> <li>- Less expensive</li> <li>- Less time consuming</li> <li>- Less likely to lose a portion of the sample over time</li> <li>- More cost effective if the program is in flux</li> </ul>	<ul style="list-style-type: none"> <li>- Recall data on many variables are weak especially for long reference periods and when attitudes and behaviors are likely to change over time</li> <li>- Requires use of less precise measures on most variables</li> </ul>

C. Other Decisions

1. What is the Estimated Level of Effort and Timeframe for the Assessment?

In planning an impact assessment, it is important to consider how much time it will take to carry out each of the IA stages, over what period of time the work will take place, and the number of people anticipated to be involved at each stage. A realistic projection is critical: how much time is likely to be required for participation by key stakeholders? How much time will be required for the design, training, implementation, and data analysis stages? Does the tentative schedule cover critical periods (e.g. rainy or planting season, holidays, or high labor demand months) when it will be difficult to locate and interview microentrepreneurs? Does the projected schedule conflict with peak demand periods on the time of program managers and staff? It also is important to consider the time required to plan and manage the overall effort and write up and disseminate the findings. ***A frequent error is to underestimate the amount of time for the a) design, b) instrument development, testing and refinement, c) analysis and d) report preparation. A rule of thumb is that the data collection phase accounts for roughly one-third of the time required for a quantitative study and usually less than this for a qualitative study component.***

For each component, it is important to estimate the number of interviews or focus group sessions that will be carried out and how much time will be required for each. It should be noted that for focus groups and participatory sessions, time is required for organizing them and planning logistics. If data will be gathered at two or more points in time, it is important to consider the time interval between the assessments, the additional time required for the other rounds of data collection, and time for analysis of changes between data collection rounds.

2. Who are the Most Appropriate and Competent People to Direct and Carry Out the Assessment?

The core team for an impact assessment should offer a combination of complementary skills. Main skills should include field experience, financial institution expertise, quantitative (statistical) or qualitative methodology expertise, computer skills and local knowledge, including languages. The person or persons responsible for a particular component should have the appropriate skills and experience in using the method and be good at analyzing the information. A serious gap in any of these skills likely will lead to problems. Durability, flexibility, and patience are required personal traits of the core team. The IA action plan or terms of reference should clearly spell out individual duties while allowing for flexibility (Gaile, Duursma, and Eturu 1999).

When the primary objective of the impact assessment is to improve the program, its personnel should be active in planning and directing the IA and to the extent possible carrying out all the stages. When key essential skills are not available in-house, the program may seek an outside facilitator to train people in the steps and secure other expertise as required. When the primary objective of the impact assessment is justification of the program, in order to increase the perception of objectivity, it is best that individuals outside the program be the main persons involved in the IA implementation and report writing stages (table 6).

Table 6: TRADE OFFS IN INVOLVEMENT OF MFI STAFF AND MANAGERS IN AN IMPACT ASSESSMENT		
	<i>Pros</i>	<i>Cons</i>
Planning and design stage	- Irrespective of IA objective(s) involvement of staff and management increases the IA's credibility, usefulness, and cost effectiveness.	- If staff lack expertise in IA, it's important to involve outsiders with appropriate skills to advise or lead design effort.
Data collection, analysis, and report writing stage	- If the IA objective is improving the program, involvement of staff and management increases its usefulness and cost effectiveness. - It's important that staff do not know the people they are interviewing.	- If objective is justifying investments, findings are less credible, useful and cost effective. - Staff involvement likely to introduce bias in client responses and in interpretation of results. - Intensive training may be required. - Takes time away from regular duties.

Irrespective of the configuration, managers of the implementing MFI ought to take an active role in planning and designing the IA. Program staff should be involved in developing and testing the questionnaire, providing information on the program, and arranging for implementation. Program managers and staff ought to be involved in interpreting the IA findings

and disseminating the report. When the implementing MFI is taking the lead, the IA plan should specify the role of the different groups within the MFI and other key stakeholders. When individuals in another organization take the lead, it is incumbent on them to communicate with the implementing MFI about their role and inputs, and provide opportunities to strengthen local capacity.

When program personnel are to be used in data collection, a guiding principle is that the staff member should not know or be familiar with those s/he interviews. For small programs, the ***guiding principle is that loan officers not interview their own clients***. Adherence to these principles helps to avoid biased results. If respondents know the interviewer, they are likely to say what they think s/he wants to hear or what s/he has told them previously. Also, respondents become confused if the interviewer asks questions they already know the answer to. An exception to this principle would be interviewing clients to better understand impact paths in order to develop a survey instrument. ***Another guiding principle is that in cultural contexts with patriarchal rules that inhibit women from talking freely with men, women interviewers should be used.***

### 3. What is the Estimated Budget and Availability of Funds?

Based on initial decisions about the method or method(s) to be used, the level of effort, the personnel requirements and logistics, a budget estimate should be made. The plan then should be revisited to see if it fits the budget (or vice versa). If necessary, adjustments or other options should be considered to ensure that the plan and the budget are a good fit. Adjusting the number of geographic locations covered, sample groups, focus groups, participatory community sessions, or case studies, is often the best option for fine-tuning a tight budget.

There are different ways to think about the budget. One way is to weigh costs against benefits in relation to the IA's objectives. For example, if the objective includes improving programs, then the benefits of using the IA data to improve management, lower operating costs, or expand the market need to be factored in. If accountability to sponsors or a supporting institution is the only IA objective, then the level of funding may be a function of their willingness to pay.

In some circumstances, it is probably best to start with a lower end middle-range approach with a modest budget. For example, this approach may be appropriate when the intent is to build local capacity or if there is no previous experience to draw upon (such as previous IA work with this or similar organizations within the same context, no client based MIS system, or the MFI is not mature). As experience is gained, more can be accomplished at a lower per unit cost.

## V.

# Guidelines for Qualitative Components of Microfinance Impact Assessments

---

### A. Purpose of Qualitative Approaches

Qualitative methods permit one to understand key relationships and how different aspects of life are interrelated. They can be used to illuminate relationships between the client, the client's household, enterprise(s) and community. These methods facilitate the examination of topics in depth and are particularly useful to:

- facilitate a better understanding of the impact paths and processes (why and how changes occur);
- permit a client perspective of changes due to program participation and valuation of these;
- illuminate intangible changes such as in gender relationships, self-esteem and confidence;
- reveal unanticipated consequences of the program;
- help clarify hard to interpret findings from a survey; and
- allow clients to examine, explain and learn from their experience.

*An understanding of the impact paths and processes* allows MFIs and their supporters to better know ways clients are using loans, in particular, why decisions are made about the use of loans and enterprise net revenue, and how these funds are allocated. This information can be used to inform the development of appropriate questions and measures for a survey questionnaire. It enables one to a) identify the key hypotheses, variables, and measures to be covered in a survey, and b) focus on complex phenomena to better understand how to formulate survey questions and to determine the most appropriate response categories to capture this phenomena. An example would be to explore and understand the most common types of financial crisis among clients in order to draw up a list of most likely responses to this question in a survey instrument.

*A client perspective on program impacts and valuation of these* can feed into development of a survey questionnaire, be used to help to interpret survey findings, or be the cornerstone of an assessment. Qualitative approaches facilitate participants' explanation of their own situation. These may also be used to illuminate intangible or subtle changes that are difficult to capture with a survey. For example, for a very poor person the ability to occasionally buy new, rather than used, clothes can be a source of great pride but the small incremental amount of net income earned as a result of the loan may appear insignificant or not be captured in a survey. Also, changes in gender relationships, such as decision-making within the household, are extremely difficult to capture in a questionnaire since these involve dynamics that are not straightforward.

Unlike surveys whereby all questions are predetermined, qualitative methods permit a greater interaction and flow of information enabling the interviewer/facilitator to follow-up on information shared by the participant. The format allows for participants to discuss their perceptions and behaviors as well as issues that are important to them. As a result, these techniques permit the unveiling of unanticipated positive and negative consequences of program participation, and factors contributing to and impeding impacts.

Qualitative methods may include focus group discussions, case studies, or participatory self-learning. Focus groups involve asking a few key questions to generate discussion that yields information on the views and opinions of participants. It works best when the group is relatively homogeneous to allow conclusions to be based on common characteristics. Case studies involve detailed studies of either a group or individuals. An interviewer usually asks a set of open ended questions based on an interview guide. Histories can trace events from prior to joining the program to the current period.

Participatory self-learning (PLA) normally involves small groups, rather than individuals, discussing and analyzing information about their own lives and communities from their perspective and local knowledge base.<sup>6</sup> A facilitator usually uses tools focused on particular topics to stimulate discussion and records the information. A participatory self-learning approach may involve clients carrying out a self-assessment of changes that have occurred and identifying changes linked to program participation. It may also be used for clients to track key changes and to analyze them. These changes may be at the individual, enterprise or household level, or within their loan group or community. The changes are usually discussed in terms of trends rather than absolute amounts. The method is particularly appropriate for programs with an empowerment objective. Also, a participatory self-learning approach is particularly suitable for sharing of information between the program and clients, with a view to identifying ways the program might be improved to increase retention rates, and expand and deepen outreach. A key principle of PLA is triangulation of results from different individuals, groups or locations, or the use of different methods. This enables one to compare and verify the results (Shah 1999).

## **B. Design Stage**

### **1. Revisit the Plan and Budget**

After studying the key research questions, the stated purpose for collecting qualitative data, and the qualitative approach specified in the plan, the person responsible for the design may determine that another qualitative method would yield better information given the intended purpose and the key research questions. (See discussion below on specific approaches.) If so, the designer should set forth a justification for changing methods and obtain the approval of the person(s) that developed the terms of reference. Also, the designer of the study may request further elaboration of the key questions and the reasons why these were selected.

---

<sup>6</sup>The participatory self-learning approach is an outgrowth of participatory rural appraisals (PRA) that involve local people in sharing and analyzing their knowledge and conditions. The PRA discussion often centers on use of tools such as group mapping and wealth ranking, and is used to involve local people in planning project or program interventions. A feeling that the term PRA is too restrictive with its focus on "rural" and "appraisal" led to the PLA label for basically the same fundamental approach. It keeps the emphasis on triangulating results through learning from different groups or locations, or the use of different methods to compare and verify results.

The person designing the qualitative study should also consider the budget for the work to be done. If the budget is fixed, then the work to be done needs to match it or there needs to be a discussion about the trade-offs between the qualitative and quantitative components and, if necessary, funds shifted. Budgetary considerations are extremely important if the designer has to develop new tools for the study.

**2. Involve the IA Planners and Other Key Stakeholders**

Throughout the design stage, the person requesting the study and other key stakeholders should be involved at critical junctures. When revisiting the plan, clarification might be sought and adjustments proposed and justified. At the end of the design phase, the requestor of the IA and the program manager or representative (if different than the requestor) should provide review comments on the design, and the design adjusted accordingly.

When the designer of the qualitative component is not the same as the key designer(s) of the quantitative component, there must be close collaboration to assure that the components are synchronized to best address the key questions. When the qualitative component is to address impact questions, the designer should be involved in the identification of the conceptual framework, hypothesis and variables developed under the quantitative component.

**3. Identify Existing Tools or Design New Ones**

EXAMPLES OF QUALITATIVE TOOLS

Planning, Training, Conduct and Analysis of Focus Groups on Client Satisfaction: Individual Interviews on Loan Use  
(SEEP Network 2000)

Purpose and Conduct of Participatory Group Appraisals Using Seasonality Tool, Life Cycle Profile, Time Series of Asset Ownership, and Wealth Ranking  
(Wright, *et al.* March 1999)

Working within the IA budget, the designer(s) of the qualitative component should decide whether to use existing tools or to develop new ones to fit the purpose and key questions of the qualitative component. (See for example, SEEP 2000 and Wright et al. 1999). The existing tools cover focus group discussions, an interview guide on loan use, and participatory guides on wealth ranking, seasonal calendars, trend analysis and livelihood analysis (SEEP Network 2000, Davies 1996, and Shah 1999). Photos have also been used successfully to stimulate discussion of certain topics.

When a participatory self-learning approach is used, the tools should stimulate and allow for critical analysis among participants. Success will depend on the relevance of the issues to the lives of the participants. For example, a tool might be a livelihood chart that stimulates participants to discuss behavior, decisions, coping strategies, and ways these have changed and why. Another type of tool might be a series of photos to stimulate discussion about changes and factors contributing to and inhibiting change.

If existing tools are not applicable, the development of new tools should follow the principles below:

- Select key question to be answered.
- Develop tools based on purpose of the PLA and key question to be answered.
- Identify ways in which the information will be validated.
- Prepare probing question guide.
- Pre-test the tool and modify as needed.

**4. Develop a Plan for Selecting Participants**

The designer(s) of the qualitative component should develop a plan and rationale for selecting participants. The basis for selecting participants is a key step in establishing the objectivity of the results and thus needs to be documented. Selection decisions will revolve around different considerations depending on the method used and purpose of the study (table 7). To assist in choices, the designer of the qualitative component ought to discuss options with field staff and review program records. What is the distribution of clients by gender, sector, and location? A quick review of a random selection of client loan files or loan applications should provide the designer with insight into certain characteristics of the clients so that s/he obtains a sense of what appears to be the norm and what are outliers. This step is crucial whether or not the designer belongs to the program organization.

TABLE 7: OPTIONS FOR SELECTING INDIVIDUALS FOR QUALITATIVE STUDIES	
Selection Basis	When To Use and Why
Convenience	When expediency is important and the data are to be used to inform survey design.
Bracketing (selecting people at the extremes)	When the aim is to know what is happening at the extremes and what explains the differences. Extreme heterogeneity.
Best cases	When the aim is to understand what accounts for an effective program.
Worst cases	When the aim is to understand negative effects of program or factors impinging on the program having positive impacts.
Cluster	To determine how different program branches or clients with different products compare with each other.
Typical	When the aim is to know what is happening and why among typical clients.
Representative	To determine important differences between individuals chosen to represent important variations.

Adapted from Table 2.5 in GAO 1990.

**Focus Groups:** A focus group is a carefully guided discussion to obtain information on a limited number of key questions (3-4) through the synergy of social interaction and discourse among the participants. A focus group should consist of between seven to ten participants, and a facilitator.

Successful focus group discussions depend on equality and trust between group participants (Gosling and Edwards 1995). First, one must decide on the “target group” or who



should be involved. Next, decisions need to be made about the composition of the group: people who do not know each other, a mix of sub-groups who know one another, or only those who know one another. The sensitivity of the topic, the heterogeneity of members in loan groups, and local social practices should help guide the decision. In many cases it is culturally unacceptable for friends to express differing viewpoints in the presence of others, and for women to articulate views that differ from those of men, or even to actively participate in a discussion on an equal basis with men. In such cultural situations, a focus group should not include both men and women.

If the purpose of a focus group discussion is to help interpret survey data, and the key information sought is not sensitive, one might decide that participants need not know one another. When in doubt, seek the advice of loan officers and pilot test the group dynamics by holding a couple of focus groups using different bases for selection of participants. In principle the number of focus group discussions should depend on the point when little new information is provided. However, in practice, the study's purpose and factors such as time and budget will most likely influence the number.

The actual selection of the individuals or loan groups from which to select individuals should follow the same selection procedures as for case studies (below). It should be noted that focus group discussions do not necessarily need to be held only with clients, especially when the primary objective is program improvement. Focus group discussions with program leavers may be useful to highlight factors associated with departing. Furthermore, focus groups with non-client microentrepreneurs from the same catchment areas as the clients may provide very valuable information on a) why they self-exclude, b) group dynamics, in cases where the group selects its members, that lead to exclusion, or c) positive and negative impacts of the program on non-participants.

**Case Studies.** The cases should be selected by purposive sampling in which the investigator carefully chooses the cases to strengthen the validity of the data. The cases should be selected in order to gain either literal replication or theoretical replication (Yin 1994). Literal replication occurs when more than one case points to similar results. The ability to replicate similar findings with two or three cases strengthens the credibility of the findings. Theoretical replication occurs when two cases point to contrasting results, but the differences between the cases are predictable and based on the underlying theoretical framework of the study (Dunn, *et al.* 1998).

To determine a priori which individuals meet the criteria for selecting the cases, certain key information about the client population is needed. Information about key characteristics can be gleaned (table 8) from program files, loan officers, or the survey results. Examples of general categories that might be used to select the cases are: gender, number of years in the program, enterprise sector, and loan product or role of enterprise in the household economy (i.e. only source or primary sources versus complementary).

Once the criteria are agreed upon with the key stakeholders, the next step is to determine a non-biased process for selecting the cases meeting the criteria. One way to do this is to identify several individuals meeting the criteria and then to randomly select among the names. The criteria and process used should be carefully documented.

	Client the sole income earner in the household	Household has multiple income earners
Clients in their second year of program participation	1 female 1 male	1 female 1 male
Clients in their third or greater year of program participation	1 female 1 male	1 female 1 male
TOTAL CASES = 8	Males = 4; Females = 4	
	2 <sup>nd</sup> year clients = 4; 3+ year clients = 4	
	Client the only household income earner= 4 More than one household income earner = 4	

**Individual Open-Ended Interviews.** Case studies are one type of individual, open-ended interview. Individual interviews also may be carried out in a more informal manner when the intent is to inform the development of a questionnaire. For these interviews a cross-section of clients should be interviewed. The number of interviews will depend on the intended use of the information and budget.

**Participatory Self-Learning.** The purpose for which the information is gathered should guide the techniques to be used, the number of persons covered, and the method used to select those to be included. In general, a stratified, random sample is preferred when this approach is to form the core of the impact assessment. When it is not intended to cover a large sample of clients, the selection may follow the same process as suggested for case studies. Or, when the intent is to inform questionnaire development or interpret survey findings the selection may be based on convenience. A number of groups, locations or individuals should be specified that will form the basis for comparison and validation of findings, especially when another validation method is not used.

## 5. Identify and Select the Most Appropriate and Competent Individuals to Do the Work

In addition to decisions about the number of individuals or groups to cover, several related decisions also need to be made. For example, how many interviewers or facilitators are needed? Who should do the training? Should special recorders be used? Who should analyze the findings and write up the presentation? The human resource decisions need to take into account: the timing of the work; language ability; the importance of consistency across

individuals involved; requisite interviewing/facilitation skills; prior experience; and use of the information. Familiarity with the program services and methods will be required in varying degrees depending on the method to be employed. In addition, in certain cultures gender considerations need to be taken into account. The use of the information should serve as a guide when considering trade-offs.

In general the interviewers and facilitators must be good at establishing rapport, be good listeners, understand the underlying purposes, be open to receiving unanticipated responses, know how to ask unbiased questions and be good at probing and following up on information obtained. In addition, the facilitator's role is to make sure participants understand the questions, keep the discussion from straying too far from the topic, and help ensure that each person participates and no one dominates the discussion or influences its outcome. As noted by Hulme (1997), the participatory learning approach requires the highest level of skills.

Determining who are the most appropriate and competent people to carry out the work depends to a large extent on the objective of the impact assessment, and the intended use of the qualitative data. It is best to use individuals with the necessary skills that are involved in the program when the main objective is program improvement, but care needs to be given so that they are not interviewing or facilitating sessions with clients whom they know. When the qualitative component is part of an overall participatory approach to learning, then program personnel should be involved. If the main objective is primarily to justify the program by determining its impacts and the qualitative data are an essential element of the IA, then the qualitative work is best conducted by individuals not associated with the program in order to enhance the perception of objectivity.

## **6. Develop a Data Analysis Plan**

After developing the individual or focus group interview instruments or participatory tools and selecting the individuals to carry out the work, the designer of the qualitative component needs to develop a plan for specifying how the information from each qualitative approach will be recorded, organized, and analyzed. To the extent possible, the information recorded should be the actual words or phrases of respondents and follow the actual sequence of the discussion. There should be an individual record for each interview, focus group or participatory learning session.

The analysis plan ought to include a framework for organizing the information from the specific records following a format that reflects the underlying purpose of the qualitative study. It serves to summarize information across a number of interviews, focus groups or participatory sessions. When more than one person is involved in carrying out the work, each interviewer/facilitator should make their own summary and then the group or person responsible for the qualitative study should be responsible for developing the master summary (an example is provided in Section 5D below).<sup>7</sup>

---

<sup>7</sup> For more information on analysis of qualitative data see the section on loan use in SEEP 2000.

**C. Implementation**

**1. Finalize Logistics**

Implementation should begin with finalizing logistics such as transport and materials. Especially important for focus groups is determining where they will be held and the best approach for inviting the participants. The following factors should be considered in determining the setting for focus groups:

- a setting that provides privacy for participants;
- a location where there are no distractions;
- a non-threatening environment;
- a location that is easily accessible for respondents; and
- a place where seating can be arranged to encourage involvement and interaction.

For individual interviews the factors listed above are also relevant; in most cases, the home of the microentrepreneur meets the above requirements. An appointment should be made for a time that is mutually convenient.

Irrespective of method, contact should be made ahead of time. A written protocol should be followed when making contact (see below). It should clearly explain the purpose of the study, the way the data are to be used, confidentiality of individual responses, and the amount of time required. Individuals should be told that they may refuse to participate or refuse to answer any questions or discontinue their participation at any time without prejudicing their status with the MFI.

An important factor in qualitative studies is the recording of the information. When the person facilitating a focus group or conducting a case study is charged with recording the information, it usually interferes with the smooth flow of the session. Tape recorders are often used, but end up presenting problems because the conversation is not audible, and if transcribed, time consuming. An option is to have an individual serve as the recorder. For focus groups, a low cost option is to have the key information recorded on flip charts rather than attempting to capture all that is said.

**2. Plan and Conduct the Training Session**

PROTOCOL
<p>It is of utmost importance that the following be explained when persons are invited to participate and again at the beginning of the focus group or interview:</p> <ul style="list-style-type: none"> <li>- the intended use of the data;</li> <li>- the confidentiality of the individual level information;</li> <li>- the responses not affecting their relationship with the program; and</li> <li>- the freedom to refuse to participate, or not reply to certain questions.</li> </ul>

The training session needs to be tailored, based on a) previous experience of the individuals to be involved, b) familiarity of trainees with the program, and c) the task. In general, the training should cover the following:

- the purpose of the study and intended use of the results;
- practice on introducing the study and explaining participants' rights (the protocol);
- instructions about un-biased probing;

- review of good interviewing or facilitation techniques, with emphasis on avoiding verbal and non-verbal behaviors that would bias responses;
- instructions about discussing and recording information in an unbiased manner;
- mock learning sessions on carrying out the assignment that provide feedback on listening skills, encouraging openness, and probing questions; and
- pilot testing the process, instruments, and data recording as necessary.

**3. Review and Write Up the Information Gathered**

The person responsible for the qualitative study should be involved in supervising at least the first interviews or sessions conducted. The purpose is to provide immediate feedback to the individuals conducting and recording these. Care should be taken to ensure that the persons are not biased in recording the information. It is all too easy for someone external or internal to the organization studied to dismiss information with which they do not agree.

Review of notes and write up should take place within 12 hours of any interview or session, while it is still fresh in the memory of the interviewer/facilitator or recorder. The review should involve making sure that the information is accurate. Also, the record should include comments about the context and situation that might have bearing on the interpretation of the results. For example, did one person try to dominate the focus group discussion or during a case study interview did other people enter the interview setting who might have influenced the responses? If more than one person from the assessment team was involved in a session, all should review and discuss the write-up to make sure it is thorough and accurate.

COMMON WEAKNESS WITH  
QUALITATIVE STUDIES

Often a description of the findings is given but the information is not analyzed to highlight patterns and common themes and deviations from the patterns are not explained.

**D. Analysis**

The analysis of the qualitative information ought to begin with the initial plan. The information gathered should be analyzed to address the key questions and reasons for the study. Qualitative modes of data analysis provide ways of examining, comparing and contrasting, and interpreting meaningful patterns or themes. The meaningfulness is determined by the objectives of the study.

In contrast to quantitative analysis, qualitative analysis is guided by words and there are few universal rules and standardized procedures. Nevertheless, good qualitative analysis is both systematic and disciplined. A well-developed plan of analysis assists with the organization of the resultant information. In the process of carrying out the study, analysis should occur simultaneously; that is, the information provided should stimulate additional questions and new connections should be unearthed (Frechtling and Sharp 1997).

The mass of information has to be organized and meaningfully reduced by selecting, focusing and/or simplifying it (table 9). Normally one wants to look for:

- patterns and common themes on specific items;
- deviations from the patterns and factors that might explain these; and/or
- interesting stories that help illuminate the broader study questions (Frechtling and Sharp 1997).

**TABLE 9: EXAMPLE OF A MATRIX FOR ORGANIZING DATA ON LOAN USE**

Respondent group	Loan one			Loan two		
	Use	Why	Results	Use	Why	Results
Female clients: individual loans						
Female clients: group loans						
Male clients: individual loans						
Male clients: group loans						

Formal data analysis may begin with summarizing the information from each individual or focus group. Thereafter, by studying the results these can be further reduced and presented in another matrix to indicate patterns and deviations, with key reasons or characteristics noted for the deviations. The extent to which the findings lend support to a particular theoretical or practical proposition also may be noted. These matrices however tend to obscure the richness of the information. Therefore, quotes or interesting stories that help underscore the key study questions should also be extracted from the original data documentation.

In most cases, the analysis should involve drawing conclusions. This means stepping back to consider the meaning of the information and its implications in relation to the key questions. When conclusions are drawn, validity centers on whether the data are credible, defensible, and able to withstand alternative explanations. The latter is extremely important: if changes occurred, were they the result of program participation? Care should be taken to make sure that the conclusions do not go beyond what the data reasonably warrant. In other cases, the information is used to illustrate interpretations and conclusions drawn from a complementary survey. In such cases, the qualitative data informs the quantitative analysis.

**E. Dissemination**

How the information is presented and disseminated depends on the purpose of the qualitative study. In instances where the information is intended to illuminate survey data, the information needs to be integrated into the impact assessment report. This means fully integrated, not as stand alone sections of the report. The report’s explanation of the methodology used should include a description of the selection process and the key questions asked.

When qualitative information is the key component of an impact assessment, it is important to not only present the findings but to draw conclusions or inferences from them without generalizing to a larger population. A major weakness tends to be presenting only descriptive data without linking them to hypotheses or key questions investigated.

See section VI-D-2 on Presentation of Information and Dissemination for a fuller discussion of this topic.

# VI.

## Guidelines for Quantitative Components of Microfinance Impact Assessments

---

### A. Purpose of Quantitative Approaches

Quantitative methods, by definition, focus on numbers rather than words. They are useful for addressing questions related to what, whom, where, how many, and how much, and can be used to measure the incidence and prevalence of a phenomena. They can yield representative and broadly generalizable information about a particular population. Depending on the sampling technique and sample size, the findings can apply to a population beyond the individuals involved in the study.

Quantitative methods are appropriate when:

- the objective of the IA is to attribute changes to participation in the program (make a strong case for plausible association);
- a program affects large numbers of clients;
- the purpose of the assessment requires estimates of program impacts across a broad spectrum of clients, and thus seeks to be able to generalize the findings to a population larger than those interviewed;
- the assessment seeks to document who the program reaches, especially in terms of poverty level or other program objectives;
- statistical comparisons are required between groups or geographic locations;
- program delivery and implementation mechanisms are operating well enough to justify investment in a quantitative assessment (there is reason to believe the program has had a positive impact);
- the program is engaged in action research or pilot testing new products or outreach strategies; or
- there is some understanding of likely impacts based on previous qualitative work, or previous studies.

Quantitative methods are not appropriate in cases when the program has a small number of clients. Nor are they necessarily appropriate when managers are concerned mainly with MFI performance or client perspectives on the program or changes in their lives.

An advantage of quantitative methods is that they can cover large numbers of respondents. Another is the ease of data standardization, aggregation, and synthesis. Quantitative methods are particularly useful in isolating non-program causes of change. And while all data are affected to some extent by the perceptions and beliefs of investigators and data collectors, quantitative surveys and techniques can reduce bias in the results.



Increasingly, quantitative impact assessment methods are based on a *quasi-experimental design*. This design estimates the impacts of a program by measuring changes that have taken place in its clients, and isolates the effects of other factors that might have contributed to changes by using a control group. Changes in the client group minus those in the control group reveal changes that can be attributed to the program participation. Quantitative data are collected primarily through sample surveys. Guidelines for designing a quantitative survey to study microfinance impacts are presented below.

## B. Design Stage

### 1. Set Out a Conceptual Framework

#### PURPOSE OF A CONCEPTUAL FRAMEWORK

It sets out a pattern by which program participation is expected to lead to certain types of impacts. If the expectations are unrealistic given the length of program participation, loan size and other factors,

The first step in designing a quantitative survey is to conceptualize the *impact chain* to be examined. It should specify the *unit(s) of analysis* to be assessed (e.g., household, individual, enterprise, community) and the *types of impacts* to be studied (Hulme 1997). Whether the conceptual framework is explicit or implicit in the study design, it provides a base for framing specific research questions related to the impact of

microfinance services, developing related hypotheses, and identifying priority variables for study. The framework also should be used in the interpretation and presentation of the findings.

Previous impact assessments provide useful conceptualizations of impact chains that can be used, adapted, or refined in future studies. Some are more complicated than others, but basically provide a foundation for conceptualizing a unit (or units) of analysis, behaviors, practices and relationships, moderating processes, expected impacts, and the role of microfinance interventions in contributing to the desired changes. This information needs to be complemented by a basic understanding of the way in which clients are likely to use money in relation to their local environment, their socio-economic status and gender (FOCAS 1999). When such information is not available a priori, qualitative studies can be used to explore these issues and help to refine the framework.

### 2. Select Appropriate Hypotheses

A statement of hypothesis that is linked to the key research questions helps to focus data collection and analysis. Hypotheses often are stated as: *participation in a microfinance program is related to Y (e.g., increase in household assets)*. Ideally, a conceptual framework about the impact process should underlay the selection of hypotheses. Some of the hypotheses are embedded in theory, others in practice. But at a minimum, they should be grounded in an understanding of the program, the clients, and the broader context in which the microfinance program operates.

In selecting hypotheses to study, it is important to consider what is likely to change in relation to the type and magnitude of program inputs received (which may vary among clients).

For example, impacts on health knowledge and practice are more likely to occur in microfinance programs with health education components. Similarly, client level impacts are likely to vary according to the length of time the client has been in the MF program or the total amount borrowed or saved.

One option is to develop hypotheses to test whether or not the MFI is meeting its objectives or mission. In such instances, the hypotheses should be based on the program's stated purpose, objectives, and/or mission. An example is provided in the box below. The most appropriate approach for selecting hypotheses will rest with the objective of the IA.

There is always a trade-off between breadth and depth in selecting the number of hypotheses. There is a tendency to want to study and learn as much as possible about impacts and not to miss anything. However, in a world where resources are limited for impact assessments, trying to cover too many impacts in one assessment does not usually pay off. *Overly ambitious studies are difficult and resource intensive and often produce less-credible and less-useful results.* One crucial consideration is the complexity (and associated cost) involved in testing a particular hypothesis. Other important considerations are: the demonstrated validity in previous impact assessments; the likely impact patterns given clientele and local environment; and whether or not the IA should center on

EXAMPLE: TURNING MFI PROGRAM  
OBJECTIVES INTO HYPOTHESES

KMBI's loan and savings services result in increased social and economic empowerment of female clients, including women from poor households.

KMBI's loan and savings services lead to the participation of microentrepreneurs from poor households.

(Adapted from FOCUS March 1999)

objectives of the program and its supporters. *In general, a small and focused set of hypotheses can lead to more credible, useful, and cost effective impact assessments.*

### 3. Select Valid Impact and Moderating Variables<sup>8</sup>

There are almost an infinite number of impact variables that can be used. In deciding on what variables to include in an assessment, it is important to establish that they are linked to hypotheses, they are defined with precision, and that they are measurable within the timeframe and budget of the assessment. The choice of variables should also consider their demonstrated validity in previous impact assessments.

Impact variables serve as 'markers' of positive changes that are anticipated as a result of participation in a microfinance program. Similar to selecting hypotheses, the selection of valid impact indicators must be grounded in the local context, and be realistic in terms of what impacts are likely to occur given the nature and extent of client-program interaction. Has the client received only credit or only savings services? Has the client received business services for a day or over an extended period of time? How many loans has the client received and what is the total amount? In selecting variables, impact assessors must be realistic in terms of the amount of time

<sup>8</sup> The information in this section also applies to the identification of expected program performance results that reflect hypotheses based on the program's objectives.

it may take for certain impacts to occur, and be content with documenting positive steps along the way.

Moderating variables represent characteristics of the client, the client's household or enterprise, the program, or the context that influence change in the impact variables, but are unlikely to be affected by the MFI program. These may include, for example, the gender of the client, the household dependency ratio, the design of the loan product, or the level of inflation.

In choosing variables, it is important to consider the timeframe required for impacts to manifest themselves. Previous studies have shown that different variables show change at different times. For example, impacts on enterprise profits may occur early and then taper off within the first year or two of microfinance program participation. Other impacts, for example the accumulation of selected household assets, may take as long as three to five years of microfinance program participation to happen. One recent study concluded that social impacts (such as changes in women's mobility) are likely to take longer to occur than economic impacts (such as changes in income). Attention to temporal issues in measuring variables in impact assessments (either through longitudinal designs or through the use of recall data) is important for ensuring valid findings. Impact variables that have proven valid in previous assessments are provided in Appendix 3.

#### 4. Select Reliable Measures

Measures show the direction of change (positive or negative; increase or decrease), pattern of change (ordinal scale), or amount of change (interval measure) in a variable. Information on the direction of change can be used for almost any variable, but interval measures are generally more robust. However, they often are more complex to measure and are more subject to measurement errors.

GENERAL RULE ON  
RELIABILITY OF RECALL

People can recall extraordinary or large lump-sum income and expenditures more easily than regular, recurrent expenditures and irregular flows of income.

Microfinance impact assessments are likely to involve a mix of measures. The inherent nature of the impact variable and the complexity of measuring change in it should guide the definition of the measure. One way that a "middle-range" impact assessment differs from a more complex approach is that it estimates change, rather than measuring it more precisely. Such assessments are likely to involve a mix of measures, but to reduce costs and complexity (while allowing for reliable estimates of change), they probably will include more measures that indicate the direction and pattern of change than measures that indicate the amount of change. The selection of variables to measure the amount of change should be guided by the relevance of the variable and the budget and timeframe of the assessment. The variables also should be selected and framed to gain the most reliable recall data. For example, attention should be given to the reference period: one month, last season or last year.

Several other measurement issues also should be considered. One is the importance of distinguishing between the perception of change and actual change in questions. This distinction sometimes gets lost. Another is to consider the precision desired in choosing a particular measure: direction of change, pattern of change, or amount of change. As a rule of thumb, it is

easier to measure the direction and pattern of change than the amount. One should prioritize and consider trade-offs in deciding on the mix of measures to use and some variables are more difficult to measure than others. In choosing the mix of variables to include in an assessment, it is important to consider how difficult it will be to measure it in relation to the methods to be used, the skills required, and the budget.

## 5. Design the Sampling Plan

In quantitative surveys, the sample design is critically linked to the ability to draw conclusions about impacts. A sample design should provide a rationale for what groups of clients and non-clients to include in the sample, how many to include, how to select them, and the location of program study sites (and non-program study sites, if included). It should consider the best ways to include respondents with and without microfinance services and in measuring changes over time.

If the aim of the impact assessment is to understand the general impact of the program, the approach should be to select a random sample from the population. However, if the key research questions relate to comparing impacts among specific categories of clients, for example by gender, by loan product, or by geographic region, then these categories should be included in the design of the sampling plan. The more analytic categories included, the larger the required sample.

***Selecting clients:*** Clustered random sampling of clients is the preferred method of sampling for a low-to-mid range study. The first step usually is to select a sample of branch offices or towns in sites where the MFI is well-established. This process might involve: a) listing branches that have both new and repeat clients, and then randomly selecting a specified number of branches, or b) listing of branches together with their percent of total clients and purposively or randomly selecting branches that represent the majority of current clients. If the assessment's objective is to know about the entire program, the branches chosen should represent the MFI's total client population and geographic portfolio.

From here, clients can be selected randomly within these clusters from client lists. Or they can be selected by stratified sampling within each cluster, for example, by years in the program. The latter is a common practice since impacts should vary by the length of time the client has been in the program. If the IA is a one-time retrospective study, then attention ought to be given to inclusion of clients who have left the program but joined at the same time as some of the current client sample. Leaving out this group may result in overestimating or underestimating impacts. For example, clients who have exited the program may have graduated to receipt of credit from other organizations or at the other extreme, may have had to sell assets to pay-off their loan.

The ease of selecting clients relates to how the institution keeps track of its clients, including how clients are registered. For example, is it done manually or by computer? Are clients registered by business address, home address, or in some other way? If the client lists do not have addresses, what other practical ways can they be found (Gaile, Durrsma, and Eturu 1999)?

**Selecting a control group:** Best practice calls for the use of a comparison group of non-clients. Comparing the non-client group with the client group establishes that a microfinance program is plausibly associated with change or difference. Changes in the client group minus changes in the control group should reveal impacts associated with the program. A key challenge in selecting a comparison group is to ensure that the client and the control group are similar on key variables and to address the issue of self-selection, which can affect the validity of the findings.

The control group may consist of either pipeline clients or non-clients. Pipeline clients refer to those who have applied for credit but will not have received it by the time of the assessment. Use of pipeline clients helps to address the issue of self-selection in the selection of the control group because it indicates that the person is motivated to join the program and has met the initial if not the final criteria for joining the program. The use of pipeline clients is attractive for impact assessments, since they have potential of serving later as the client group, and hence providing good ‘before project’ data. Problems arise, however, when the pipeline group will not yield a large enough sample vis a vis the number in the client sample drawn from the same geographic area (or branch office). It also is not very suitable for longitudinal studies because the control group in the first round turns into a client group in the second round.

In general, if the budget and timeframe allow, it is preferable to use a non-client control group and to collect data on both groups at two points in time. In selecting the control group, individuals should be screened against a small set of key variables to establish the basis of comparability with program clients. Data on both groups at two or more points in time can provide a more reliable basis for assessing not only the differences between the two groups, but also the pattern and direction of change for both, thereby improving the plausibility of the impact findings.

The non-client sample should be selected on the basis of a few criteria so that they are similar to the client sample. The basic criteria should include same geographic location, enterprise sector and gender, and may also include easily obtained data that indicate that they would be eligible to be program participants. In most middle-range IAs, attention should be given to exclusion of microentrepreneurs who have received loans for their enterprise from other formal sources; by excluding them, the non-client sample represents a “non-treatment” group, that is, those without loans, to facilitate identification of the impact of MFI financial services.<sup>9</sup>

EXAMPLE: CRITERIA FOR SELECTING A NON-CLIENT  
CONTROL GROUP OF MICROENTREPRENEURS

A priori, match non-client microentrepreneurs to client microentrepreneurs according to the following factors.:

- Have a residence or business in the same geographic location or area as the clients;
- Have a microenterprise in the same sector as clients’ microenterprise; and
- Are the same gender as the clients.

Rather than a one to one match, the proportion of non-client microentrepreneurs with these characteristics should be in the same as the proportion of the client sample

<sup>9</sup> See sampling section of SEEP manual (2000) for further discussion of this method.

Examples of sampling designs commonly used for microfinance impact assessments are presented in Table 10.

TABLE 10: ALTERNATIVE SAMPLE DESIGNS FOR MIDDLE-RANGE MICROFINANCE IMPACT ASSESSMENTS		
	Advantages	Disadvantages
<i>Design 1 (longitudinal)</i> - Clients (T1 and T2) - Clients no longer in program (T2 only) - Non-clients (T1 and T2) - Non-clients who join the program (T2 only)	- Less reliance on recall - Second round includes persons no longer in the program - Second round allows for analysis of impacts by length of time in the program	- To control for seasonal variation, T2 needs to be done in the same season as T1 - Time delay in getting the analysis on impacts
<i>Design 2 (longitudinal)</i> - Clients (T1 and T2) -- Loans only -- Loan plus other services - Non-clients (T1 and T2)	- Same as above - Provides information about added impacts from other services	- Same as above - Development of questions to highlight differences between clients groups requires careful consideration
<i>Design 3 (retrospective)</i> - Clients (one time) -- Continuing clients - Clients (one time) -- Former clients - Non-clients (one time)	- Shorter turn around time from design to final report - Captures information on current and former clients to determine overall impact	- Depends on recall so restricted to more general trend or direction measures for most variables - Difficult to determine extent clients and non-clients were similar at the beginning of the reference period
<i>Design 4 (retrospective)</i> - Clients (one time) -- Continuing clients - Non-clients (one time)	- Shorter turn around time - Ease of sampling and locating clients - Non-client group helps to attribute changes between the groups to program participation	- Same as above - Likely to over- or under- estimate impacts since it does not include persons no longer with the program; Serious problem when dropout rates are more than 25 percent

*How many respondents should be included in the sample?* For quantitative surveys, the choice of *sample size* is critically linked to the ability to draw sound conclusions about the population sampled, not just those interviewed, and to establish plausible association. This is where limiting the scope of an impact assessment becomes important because the broader the range of issues covered, the more time required for data entry, cleaning, analysis, and report preparation. The general rule of thumb is that the sample size should be large enough to ensure effective use of control variables, allowing for at least 30 in any sub-sample of interest in the study and 50 is preferable. Each control variable used increases the minimum sample size that is required. The sample also should be large enough to allow for invalid data issues and to account for dropouts if the study is longitudinal. It should be small enough to fit the budget. In determining the sample size, trade-offs must be made between the margin of error, the confidence interval, and the budget.

	<i>Pros</i>	<i>Cons</i>
Larger sample size (n > 300)	<ul style="list-style-type: none"> <li>- Results more reliable</li> <li>- Scope of IA is wider because it can use more control variables</li> <li>- Can more effectively accommodate invalid data issues</li> <li>- Can account for dropouts if the study is longitudinal</li> </ul>	<ul style="list-style-type: none"> <li>- More expensive</li> <li>- More time consuming to collect data</li> <li>- Data may be more difficult to manipulate and work with</li> </ul>
Smaller sample size (n < 200)	<ul style="list-style-type: none"> <li>- Less expensive</li> <li>- Less time consuming</li> <li>- Data may be easier to manipulate and work with</li> </ul>	<ul style="list-style-type: none"> <li>- Results less reliable</li> <li>- Scope of IA is limited by the small number of control variables that can be used</li> <li>- Difficult to accommodate invalid data issues</li> <li>- Usefulness of data from a longitudinal IA is limited when there is a relatively high non-find/non-interview rate in the follow-on round(s).</li> </ul>

Determining the sample size can easily get complicated and ultimately rests on the forethought and judgment of the designer. Considering the examples in Table 10, it is probably safe to say that a total sample size of 300 or more can meet the information needs for most basic middle-range retrospective and longitudinal studies (designs 1 and 4) and a total sample size of 450 would be needed for Designs 2 and 3. This is a general rule of thumb that would probably need to be modified if sophisticated analyses are required.

*What sample selection method should be used?* Table 12 presents sampling options and guidance in choosing a method.

Selection approach	When to use and why
Simple	Each unit of the population (client or non-client) has an equal chance of being selected
Cluster	The population is divided into clusters, such as branch offices or towns, and then a random or purposive (based on specific criteria) sample is taken of the clusters
Stratified	Groups of special interest are separated and then a random or purposive sample taken within each group

## 6. Design the Questionnaire

The questionnaire ought to yield information that helps answer the key research questions and addresses the objectives of the impact assessment. It should be clear, not too long, and comprehensive. There is a trade off between brevity and comprehensiveness. The first step should be to clearly define a small set of focused hypotheses and variables and stick to them. Next, appropriate measures need to be identified, along with the recall reference period. Thereafter, attention needs to be given to crafting the questions. The questions must be written in a way to facilitate the respondents' ability to understand the question and to report the information to the best of their ability (GAO 1993).<sup>10</sup> Moreover, the questionnaire ought to be constructed so that the questions do not bias the responses. It should include some open ended questions to allow for issues to be brought forward that may have been missed in the structured part of the questionnaire. The questionnaire should be reviewed by the core team members and remain a working document until the last possible minute (Gaile, Durrsma, and Eтуру 1999). When time permits, key stakeholders ought to review and provide feedback on the questionnaire.

*The quality of the data in a middle range approach can be enhanced if the survey questions are sequenced effectively (sensitive questions at the end), the interview time is not too long (45 minutes maximum), the data collectors have the appropriate skills for the task (collecting information from case studies requires different skills than collecting information through structured questionnaires), and the total number of data collectors involved in the assessment is kept to an optimal minimum.* Closed-ended survey questions reduce errors and are easier to analyze. Good questionnaires also require translation into local languages and then reverse translation to cross check the accuracy of the translation. They further require standardized instructions for recording responses and protocols for introducing the study to respondents to avoid biased responses. The introduction should stress the confidentiality of individual responses and assure clients that there is no link between status in the program and responses.

Once the questionnaire is drafted, *it should be pre-tested* by a small core group including those involved in its development. The pre-test phase should entail interviewing several people, revising the questionnaire, and further interviewing and revising. The point of the pre-test is to help ensure:

- the questions are well sequenced with sensitive questions placed near the end of the questionnaire;
- the questions are clearly understood by respondents;
- the questions cover one thought at a time, rather than combine issues;
- the questions tend to be applicable to almost all of the respondents;
- the closed-ended response categories capture almost all of the responses, with only a few or no responses falling into the 'other' category;

---

<sup>10</sup> Free sources focused on questionnaire development can be located on the internet and a free copy of Developing and Using Questionnaires, by the US General Accounting Office can be ordered at US GAO, P.O. Box 6015, Gaithersburg, MD 20884-6015 or by fax at 301-258-4066.



- the questions do not require unreasonable effort to answer; and
- the questions can and will be answered as accurately as possible.

## 7. Develop an Initial Data Analysis Plan

After the sample is designed and the questionnaire developed, an initial data analysis plan should be developed. Thinking through the data analysis plan may lead the assessors to change their questionnaire and even their sampling plan. The analysis plan forces the assessors to decide what findings they do and do not need. The use of every question should be justified at this stage and non-essential questions eliminated (for example, we may not need to know Grandpa's education level). Also, development of the plan may reveal potential difficulties, such as how multiple response questions will be analyzed, which in turn, may lead to refinement of questions.

The plan should indicate the survey specific questions and variables that will be used to address the key research questions. It should specify the type of statistical analysis required. After the data are collected, the plan may be expanded to test different relationships and specify more complex analyses (this is discussed further in section VI-D). This also is a good time to identify combinations of quantitative and qualitative information to address key research questions. The initial data analysis plan can help to reveal where quantitative data may need to be backed up with qualitative.

## 8. Involve Program Management and Staff in the Design Process

Whether people inside or outside the program design the impact assessment, it is critical to involve key MFI management and staff in the design phase. They can substantially improve the credibility and ultimate usefulness of the assessment. Their knowledge of the program, the clients, how the clients interact with the program, and the context in which the clients and program operates can be invaluable in informing many aspects of the research design (hypotheses, variables, research sites, sampling categories, etc.). For example, if an individual from the program's headquarters is responsible for the design, this person should seek the involvement of key managers and staff of headquarters as well as personnel in the branch offices to be covered by the survey. To the extent feasible, the key persons who are expected to be the main users of the findings ought to be engaged in this process or provide review comments.

## 9. Re-Evaluate the Budget and Adjust Accordingly

At the end of the design process, the budget should be re-evaluated and it or the design adjusted to 'fit' while staying true to the objectives of the impact assessment.

## C. Implementation

### 1. Make Basic Logistical Arrangements

There is no standard practice when it comes to logistics. But planning ahead can save a lot of time, trouble, and added expense in conducting an IA. This is especially important in the context of a middle range approach with a tight timeframe and budget.

One important issue to consider is planning, scheduling and booking transport for the IA team and enumerators. The team should not have to rely on unscheduled transport unless it is absolutely necessary for flexibility's sake. Another practical issue is the need for careful consideration of the time needed in each field site to achieve the study's goals. Plan for an extra day if possible. It is also important to plan for a place to train enumerators. It should be in a place that offers privacy, a board or room for flip charts, and space for working in small groups.

When the survey is directed and implemented by non-program personnel, establishing a field 'headquarters' near to, but separate from, the microfinance institution being assessed is useful. For example, a small restaurant or hotel can provide a welcome meeting point with privacy for core team members and enumerators. It can provide space for training, tables for work, and food and drinks. If the field headquarters is close to where the core team members are staying, members can trade off between meeting the enumerators as they return from the field and entering data. It is also important to anticipate the need for supplies (paper, pens, clipboards), photocopying facilities, and power for computers and light.

Especially when the core IA team is composed of individuals who are not part of the program, it is always a good protocol for a team member to pay a courtesy call to local officials soon after arriving at a field site. Government officials are responsible for knowing about activities within their jurisdictions. Local officials also may be able to help identify enumerator candidates and sort out other logistics.

## 2. Use Personnel with the Appropriate Skills and Experience

The quality, training, and attitudes of those responsible for collecting the data are critical. Objectivity and skill on the part of enumerators are important for collecting credible quantitative data. It is good to have the gender of the enumerators roughly proportionate to the gender of the respondents. For example, if the sample is over 80 percent female, then ideally most of the enumerators should be female. However, quality and experience should not be sacrificed for gender balance, except in cases where culture seriously inhibits females talking freely to males whom they do not know.

When program staff are not used to collect data, good enumerator candidates can be identified through the local research community, government statistics offices that conduct household surveys, program staff, local secondary school teachers, and/or local officials. ***The selection of enumerators, irrespective of whether or not they are program staff, should be based on good communication skills, survey experience, ability to speak the local language, the ability to understand and write in English (or the language used in the assessment), and availability during the survey period and willingness to work outside normal hours. In addition, it is particularly helpful if the enumerator understands the key economic terms used in the questionnaire.***

Objectivity and skill on the part of the enumerators is important for collecting credible impact data. Using program staff introduces the risk of biased responses and of compromising

the validity and reliability of the data.<sup>11</sup> This is a particularly important issue when the primary use of the findings is to justify program expenditures and to solicit donor funding. On the other hand, if the primary use of the findings is to improve the program and the secondary use is justification of program expenditures, then the experience and insights gained during the collection of survey data can be used to help inform decisions for program improvement. This approach is particularly appropriate when the program is open to learning and improving its performance. In such cases, program staff may be used, but it is essential that no one interview their own clients or people they know. The latter principle is important for two reasons: interviewing someone you know is likely to bias responses and the respondent often is confused when the person asking the question already knows the answer.

### 3. Back Translate the Questionnaire

The need for careful translation of the questionnaire into the predominant local languages is very important. It is better to do this well ahead of time. It may entail choices when there are several languages spoken in one area. A good way to check on the accuracy of translation is through back-translation of the questionnaire into English. Where the back translation does not match the original English, there is a problem. *Back translation is inexpensive and not very time consuming and can save a lot of problems later.*

### 4. Develop Written Instructions as a Reference Guide for Enumerators

An instruction manual is extremely useful. In particular, it should clearly set forth the protocol for explaining the study and seeking participation. The protocol should highlight that individuals are to be invited to participate and are free to refuse.<sup>12</sup> Also it should promise that the information provided will be treated as confidential and will not affect their status with the MFI. The instructions guide should also serve as the basis for training the enumerators on the purpose of each question and how to deal with unusual cases. Also, it should be a reference guide when questions arise during the interview stage. This is especially important when those actually carrying out the fieldwork differ from those involved in the design of the questionnaire.

The instruction manual may need to be revised after the questionnaire is pilot tested. When additional instructions are given during the interview process, the supervisor should record these so they are available to the person analyzing and interpreting the data.

### 5. Train Those Conducting the Survey

Enumerator training should:

- review and practice best practices for conducting interviews and establishing rapport;
- discuss the purpose and objectives of the survey;

---

<sup>11</sup> In a 1993 survey of microfinance clients in two programs in Egypt, half of the enumerators were program staff and half of the enumerators were not. In analyzing the results, the research team established through statistical analysis that there was a clear bias in the responses given to the program staff.

<sup>12</sup> See the discussion of protocol in Section VC.

- review written protocol on how to introduce the study to respondents and role play how to introduce the survey and respond to questions about it;
- discuss the meaning of each question;
- role play in groups of three: one interviewer, one respondent, one observer to record and later explain problems with the interviewing techniques;
- provide practical experience by interviewing one or more microentrepreneurs who are not part of the study; and
- pilot test the questionnaire and at the end of each day discuss the difficulties encountered and unusual circumstances that require special instructions.

It is critical that enumerators understand the meaning of each question and how to record the answers to the questions. Trainers should make sure that enumerators clearly understand the terms and questions embedded in the questionnaire. Training should involve a review of the written instructions that serve as a reference guide for the enumerators. In addition, ethical issues should be discussed.

One training technique for a middle range approach is to begin by explaining the questionnaire and then having the enumerators conduct a few surveys in the vicinity of the field headquarters. Following this, they can review and discuss ambiguities and other difficulties with the core team members.

## 6. Pilot-test the Questionnaire

The questionnaire for collecting quantitative data should be pilot tested with both clients and non-clients who are not part of the sample, taking care not to pollute or influence the environment in which the actual survey will be conducted. The purpose is to ensure that the questions are clearly understood by data collectors and respondents, the questionnaire is introduced properly, and responses are recorded correctly.

A field test of the questionnaire can reveal problems with the questions and codes. It also may reveal questions that are inappropriate for the context (for example, if a large majority of business premises have electricity, it is probably not very useful to include a question on electricity at the business premise). A field test of the questionnaire also is a way of training enumerators. In a middle range approach, time and budget constraints mean little time between pre-testing and conducting the survey, so close monitoring and coaching of enumerators becomes very important.

## 7. Refine the Questionnaire and Instructions Manual

After completing the steps of back translation, enumerator training, and pilot-testing, the questionnaire and the instructions manual will probably need fine-tuning. Problematic questions should be changed or eliminated; new questions may be added; and changes in words or codes may be made. Questionnaire production should be done only after this point. The survey is now ready to go!

## 8. Supervise the Enumerators

Adequate supervision of the data collection process is important for reliable and valid data. In practice, it is not always possible to monitor enumerators directly during interviews if the survey is geographically dispersed or personnel resources are limited. An alternative is for the IA core team members or a supervisor to meet the enumerators on a daily basis and review their work with them (for example, go through one or two questionnaires to discuss possible problems). The core team members should provide feedback on their review of the previous day's questionnaires and any common problems identified. Also, this is a good time for the enumerators to share information on clients with interesting stories, important experiences, or certain issues that might be explored in more depth through qualitative methods.

## 9. Enter the Data

Data entry is often viewed as a tedious task that should be carried out by someone paid to do tedious tasks. In fact, data entry of a limited sample is an excellent way for core team members to identify problems before they become impossible to solve. If data entry begins immediately after the first day of survey work (or even better, after pilot testing the questionnaire), problems with the questionnaire and/or enumerators can be identified and rectified quickly.

There are a number of advantages if one or more core team members enter the data themselves. They can get a real 'feel' for it, which is very useful during the analysis stage. Involvement of core team members in data entry is an effective means for them to check the validity of the data and facilitates effective coding. It eliminates the need to train and monitor data entry personnel. A major drawback is that the first few days in the field are the busiest, and adding data entry to the workload of team members can be onerous. Alternatives include hiring a person experienced in data entry and analysis, and training one or more MFI staff in data entry.<sup>13</sup>

Most data entry programs are now standardized globally. The data entered on a major spreadsheet program, such as Excel, Lotus, Dbase, EPI, and others can be easily moved to another type of program. For example, data entered on EPI can be analyzed using SPSS or SAS. If at all possible, it is recommended that one of the "global" programs be used for ease of transferability. The quality and validity of any translation should be checked and included in the process when data are validated, cleaned, and the coding checked.

## 10. Document Problems Encountered During the Implementation Phase and Responses

The credibility of an IA can be improved by keeping track of problems encountered during the implementation phase and describing how they were dealt with. This makes the

---

<sup>13</sup> An important instruction is to enter everything written for the "other" category in closed-ended questions. This will permit the analyst to review the answers when there is a relatively high percentage in this category to determine if the answer really fits another specified category and to determine if another category ought to be established.

process more transparent and provides information that outsiders can use in judging credibility of a study. It also provides useful lessons for future impact assessments.

## D. Analysis

### 1. Analyze the Data

Systematic data analysis is a pivotal link in establishing the credibility of middle range IA. Good analysis starts with the quality and reliability of the data, the manageability of the data in terms of quantity and ease of manipulation, and whether sufficient resources for analysis (time, money and people) have been planned for. Much can be learned from previous assessments in terms of what types of questions did not work in generating the information needed for analysis. Documentation of these lessons during the analysis is useful for future impact assessors.

Quantitative survey data can be analyzed in many ways but should focus on the key research questions and hypotheses. The initial round should explore differences in averages and distributions for all variables by key analytic category (e.g., clients/non-clients). The next round should involve simple statistical tests (T-tests or chi-square tests) on initial run data showing differences in the characteristics of the sample and in the impact variables. If the data do not fit the assumptions or the hypotheses, the analysts should further explore what the data reveal and may identify issues that should be covered through qualitative studies such as focus groups or case study interviews. Particularly in such cases, subsequent analyses should look at clusters or disaggregations of interest (such as men/women; location A/location B; client poverty levels) to determine the extent to which impacts vary by group.

Certain areas of analysis are of more interest than others (for example, changes in assets) in a microfinance impact analysis. The analysis should explore these variables in greater depth relative to other variables in the study. Simple statistical tests like T tests (comparing means) or cross tabulations (or chi-squared tests comparing across categories) likely will yield useful information even if they are not statistically significant. They also can lay the groundwork for more complex multivariate testing using controls and comparisons or other more sophisticated tests in cases when such analysis is called for and the data lends itself to such analysis.

In middle range impact assessments, T tests and chi-square tests help to determine if the findings between clients and non-clients or other analytic categories are statistically significant. Statistically significant results between clients and non-clients make a plausible case for causation when the data compare the change that has occurred in the client and non-client samples between two time periods.

At this point it is appropriate to begin the process of integrating the quantitative findings with qualitative findings. When data from more than one approach addresses the same hypotheses or issue, then it should be triangulated. In some cases, the survey results may suggest issues to explore through additional qualitative interviews. Or the qualitative findings may suggest certain issues or variables to probe further in the quantitative analysis.

It is also a good time to get initial feedback from MFI program staff and clients on the preliminary findings. Involving program management and staff at certain stages in the analysis

process can help to promote co-ownership of the findings and provide them with immediate feedback.

## 2. Present and Interpret the Findings

Impact assessment reports should both present and interpret the findings and provide suggestions for program improvement. Clear presentation of survey data can be challenging. It is necessary to make choices about what data to present, what level of detail to provide, and the balance between data presentation and interpretation in the write up. The audience for middle range IAs is likely to be comprised largely of practitioners and development planners rather than academics and so it is important to define key terms and concepts and avoid overly technical terms.

At this stage, the report writers should review the key research questions, study hypotheses, and questions that address rival hypotheses and provide a profile of the sample. While the sample profile may be presented on its own, the other data should be organized to address the hypotheses and key research questions. In this way, the findings address and illuminate the phenomena studied by telling a story. The qualitative data and information on program features and context should be woven into the text to illuminate and interpret the survey findings. In instances when the qualitative data address additional questions, the information should be incorporated in the most appropriate location in the text.

To be user-friendly it is important to simplify the presentation by using a select number of tables, graphs, or charts. Quantitative findings should be presented in a way that makes it easy to see differences between clients and non-clients. For both groups, averages should be complemented with distributional data to provide a fuller picture of the finding and to identify possible outliers. Table 13 shows how survey findings could be presented.

It is challenging for one report to meet the needs of all stakeholders. For example, implementing organizations often want more detailed information than funding organizations. A best practice for addressing this is to have an annex that provides tables with both numbers and percentages to enable others to analyze and use the data. A description of the methodology should be summarized in the report, but explained in more detail in an annex.

Variable	Clients Average N=203	Non-Clients Avg. N=112
Composite impact index**	19.59	18.08
Change in amount of business**	3.37	2.88
Change in business income**	3.32	2.97
Change in sales	3.27	3.17
Change in enterprise assets**	3.33	3.09
Change in household assets**	3.35	3.18
Change in savings	3.03	2.83

\*\* Statistically significant difference at the 95% confidence level. Gaile, Duursma and Eturu 1999.

### 3. Document Problems Encountered, Responses, and Lessons for Future

Impact assessment reports should document the quantitative survey process, including initial plans, problems encountered, changes made, and lessons learned. This can enhance the transparency and credibility of an impact assessment, help future assessors avoid the same problems, and contribute to the evolution of better impact assessment methodologies. While it is not necessarily cost-effective to document the research process in all IAs, the inclusion of learning components in some impact assessments is important for the development of more efficient and effective impact assessment methodologies.

#### E. Dissemination

To be useful for decision-making, impact assessments must be disseminated effectively and in a timely manner to the intended audience. Dissemination strategies may include bullet point summaries, snappy presentations, and strategic cups of coffee (Hulme 1997). The time lag between data collection and presentation should be reduced to a minimum. While some assessors recommend a period not exceeding nine months, others argue for the advantages of more immediate feedback. In fact, having program staff directly involved in all of the stages is one way to ensure immediate feedback. In addition, an impact assessment linked to a larger program evaluation is likely to reach a wider audience.

General rules of thumb to follow in dissemination:

- develop a specific dissemination plan;
- leave time and resources for dissemination to the intended audience;
- disseminate information on the research process; and
- disseminate the report as soon as possible after data collection and analysis.



## VII.

# Checklist of Key Steps for Planning and Implementing a Microfinance Impact Assessment

---

### At the planning stage...

- ◆ P-1. Choose a *program or program branch that is relatively stable* in its strategy, operations and staff.
- ◆ P-2. Clearly state the *objectives and intended users* of the impact assessment. Ideally the objectives should include both justifying program investments and improving programs.
- ◆ P-3. Review the *availability of important background information* on the program and context to use in the impact assessment. During the planning stage, it is important to review information that establishes the extent to which the program operates effectively in its environment and is financially sustainable. The planning stage also should specify how the context information to be collected during the IA is likely to influence program outcomes (outreach, sustainability and impact) by affecting the operations of the program and the economic activities of clients.
- ◆ P-4. State the *key research questions*. They should be guided by the intended use of the information and limited to those most important to be addressed and fit with the resources available for the IA.
- ◆ P-5. Determine the *method or mix of methods* for the IA. The choice of methods should be based upon the objectives of the impact assessment, the key research questions, the availability of resources (time, money and people), and the degree of precision and generalizability required.
- ◆ P-6. Estimate and provide a rationale for the *sample size* based on the objectives, choice of methods, total number of clients, and degree of reliability required. This will be an important factor in determining the budget. In most cases, the survey sample should include clients currently in the program, clients who have left, and a quasi-experimental design control group.
- ◆ P-7. *Locate the sample* to represent the MFI's overall geographic portfolio. The cost of transport of the data collection team to sites is an important budget consideration.
- ◆ P-8. Decide whether to use a *longitudinal or retrospective design* for measuring change. Data gathered at at least two points in time generally yield more reliable information but longitudinal studies are more expensive than one-time surveys or qualitative interviews based on recall. Collecting data only once yields less reliable data on certain questions. If a retrospective design is used, emphasize direction of change questions and others that lend themselves to reliable recall (e.g., money spent on children's education and investments in

housing). If a longitudinal design is used, include recall questions on some quantitative variables, such as income and expenditures, but use relatively short reference periods.

- ◆ P-9. Estimate the *level of effort and timeframe* for the IA. Sufficient time should be included for each research method for the design, training of those involved, implementation, analysis of the data, interpretation and presentation of the results, and dissemination. The timeframe should consider the time interval between two data collection rounds (if a longitudinal assessment is planned) and seasonality issues, as well as the schedules and availability of MFI management and staff.
- ◆ P-10. Define the *key personnel needs and availability* for the IA. Skills of the core team members should be complementary and include field experience, financial institution experience, quantitative (statistical) or qualitative methodology expertise, computer skills, and local knowledge including language. If program staff are to be used to collect data, it is critical to use staff members who do not know the people they are interviewing.
- ◆ P-11. Develop a *terms of reference* for the study, including a budget, which reflects the above points. Involve critical intended users in the planning process and in developing the terms of reference.

---

## At the design stage

### ... for qualitative methods

- ◆ D-1. *Revisit the plan and budget.* Reconfirm or propose a different qualitative method(s) to address the purpose and key questions guiding the qualitative component.
- ◆ D-2 *Involve* the person(s) requesting the study, the key program managers and other key stakeholders intended to use the information, throughout the design stage.
- ◆ D-3 *Identify existing tools or develop new instruments* to elicit the information desired and fulfill the purpose of the qualitative component. When centering on impacts, draw on the conceptual framework and hypotheses and variables identified in the quantitative component. Work in conjunction with those designing the quantitative component.
- ◆ D-4 *Design a plan for selection of participants* taking into account geographic differences and document rationale. If interviews are to be held or a participatory self-learning approach is to be used, determine the number of total persons to be covered and their geographic dispersion. For focus groups, determine the number of sessions in each geographic area and the total number of groups to be held.
- ◆ D-5. *Identify and select the most appropriate and competent individuals to do the work.* The decision needs to take into account the timing of the work, language ability, importance of consistency across the interviews/focus groups/participatory sessions, interviewing/facilitation skills, and use of the information. Also, attention should be given to who should do the training, recording, and who should analyze and write up the data.

- ◆ D-6. Design a *data analysis plan*. Set out a format for initial analysis of the resultant information.

---

### At the design stage

#### .... for quantitative methods

- ◆ D-7. Conceptualize the *impact chain* to be studied including the *units of analysis* to be used and the *types* of impacts to be studied. This should provide a base for framing specific research questions, identifying hypotheses, and selecting impact variables to be addressed in the quantitative survey
- ◆ D-8. Select a small set of meaningful *hypotheses*. The hypotheses should be linked to the key research questions and the conceptualized impact chain.
- ◆ D-9. Select valid *impact and moderating variables*. These variables should be linked to the hypotheses, defined with precision, and be measurable within the timeframe and budget of the assessment. They should be grounded in the local context, be realistic in terms of what impacts are likely to occur given the nature and extent of client-program interaction and the amount of time it may take for a particular impact to occur. The use of variables with demonstrated validity from previous impact assessments should be considered.
- ◆ D-10. Select *measures* to reliably estimate change. Middle range IAs generally will typically include more measures that indicate the direction and pattern of change than measures that indicate the amount of change. The relevance and inherent characteristics of the variable, the budget, and the timeframe of the assessment should guide the selection of measures. Remember that to make claims about impact in a one-time survey, the data must represent changes over a specific recall period.
- ◆ D-11. Identify *information needs related to improving the program and studying rival hypotheses*. This may include, for example, the poverty level of clients, their perception of positive and negative factors associated with program participation, and former clients' reasons for leaving the program. Develop questions to elicit this information.
- ◆ D-12. The design of the *sampling plan* for a quantitative survey should include a rationale for:
  - which clients and which non-clients to include in the sample;
  - how many to include;
  - how to select respondents with and without microfinance services;
  - the location of program study sites (and non-program study sites if relevant); and
  - the best way to measure change over time.

It is also important to identify key analytic categories of interest (gender, poverty level, length of time in the program) to ensure that sufficient numbers of each group are included in the sample.

- ◆ D-13. The *questionnaire* should be *clear, easy to use, and not too long*. Also, questions should be stated in a way as to not bias the responses. In addition to *closed ended questions*, it should include a *few key open ended questions*, keeping in mind that closed ended questions reduce errors and are easier to analyze, while open ended questions allow for new or different issues to be brought forward. The questionnaire should be *translated* into the appropriate local language and *back-translated* to cross check the accuracy of the translation. *Standardized instructions* for recording responses and the *protocol for introducing the study* to respondents should accompany the questionnaire. The questionnaire should be *pre-tested* with several microentrepreneurs.
- ◆ D-14. An *initial data analysis plan* helps to ensure that all the data to be collected are necessary and useful. It should identify specific questions in the survey instrument and qualitative information that will be used to address each of the key research questions. The development of the plan may induce designers to reconsider and change the questionnaire design, the mix of methods, or the sampling plan. At this stage, every question in the survey questionnaire should be justified. If it is not justified, it should be eliminated.
- ◆ D-15. *Involving MFI managers and staff in the design process* is very important for improving the credibility and ultimate usefulness of an IA. Their knowledge of the program, the clients, how clients interact with the program, and the context can be invaluable in informing many aspects of the research design.
- ◆ D-16. The final stage in the design process should involve *re-evaluating the budget* and adjusting it and/or the design to fit.

---

## At the implementation stage

### ... for qualitative methods

- ◆ I-1. *Finalize logistics*. For focus groups and participatory learning group sessions, select a convenient site, taking into account the need for privacy, no distractions, a non-threatening environment, accessibility and suitability for small group work. Contact the selected microentrepreneurs and, following a written protocol, clearly explain the purpose of the study, the amount of time required, confidentiality of responses and the location. Then, invite the person to participate. Also, make arrangements for training interviewers/facilitators. Organize transport, materials and other items related to the successful conduct of the qualitative component.
- ◆ I-2. Plan and *conduct the training session* for those who will serve as facilitators, interviewers and/or recorders. Tailor the session to the trainees' previous experiences and their familiarity with the program. At a minimum, the training should cover purpose and use of the resultant information, review of good interviewing/facilitation techniques, protocol for introducing the study, key questions to be asked, and mock learning sessions. It should also cover instructions for writing up the resultant information.
- ◆ I-3. The *review of notes and writing up* should take place within 12 hours of any interview or group session, while the information is still fresh in the recorder's mind. This should

provide a written record using, to the extent possible, the exact words used by participants and the same sequence in which items were discussed.

---

## At the implementation stage

### ... for quantitative methods

- ◆ I-4. *Logistical planning* is very important for credible, timely, and cost effective IAs. Examples of things to think about ahead of time:
  - scheduling and booking transport for the IA core team and enumerators;
  - the amount of time needed in each field site;
  - a place to train the enumerators;
  - a place where core team members and enumerators can meet daily;
  - the need for supplies (paper, pens, clipboards) and photocopying facilities; and
  - the availability of electricity for computers and light in the field.
- ◆ I-5. The selection of *enumerators* should be based on their previous survey experience, knowledge of the local language, ability to understand and write in English (or the language used in the assessment), and their availability during the survey period and willingness to work outside normal hours. Their objectivity is important to collecting credible data. The gender of the enumerators should be roughly proportionate to the respondents, but quality and experience should not be sacrificed for gender balance.
- ◆ I-6. The *enumerators should be trained* to understand the purpose and objectives of the survey, how to introduce the survey to respondents, and the meaning of each question. They also should learn best practice interview techniques and how to record the answers.
- ◆ I-7. Following (or as part of) enumerator training, the questionnaire should be *pilot tested* with a small number of both clients and non-clients who are not part of the sample. Close monitoring and coaching of the enumerators by core team members during the pilot test helps to ensure that the questions are clearly understood by the data collectors and respondents, the survey is introduced properly, and responses are recorded properly. It may reveal problems with the questionnaire that require changes. After this, the questionnaire and the instruction manual can be *fine-tuned and, finally, photocopied for field use*.
- ◆ I-8. Adequate *supervision of enumerators* is critical for valid and reliable data. An IA core team member should meet with enumerators on a daily basis to review their work, discuss problems, and provide feedback on the previous day's questionnaires. This needs to be done irrespective of whether or not the enumerators are program staff.
- ◆ I-9. *Data entry* should begin as early as possible, ideally after the pilot test of the questionnaire. If a computer and electricity are available in the field, *data entry can start after the first day of survey work* and problems with the questionnaire and/or enumerators can be detected and rectified quickly. If data are being entered on more than one computer, it is important to establish that the files can be merged early in the process. *Involvement of core IA team members in data entry* is an effective way for them to check the validity of the data

and facilitate coding (assuming the questionnaire is a manageable length, the sample is a manageable size, and the core IA team members are computer literate). It also precludes the need to train and monitor data entry personnel.

- ◆ I-10. Documenting *problems encountered* during the implementation stage and how they were dealt with makes the process more transparent and provides a stronger basis for establishing the credibility of a study.

---

### Analyzing the data ...

- ◆ A-1. Good analysis requires a *data set that is manageable* in terms of quantity and ease of manipulation. It further requires sufficient resources (time, money and people) that have been planned for ahead of time.
- ◆ A-2. The *analysis of quantitative survey data* should focus on the key research questions and hypotheses and follow the initial data analysis plan. First rounds should explore descriptives, such as *averages, and frequency counts and distributions* for all variables by key analytic categories. The next round should involve *simple statistical tests* (T-tests or chi-square tests) on initial run data showing differences in the characteristics of the sample and in the impact variables. Subsequent analysis should look at clusters or disaggregations of interest (such as men/women; location A/location B; client poverty levels) to determine the extent to which impacts vary by group. If the quantitative survey data do not fit with the hypotheses, the analysis should proceed to explore what the data do reveal.
- ◆ A-4. The *qualitative analysis* should be guided by the initial plan, but additional information and analysis that would illuminate the underlying purpose of the qualitative component should be included. After summarizing the findings, the information should be organized and meaningfully reduced by selecting and focusing on:
  - patterns and common themes on specific items;
  - deviations from the patterns and factors that might explain these; and
  - interesting stories that help illuminate the broader study questions.

A more informal method can be employed when the qualitative data are used to develop a survey instrument.

- ◆ A-5. The analysis of the quantitative data is likely to reveal issues to explore further through qualitative methods. Similarly, the qualitative findings also may suggest certain questions or variables to probe further in the quantitative analysis. *Triangulating* the quantitative findings and the qualitative findings deepens understanding of impacts and helps to establish the reliability of the findings.

---

### Presenting and interpreting the findings...

- ◆ F-1. The presentation and interpretation of quantitative and qualitative IA findings should center on key research questions, study hypotheses, and questions that address rival hypotheses. It is important to define key terms and concepts and avoid the use of overly technical terms. Information on key features of the program and context should be woven in to illuminate and interpret the findings.
- ◆ F-2. In presenting and interpreting *quantitative findings*:
  - provide a profile of the sample;
  - organize the other data to address the hypotheses and key research questions;
  - simplify the presentation by using a select number of tables, graphs, or charts;
  - present the data clearly and accurately, and use averages complemented with distributional data to provide a fuller picture and identify outliers; and
  - make it easy to see differences between clients and non-clients.
- ◆ F-3. In presenting and interpreting *findings from qualitative components*, the purpose of the IA should guide the approach. In general, qualitative findings should be integrated as appropriate with the quantitative data to illuminate and explain quantitative findings, address questions that complement the quantitative findings, and provide a human face and tell a story. In some cases discussion of qualitative findings may dominate, with the quantitative data used to generalize these findings to a larger population. Boxes can be an effective way to highlight findings or cases that address a particular issue.
- ◆ F-4. In interpreting the findings, never draw conclusions without data to support them. Always *support conclusions with data*.
- ◆ F-5. At a minimum the IA *report should include*:
  - a clear statement of the objective of the impact assessment;
  - a description of the program including its objectives, products, methods, drop-out rates and financial data;
  - a description of the methodology used for qualitative and quantitative components, especially criteria and process for selecting participants;
  - findings, interpretation of results, conclusions, and when appropriate, recommendations for program improvement;
  - difficulties encountered, how these were addressed, and lessons learned; and
  - appendices with more detailed information.

---

### Dissemination . . . .

- ◆ S-1. Have a specific dissemination plan.
- ◆ S-2. Leave adequate time and resources for dissemination to the intended audience.

- ◆ S-3. Disseminate information on the results of the IA and on the assessment process.
- ◆ S-4. Disseminate findings as soon as possible after data collection and analysis.
- ◆ S-5. Consider the use of bullet point summaries, snappy presentations, and strategic cups of coffee.



## REFERENCES\*

- Chen, Martha Alter. 1997. A Guide for Assessing the Impact of Microenterprise Services at the Individual Level. AIMS Project. Management Systems International: Washington, D.C.
- Davis, Philip. 1996. Qualitative Tools and Microenterprise Impact Evaluation: Discussion Paper. AIMS Project. Management Systems International: Washington, D.C.
- Dunn, Elizabeth, William Matthews and J. Arbuckle. 1998. Assessing the Impacts of MicroEnterprise Services: Research Plan for Case Study and Focus Group Components of Core Impact Assessments. AIMS Project. Management Systems International: Washington, D.C.
- Edgcomb, Elaine L. 1996. PVO/NGO Tools for Household and Business Impact Assessment: Report of a Planning Meeting. Organized by the AIMS Project. Management Systems International: Washington, D.C.
- Frechtling, Joy and Laure Sharp. 1997. User-Friendly Handbook for Mixed Method Evaluations. Directorate for Education and Human Resources. National Science Foundation. www.ehr.gov.
- Gaile, Gary L. and Jennifer Foster. 1996. Review of Methodological Approaches to the Study of the Impact of Microenterprise Credit Programs. AIMS Project. Management Systems International: Washington, D.C.
- Gosling, Louise with Mike Edwards. 1995. Toolkits: A Practical Guide to Assessment, Monitoring, Review and Evaluation. Development Manual #5. Save the Children: London.
- GAO. November 1990. Case Study Evaluations. Program Evaluation and Methodology Division, U.S. General Accounting Office: Washington, D.C.
- GAO. May 1991. Designing Evaluations. Program Evaluation and Methodology Division, U.S. General Accounting Office: Washington, D.C.
- GAO. June 1992. Quantitative Data Analysis: An Introduction. Program Evaluation and Methodology Division, U.S. General Accounting Office: Washington, D.C.
- GAO. October 1993. Developing and Using Questionnaires. Program Evaluation and Methodology Division, U.S. General Accounting Office: Washington, D.C.
- Hyman, Eric and Kirk Dearden. 1996. A Review of Impact Information Systems of NGO Microfinance Programs. AIMS Project. Management Systems International: Washington, D.C.

- Hyman, Eric and Kirk Dearden. September 1998. "Tracking the Impact of Microenterprise Programs: The Experience of Non-governmental Organizations" IN Impact Assessment and Project Appraisal. Vol. 16 (3), p. 195 – 202.
- Rossi, Peter and Howard Freeman. 1989. Evaluation: A Systematic Approach:. SAGE Publications: Newbury Park, California.
- Sebstad, Jennefer, Catherine Neill, Carolyn Barnes with Gregory Chen. 1995. Assessing the Impacts of Microenterprise Interventions: A Framework for Analysis. USAID Managing for Results Working Paper No. 7. Center for Development Information and Evaluation, US Agency for International Development: Washington, D.C.
- SEEP Network. 2000. Learning from Clients: Assessment Tools for Microfinance Practitioners. AIMS Working Paper. AIMS Project. Management Systems International: Washington, D.C.
- Shah, Meera Kaul with Rose Zambezi and Mary Simasiku. 1999. Listening to Young Voices: Facilitating Participatory Appraisals on Reproductive Health with Adolescents. FOCUS Tool Series No: 1. CARE International: Lusaka.
- Snodgrass, Donald. 1997. Assessing the Effects of Program Characteristics and Program Context on the Impact of Microenterprise Services: A Guide for Practitioners. AIMS Project. Management Systems International: Washington, D.C.
- Yin, Robert K. 1994. Case Study Research: Design and Methods. Applied Social Research Methods Series, Vol. 5. SAGE Publications: Thousand Oaks, CA.

\* AIMS papers are available on the internet at [www.mip.org](http://www.mip.org) under the AIMS publications section.

## APPENDIX 1

### LIST OF DISCUSSION PAPERS AND HIGHLIGHTS OF MEETINGS PREPARED FOR CGAP IMPACT ASSESSMENT METHODOLOGIES WORKING GROUP, 1997-1999 \*

- Cohen, Monique and Gary Gaile. 1997. Highlights and Recommendations of the Virtual Meeting of the CGAP Working Group on Impact Assessment Methodologies April 7-19, 1997. AIMS Project. Management Systems International: Washington, D.C.
- Cohen, Monique and Gary Gaile. 1998 Highlights and Recommendations of the Virtual Meeting of the CGAP Working Group on Impact Assessment Methodologies April 14-28, 1999. Organized by the AIMS Project. Management Systems International: Washington, D.C.
- Hulme, David. 1997. Impact Assessment Methodologies for Microfinance: A Review. AIMS Project. Management Systems International: Washington, D.C.
- Sebstad, Jennefer. 1998 Towards Guidelines for Lower Cost Impact Assessment Methodologies for Microenterprise Programs. AIMS Project. Management Systems International: Washington, D.C.

\* AIMS papers are available on the internet at [www.mip.org](http://www.mip.org) under the AIMS publications section.

## APPENDIX 2

### LIST OF BACKGROUND PAPERS SUBMITTED FOR CGAP IMPACT ASSESSMENT METHODOLOGIES VIRTUAL MEETINGS, 1997-1999

- Adams, Mia. August 11, 1999. Letter to Carolyn Barnes commenting on the CGAP IA guidelines.
- Beroff, Renee Chao. 1997. "Impact Assessment As Seen From A NGO Practitioner's Scope." Paper submitted by CIDR for the 1997 CGAP Virtual Meeting on Impact Methodologies.
- Bousson, Pacal, Pierre Daubert, Nathalie Gauthier, Martin Parent, and Cecile Ziegler. 1997. The Micro-Economic Impact of Rural Credit in Cambodia. Collection Etudes et Travaux.
- Bruntrup, M. S.M. Alauddin, A. Huda, M. Rahman. 1997. Impact Assessment of the Association for Social Advancement (ASA). Volume 1 Main Report.
- Chua, Ronald T. 1997. Impact of Access to Credit on the Poor: A Research Design and Baseline Study. Foundation for Development Cooperation: Brisbane.
- Chua, Ronald T. 1998. The Performance and Sustainability of Two Philippine Microfinance Institutions: An Assessment. Foundation for Development Cooperation: Brisbane.
- Edgcomb, Elaine L. and Carter Garber. 1998. Practitioner-Led Impact Assessment: A Test in Honduras. AIMS Project. Management Systems International: Washington, D.C.
- Espegren, Oddvar. 1997. Microfinance Impact Assessment: A Dynamic Dialogue Approach. M.A. Dissertation submitted to the School of Development Studies of the University of East Anglia.
- Feinstein, Osvaldo N. 1997. "A Transactions-Cost Approach for Microfinance Project Assessment." Paper sponsored by IFAD for the 1997 CGAP Virtual Meeting on Impact Methodologies.
- Focus Pty Ltd. 1999. Philippines Impact Assessment of Kalibalikat par sa Maunlad na Buhay (KMBI). Prepared for AusAID.
- Gaile, Gary, Marjan Duursma and Laurenti Eturu. 1999. Centenary Rural Development Bank Ltd. Microenterprise Impact Assessment Final Report. HIVOS.
- Gheen, William. 1998. Review and Summary of Consultant's Report, Survey Results, Ecuador. Inter-American Development Bank.
- Holvoet, Nathalie. 1998. Intrahousehold Resource Allocation Processes and Their Outcomes: Impact of Credit and Women Groups. Case Study for Rural North Tamil Nadu. General Outlook and Design of the Case Study. College voor de Ontwikkelingslanden, RUCA.

- Husain, A.M. Muazzam. 1997. Poverty Alleviation and Empowerment. An Impact Assessment Study on BRAC's Rural Development Programme (IAS-II). BRAC Research and Evaluation Division: Dhaka.
- LaRoy, Pascale, 1999. GRET Impact Assessment Methodology An Abstract and Some Comments on CGAP IA Guidelines (draft document). Groupe de Recherche et d'échanges Technologiques (GRET): Paris.
- Little, Peter. 1997. "Income and Assets as Impact Indicators." Paper submitted by USAID for the 1997 CGAP Virtual Meeting on Impact Methodologies. AIMS Project. Management Systems International: Washington, D.C.
- Mayoux, Linda. "Impact Assessment and Women's Empowerment in Micro-Finance Programs: Issues for a Participatory Action and Learning Approach." Paper submitted by ODA (DFID) for the 1997 CGAP Virtual Meeting on Impact Methodologies.
- McNelly, Barbara and Karen Lippold in collaboration with Ayele Foly and Robin Kipke. 1998. Practitioner-led Impact Assessment: A Test in Mali. AIMS Project. Management Systems International: Washington, D.C.
- Mosley, Paul. "The Use of Control Groups in Impact Assessment for Microfinance." Paper submitted by ILO for the 1997 CGAP Virtual Meeting on Impact Methodologies.
- Ouattara, Korotoumou, Genevieve Thi-Dieu-Phuong Nguyen, Claudio Gonzalez-Vega, and Douglas H. Graham. 1997. The Caisses Villageoises D'épargne et de Credit Autogerees in the Dogon Region of Mali: Elements of Impact. Centre International de Developpement et de Recherche and Ohio State University.
- Simanowitz, Anton. 1999. "Understanding Impact Experiences and Lessons form the Small Enterprise Foundation's Poverty Alleviation Programme, Tshomisano." Prepared for the CGAP Working Group on Impact Assessment Methodologies. Small Enterprise Foundation.
- Sultana, Monawar and Ashok Nigam. 1999. Impact Assessment Study of the Family Development Fund, Egypt. UNICEF Staff Working Papers, Evaluation, Policy and Planning Series, Number EPP-EVL-99-003: New York.
- Wiig, Arne. March 1997. "Credit Expansion in Microcredit Programs: Dilemmas and Feasible Methods for Studying Them." Paper submitted by the Norwegian Ministry of Foreign Affairs for the 1997 CGAP Virtual Meeting on Impact Methodologies.
- Wright, Graham, Deborah Kasente, Germina Ssemogerere and Leonard Mutessasira. 1999. Vulnerability, Risks, Assets and Empowerment the Impact of Microfinance on Poverty Alleviation. Contribution to World Development Report 2000/1. MicroSave Africa and Uganda Women's Finance Trust: Kampala.

Yaron, Jacob, McDonald Benjamin, and Gerda Pipek. February 1997. "Rural Finance: Issues, Design and Best Practices." Paper submitted by KFW for the 1997 CGAP Virtual Meeting on Impact Methodologies.

## APPENDIX 3

### EXAMPLES OF HYPOTHESES AND VARIABLES WITH DEMONSTRATED VALIDITY IN PREVIOUS IMPACT STUDIES <sup>14</sup>

---

*“Participation in microfinance programs leads to increases in household welfare”*

#### **Related variables**

- Household income (level and sources)
- Household assets (e.g., contextually defined assets that indicate improved economic status, such as refrigerators or transport vehicles)
- Expenditure patterns reflecting more investments and expenditures that improve the quality of life such as food expenditures (contextually defined)
- Children’s education (e.g., increased expenditures on education; increase in the proportion of school age children in school)
- Seasonal fluctuations in per capita food expenditures
- Household’s effectiveness in coping with shocks (types of shocks and coping strategies)
- Borrowing, saving and lending patterns (changes in relation to contextually defined borrowing and savings systems; in relation to contextually defined production, investment, and expenditure patterns)
- Practices related to non-financial program services (e.g., health, nutrition, or family planning practices)
- Poverty gap (difference between household poverty level and poverty line)

---

*“Participation in microfinance programs leads to increased enterprise stability and growth”*

#### **Related variables**

- Microenterprise profits (reflected in marketing margins, sales value, volume of sales)
- Microenterprise fixed assets, especially among repeat borrowers
- Paid and unpaid employment generated by the microenterprise
- Business practices (including use of technology)

---

<sup>14</sup> Many of these hypotheses and variables were identified in the AIMS research plan for its three core impact assessments. They were identified on the basis of their demonstrated validity in previous studies, their cross cutting nature, and their meaningfulness in understanding the contribution of microfinance programs to household economic welfare, enterprise stability and growth, and individual empowerment.

*“Participation in microfinance programs leads to increased individual empowerment”*

**Related variables**

- Personal income
- Client’s control over use of money they earn (over purchases)
- Assets owned and controlled by client
- Pattern of savings towards more formal channels
- Self-confidence and self-esteem for women (improved self-image; increased capacity to manage specific aspects of the enterprise; extent to which the respondent values her own contribution to the household; the extent to which others value the respondents capacity and abilities)
- Increased mobility, especially for women
- Participation in decision-making

**EXAMPLES OF MODERATING VARIABLES WITH DEMONSTRATED VALIDITY IN PREVIOUS IMPACT ASSESSMENTS**

**Household level**

- Socioeconomic status
- Poverty level
- Dependency ratio (workers to dependents in household)
- Initial endowment of household (e.g., assets)
- Gender of head

**Enterprise level**

- Initial endowment of enterprise (contextually defined)
- Gender of owner
- Location of enterprise (urban or rural)
- Type of enterprise
- Increased labor supply in enterprise

**Individual level**

- Gender
- Control over loan

**Program**

- Program methodology
- Loan size
- Number of loans
- Repayment cycle (weekly vs. monthly)



- Term of loans (short term, medium term)
- Length of membership in program
- Client use of the loan (e.g., for working capital, enterprise investments, household investments, household expenditures)
- Client satisfaction with program

**Context**

- Inflation
- Natural disasters

## ANNEX 4

### MFI PERFORMANCE INDICATORS AND RELATED IMPACT INFORMATION

MFI INDICATOR	RELATED IMPACT INFORMATION	HOW COMBINED INFORMATION COULD BE USED TO ENHANCE PERFORMANCE
<b>Outreach and Effectiveness</b>		
Disbursements — number of borrowers reached by program	by target group characteristics (e.g., gender, socio-economic status, location, microenterprise type, etc.)	to assess effectiveness in reaching the intended target group; to develop strategies for improving outreach to target group
Number of borrowers currently active	By target group characteristics	to assess effectiveness in reaching target group; to identify strategies for improving outreach to target group
Average outstanding loan size	By target group characteristics	to assess effectiveness in reaching target group with appropriate loan size; to identify sub-groups likely to benefit more (previous studies suggest a relationship between larger loan size and positive impact)
Percent women borrowers	by other target group characteristics	to assess effectiveness in reaching women in all categories of the target group; to identify strategies for improving outreach to women
Number of savings accounts	by target group characteristics	to assess effectiveness of savings services in reaching the target group; to identify strategies for improving savings services for the target group
Average amount of savings	by target group characteristics	to assess effectiveness of program in mobilizing savings of the target groups; to identify strategies for improving the design of savings services to increase the volume of savings
<b>Effectiveness</b>		
Re-loan rate	by target group and loan characteristics	to better understand who repeat borrowers are; to identify groups that may be benefiting more through repeat loans; to identify strategies for expanding the volume of lending by reducing turnover of clients
Dropout rate	by target group characteristics by loan size by number of loans by reasons cited for exiting program	to better understand who is dropping out and why; to identify strategies for retaining more clients
Repayment rate Arrears by age Defaults	by target group characteristics by loan size by number of loans	to better understand the characteristics of clients experiencing repayment problems and why they may be having problems; to identify strategies for improving portfolio quality
Credit and deposit flows	by local economic cycles by target group characteristics	to better understand the role of credit and savings in smoothing investments, income, and consumption (i.e., managing risks)

MFI INDICATOR	RELATED IMPACT INFORMATION	HOW COMBINED INFORMATION COULD BE USED TO ENHANCE PERFORMANCE
<b>Viability</b>		
Costs per amount lent Cost per new loan made	by target group characteristics	to pinpoint potential problem areas; to reassess service delivery strategies for various groups; to identify strategies for reducing service delivery costs
<b>Sustainability</b>		
SOURCES OF FUNDS Grants and donations	impact information related to the socio-economic investment objectives of donors	to justify, encourage investments
SOURCES OF FUNDS Members' capital	impact information related to the socio-economic investment objectives of members	to justify, encourage investments

Sebstad 1998, Table 3.