Working Paper No. 7

Impact Assessment of Microfinance:

Towards a New Protocol for Collection and Analysis of Qualitative Data

June 2002

James Copestake, Susan Johnson & Katie Wright, University of Bath

CONTENTS

Impact Assessment Of Microfinance: Towards a New Protocol for Collection and Analysis of Qualitative Data

James Copestake, Susan Johnson & Katie Wright,
Centre for Development Studies, University of Bath, UK

1. Introduction

Section One sets the context by discussing the nature of demand for and supply of information on the impact of microfinance. On the demand side, it makes a distinction between demand from within MFOs themselves for organisational development, and from donors and regulators for public policy purposes. On the supply side, it argues that there is a case for more use of rigorous qualitative methods that occupy an intermediate position between "positivist/quantitative" and "participatory/interpretive" approaches. The *Imp-Act* programme is presented as an appropriate network within which a protocol to popularise the use of such an approach might be developed.

Section Two further clarifies the distinction between positivist/quantitative, participatory/interpretive and more rigorous qualitative methods. To illustrate the nature of the latter it presents brief case studies of how such tools have been used by the authors in Kenya, Zambia, Malawi and Peru. Section Three then takes a further step by suggesting what a standard protocol for a more rigorous qualitative impact protocol (referred for convenience as *QUIP*) might look like. Section Four concludes by emphasising the key reasons why *QUIP* might meet unsatisfied demand, then outlining the steps required for its future development.

2. The context - microfinance impact assessment

2.1 Demand for impact assessment information

Microfinance services contribute to development by adding value: benefits to society should exceed costs. Most of the costs and benefits are borne by the providers and users of the services, but there may also be important side effects, or what economists call externalities, on other people and society at large. These are referred to below as wider impact. The day-to-day business of microfinance is primarily concerned with direct impacts. Here the key equation is whether costs of provision are less than the benefits (net of transaction costs) to clients. If they are then business is possible, with benefits shared between provider and user according to the price that is struck between them. The scope for business can be enhanced by lowering the cost of provision of the services, as well as by developing new products that better serve the needs of clients.

This takes us to the first rationale for seeking to understand the impact of services on users. The better is the understanding of users – their livelihoods and relationships, values and norms, particularly in relation to

management of money – the more likely it is that providers will be able to improve the quality of the services they can offer.

But there is a second reason why impact assessment may be important. This is that where the indirect impact of microfinance is large and positive then there may be a case for or public subsidy of microfinance services. Likewise if there are significant negative effects then there may be a case for taxing them. Either way, policy makers need information on the nature and magnitude of wider impact.

The literature on this topic tends to divide into two, each corresponding to the two rationales for impact assessment outlined above, and the distinct information and decision-making systems corresponding to each. This distinction is illustrated in Table 1. The first places more emphasis on how to convince more detached decision-makers in public policy domain. The second is concerned with management of microfinance organisations. The two columns may also be viewed to some extent as a historical progression. For as the so-called new wave of microfinance has spread and deepened, so has the emphasis on IA for public policy given way to a preoccupation with IA for internal organisational learning.

Table 1: Contrasting sources of demand for impact assessment

	Impact assessment for public policy	**Impact assessment for product development**
Who for?	Donor	Service provider
When? (Historical context)	Donor led agenda of promoting increased provision of microfinance where supply of services is deficient (weak competition).	Organisation led agenda of seeking to cut costs and improve customer loyalty in order to survive (strong competition).
What for?	Is public money better used (at the margin) to subsidise microfinance services or for other purposes?	What can be learnt from clients in order to improve the prospects for organisational survival?
Outcomes	Green/red light for replication of "best practice" models (e.g. village banking, solidarity groups….) without much differentiation of users.	Product differentiation in response to diverse local demands of different categories of user.

Perhaps the most important rule of impact assessment is to be clear about why it is necessary, and for whom; hence this distinction is useful. But an overemphasis on such a simplistic classification of IA may also be dangerous. More specifically, it is easy to associate each source of demand for IA with distinct methodologies for its supply.

2.2 Supply of impact assessment information

Impact assessment for public policy tends to be associated particularly with "impact proving" using more positivist methods, or quantitative data collection, analysed statistically (Hulme & Mosley, 1996). Impact assessment for product development or organisational learning, in contrast, tends to be associated with "impact improving" using more participatory methods, or reliance on interpretation of mostly qualitative data (Woller, 2002). The former is regarded as more expensive but more scientifically rigorous; the latter is cheaper and aims to be credible. "Rigour" here may be taken to mean that findings can be logically deduced from stated assumptions and empirical evidence presented for peer review. "Credibility", in contrast, consists of the use of acceptable processes for the generation of findings that are plausible when compared with knowledge from other sources, including the personal experience as participant observers of microfinance providers themselves.

A historical dimension can again be added to the distinction. Expensive, rigorous, positivist IA can be set up as a 'straw man' to justify the shift to more cost-effective, credible, interpretive IA (Cohen, 1999), (Hulme, 2000). The former can then be linked to top-down replication of microfinance services (demanding impact information for public policy), and the latter to the more participatory processes of product development (demand for impact assessment for organisational learning and innovation). This can then be conflated with a more general 'paradigm shift' or fashion change from a top-down, blue-print, transfer of technology vision of development practice to a bottom-up, process oriented and participatory vision (Chambers, 1997).

The problem with this line of thinking is that it discourages the development of methodologies for impact assessment that do not fit neatly into either category.[1] For example, participatory and interpretive methods of impact assessment also play an important role in moulding public policy. Likewise, quantitative methods are also integral to much market research. The trade-off between rigour and cost-effectiveness may also be more complex. For example, combinations of quantitative and qualitative data collection, may generate data relevant for both internal and public policy purposes (Copestake, Bhalotra, & Johnson, 2001b). The distinction also implies that only positivist methods can be rigorous, thereby neglecting many traditions within the social sciences of rigorous analysis and interpretation of qualitative data (Moris & Copestake, 1993).

The purpose of this paper can now be restated with greater clarity. This is to contribute towards development of microfinance impact assessment methods that:
(a) are *both* impact proving (rigorous) and improving (cost-effective and useful), and (b) can meet *both* internal (organisational development) and external (public policy) demand. In particular, we argue the need for a clearer protocol for qualitative analysis based on the in-depth, semi-

---

[1] See Kanbur (2001) for a parallel debate relating to poverty assessment more generally.

structured, narrative or long interview. Before doing so a brief digression is necessary about the process of methodological innovation.

2.3 The process of methodological innovation

Like all innovation, development of methodologies, tools, instruments and protocols cannot be viewed as an entirely rational process. Promising ideas have to be elaborated and thoroughly tested. It must be possible to explain them easily to potential users, they must acquire currency and legitimacy within communities of users. Semantics, relationships, resources and power are important. A key issue is the extent to which flexibility, heterogeneity and sophistication in product design has to be traded against uniformity, scale of production and robustness.

In the field of development practice there is of course a wealth of experience with impact assessment methods: good, bad and ugly. Many NGOs and aid organisations have their own in-house methods, including some adapted specifically for microfinance (Simanowitz, 2001). There is also a vast and diverse academic literature on microfinance impact assessment. However, it is also possible to identify a relatively small number of more influential institutional players in this field during the last decade. Perhaps the most widely known are the five *SEEP/AIMS tools* developed by USAID in collaboration with SEEP (a US network of enterprise oriented NGOs). The Consultative Group to Aid the Poorest has invested in promoting first an intermediate impact survey tool, and more recently a standard tool for assessing the relative poverty of microfinance clients. MicroSave Africa, with support from DFID and UNDP, is perhaps the market leader in establishing and branding 'industry standards' for market research using participatory methodologies.[2]

It is beyond the scope of this paper to assess the relative usefulness of these different brands and their distinctive products. Nor is it appropriate to explore critically the tendency towards establishment of global standards.[3] However, it is important to locate any new proposal in the context of an already crowded and confused quasi-market place. First, the key question about a new proposal is not whether it constitutes a truly new and original methodological idea - if such a thing is indeed possible. Rather, is there scope for selecting and developing known but diffuse ideas into a 'product' that augments the range of such products that have 'industry' recognition and acceptance? Second, achieving this requires substantial investment not only in product design, but also in product testing and marketing. Are such resources available and likely to be adequate? Third, and particularly in the context of development practice, there is the issue of legitimacy of any new product. Who has been involved in its development and how?

---

[2] All these organisations maintain active websites. A useful single gateway into all of them is the Enterprise Development Website: www.enterweb.org/microcre.htm

[3] For a discussion of these issues in relation to global standards in higher education see Room (2000). For an economic theory perspective see Hudson and Jones (2002). Skevington (2002), in contrast, explores procedural issues in development of a universal yet cross-cultural protocol for assessing quality of life.

The institutional context for this particular proposal is the participation of the authors in *Imp-Act*, a three-year action research programme involving twenty-nine microfinance organisations (MFOs) around the world, sponsored by the Ford Foundation.[4] The programmes main aim is to improve the effectiveness of MFOs by strengthening the mechanisms by which they learn about the impact of their services, particularly on poverty. It emphasises that progress in impact assessment is partially determined by techniques and skills of data collection and analysis, but also by the relationships and feedback processes within which IA tools are used. *Imp-Act* sets out to move away from donor oriented impact assessment towards practitioner-focused processes of listening and learning. Each participating MFO has developed its own pilot impact assessment plans, to fit specific organisational context, objectives and stakeholder needs. A mixture of quantitative, qualitative and PRA/ PLA tools are being used. At the same time, however, the project seeks to facilitate interaction between participants in order to development of generic solutions to the problems and challenges each organisation faces.

One of the most important issues emerging from the first year of the programme has been the need to develop a simple, cost-effective, replicable and credible method for collection and analysis of qualitative data. This paper is part of the process of responding to this need. It is hoped that it will contribute towards development of a standard protocol that can be tested and developed during the remaining two years of the programme. As such it is just one of a number of themes that are in the process of discussion.

3. Towards more rigorous qualitative impact assessment: illustrative case studies

3.1 Overview

The previous section suggested that there is scope for development of a standard protocol for more rigorous qualitative impact assessment. Table 2 suggests that such a methodology would offer a 'third way', with distinct advantages over more established positivist and interpretive tools. In the context of the microfinance industry there are particular reasons for obtaining a better understanding of causal pathways linking provision of services to diverse effects. The industry still tends to assume a very direct chain of impact leading from access to loan capital, to improved micro-enterprise performance and to increased income and poverty reduction. But impact research to date has exposed a far more diverse range of pathways and outcomes.[5] This diversity poses major problems for the design of fully pre-coded questionnaires for quantitative surveys. Participatory methods, even when used by highly experienced facilitators are also restricted in scope by the extent to which discussion of individual experience is constrained by the presence of third parties, particularly other group members. Qualitative approaches can offer greater rigour, by

---

[4] *Imp-Act* stands for Impact Assessment of Microfinance: an action research programme. For more information see its web site: www.Imp-Act.org
[5] See for example, Sebstad (2001) and Marr (2002).

which we mean that there is greater scope for peer reviewers to audit the process by which conclusions are derived from documented evidence and stated assumptions.

Table 2. Methods for supply of impact assessment information

	Positivist methods	*Rigorous qualitative methods*	*Participatory methods*
Data collection method	Sample survey using a closed questionnaire	Quota sample of semi-structured narrative interviews.	Case study focus group discussion and semi-structured games
Data analysis	Interpretation of statistical tables and multiple regression analysis.	Systematic scoring of types of impact from interview notes and transcripts.	Verbal and experiential learning among participants. Interpretive reports.
Epistemology	Rigorous statistical inference and peer review.	Rigorous qualitative analysis and peer review.	Process transparency and expert judgement.
Potential strengths	Rigour. Possibility of quantitative estimates of impact. More convincing to sceptical outsiders.	Richness in detail and understanding of differential impact. Ability to pick up unexpected and unmeasurable impact.	More timely and cost-effective, with potential for shorter feedback loops from suppliers to users of information. Can be positive learning tool for respondents too.
Potential weaknesses	High cost and time lags. Restricted to measurable impact indicators. Reveals little about causation. Difficult to counter selection bias problems.	Demonstrating that findings are representative of wider populations. Lack of clarity and consensus about how to achieve rigour.	Participants may hide important facts from peers as well as facilitators. Risk of response bias makes it hard to convince outsiders of the reliability of findings.

There is of course a vast literature on the use of qualitative methodologies in the social sciences.[6] There is also no shortage of literature on the use of open-ended individual interview, particularly if this is broadly defined to

---

[6] See for example (Bernard, 1994), (Denzin & Lincoln, 1994), (Flick, 1998), (Hammersley, 1992), (Maynard & Purvis, 1994), (Moris & Copestake, 1993), (Silverman, 1997).

include, for example, "the long interview" (McCracken, 1998), the "life history" (Atkinson, 1998), the "in-depth interview" (Chirban, 1996), and "oral testimonies" (Onselen, 1993). One dimension of variation within this literature is the extent to which interviews are structured. In general, it can be argued that the more open the format then the more skilled and time consuming are the tasks of interviewing, data analysis and interpretation. Given the emphasis of this paper on the need for methods that are relatively cost-effective and useful if follows that the emphasis in this paper is on methods that are towards the more structured end of the spectrum. The remainder of the section elaborates by presenting brief case studies of the use of semi-structured interviews by the authors in Kenya, Zambia, Malawi and Peru.

3.2. An example from Kenya

In Kenya the initial demand came from the "British Aid for Small Enterprise" programme of the UK Department for International Development (DFID) as part of its requirement to assess impact of a programme of grant support for six Kenyan MFOs. (REME, 1999) An independent marketing research firm was hired to conduct an initial round of single-visit qualitative interviews with a random sample of 15 clients per MFO (Research International, 2000). The work was designed both to provide useful feedback to MFOs themselves about their own clients, as well as build up a broader sector-wide picture of direct impact for DFID.

The questionnaire schedule comprised three closed-question sections, and three open-ended narrative sections, as follows. Section A (closed) collected basic information about the identity of the respondent. Section B (open) started with the generative question "To start off, tell me a little bit about yourself, and the people you live with." The interviewer was also given four pre-set probing questions to use if necessary to ensure that the respondents narrative account covered all key aspects of the life and livelihood of the respondent and her family. Section C (closed) then supplemented this narrative with key quantitative data on household composition and income - though in many cases it was possible to fill these in during the previous open narrative section. Section D (open) then explicitly raised the issue of access to and use of financial services, including those of the MFO. The prompting question was "I now want to go into some details in your own words about the financial services that you obtained from…. The best way to do this would be for you to start by explaining how you first heard of it and became a client. You can then tell me all the things that happened after that and up to now." Again the interviewer was also given optional probing questions to use to maintain the flow of the narrative if necessary. All questions were printed in Swahili as well as English. Section E (closed) again supplemented the narrative with key quantitative data on loan use, but the fact that these questions came after the narrative section ensured that they did not guide the narrative section. Section F (open) closed the interview by asking the respondent to explain in their own words what plans they had for the future - for themselves personally, for other members of the family and for their business activities.

Where respondents agreed, narrative sections were tape-recorded. The interviewer also typed up a summary of the narrative sections to a tight word limit, drawing upon their memory, field notes and the tape recording as needed.[7] The field supervisor and interviewer then met to review the narrative summary and agree on what level of impact it represented against agreed criteria, of which the two most important were sustainable income generation and employment creation. These were scored on scales of one to five, and then entered onto a single database along with pre-coded data from the three closed sections of the interview schedule. In other words, the attribution problem was addressed by making a judgement about the plausibility and consistency of respondents' own narratives. The impact scores could then be analysed through cross-tabulation against pre-coded variables.

Each MFO received a report comprising a six-page summary of the findings of each interview, together with a summary of what they collectively revealed about impact in relation to seven hypotheses. Findings were generally more negative than they were used to hearing, and evoked a defensive reaction including criticisms of the research team. However, this was perhaps more a reflection of lack of ownership of sensitive findings on the part of the MFOs than weaknesses in the underlying methodology. DFID scored impact against each hypothesis so that it could be entered into the date-base for analysis along with the data obtained from the closed sections of the questionnaire.

3.3 An example from Zambia

In Zambia, semi-structured interviews were used as part of an ongoing three year programme of impact monitoring and assessment for a DFID sponsored MFO starting up on the Copperbelt (Copestake, Gosnell, Musona, Masumbu, & Mlotshwa, 2001a). An initial impact survey was carried out of a matching sample of one-year-old and new or "pipeline" clients using a closed questionnaire. In-depth interviews were then carried out a year later for a randomly selected sub-sample of this survey. Their principal aim was to improve understanding of the causation of impact. They also elicited data on information less easily collected through a standard questionnaire, such as changes in intra-household relations. Data was mainly collected through semi-structured interviews with individual clients, supplemented where possible with group discussions at the group level and discussions with other members of selected clients' households.

Attribution of impact was again based primarily on the coherence of respondents' own accounts of causal chains linking their membership of credit groups to changes at the individual, business, household, and community level. Particular care was taken to encourage respondents to discuss changes arising specifically from their participation in the programme, rather than changes that might have occurred anyway. The plausibility of their account was also cross-checked against information

---

[7] The tape recording was then filed so that the research supervisor could carry out a random audit of whether typed narrative summaries were consistent with them.

from direct observation, the original sample survey interview, and discussions with loan officers and other group members.

The first section of the interview format recorded basic factual and coding information about the client. Separate sections then dealt in turn with changes in household circumstances, changes in business activities, experience as a group member, impact on individuals and views of other household members. Each section was initiated by an open-ended generative question designed to encourage respondents to describe their experience in their own words in as free and fluent a way as possible. A standard set of prompting questions was also prepared for each section to assist the interview as and when the respondent stopped talking. In many cases it was not necessary for the interviewer to ask all these questions; rather they served as a checklist of topics to be covered. Each section closed with one or two pre-coded attitudinal questions, which asked respondents directly, taking into account the previous conversation, to sum up how participation in the programme had affected them.

As in Kenya, notes on each narrative section of the interviews were immediately typed up after the interview, and used as a point of reference for attributing an impact score to the client against different criteria. This data was then entered into a database along with their responses to the attitudinal questions and statistical data about them from the survey a year earlier. An aggregated index of impact was also calculated. This enabled the sample to be divided into a group for whom overall impact was "clearly positive" and a group for whom impact appeared either negligible or negative. Statistical data from the first survey round was then used to analyse possible of the determinants of this differential impact. Finally, the data was also analysed and written up as a short report by cutting and pasting data from each narrative summary into sections covering different types of impact, including the four dimensions of individual impact (material, cognitive, attitudinal and relational) highlighted by Chen (1997).

3.4 An example from Malawi

The third case study also comes from DFID sponsored impact research, in this case of a microfinance programme in Malawi (Johnson, Copestake, & Consultants, 2002). Research in the Central Region of the country had two components. The first was a quantitative survey of approximately 400 interviews, comprising borrowers and a comparison group. This was based on two rounds of interviews carried out a year apart using a fully pre-coded questionnaire. To complement this a panel of 20 borrowers were selected for qualitative interviewing. The purpose of this component was to better understand impact pathways and variables mediating outcomes on clients. These interviews were carried out quarterly over a year involving participants in a total of five interviews.

Sample selection aimed to capture diversity in members' experience of the programme and of group dynamics. This was done through a two-stage process of selecting groups with varied performance and selecting

members within them of different wealth levels. These selection processes were carried out using participatory ranking methods with MFO staff and group members in turn, followed by random selection within the ranked categories.

All interviews were semi-structured. They covered respondents' basic household and livelihood information, participation in the village bank in terms of loan history and performance, experience and views about the operation of the VB, their experience and views on any other effects that the VB had had on their lives. Basic household data was updated on each occasion to ensure that the current status of the household could be captured. The interviews were taped and later transcribed as narratives. Analysis involved the overall assessment of the information given by the analyst as to whether the impact of the programme had been positive, negative or ambiguous (in some cases opinions contradicted reported facts) on material welfare. Narratives were reviewed both against a pre-existing set of questions about material, cognitive, perceptual, and relational impact. Explanation of the impact pathways and mediating variables was established by analysing the narratives for emerging themes and patterns that could explain the impact outcomes.

The re-visits were of critical importance to the research. Through them the interviewer developed a close rapport with the respondents and was able to probe more deeply into the reasons why decisions were taken. They also enabled a much more consistent picture of household dynamics to be built up and the development of a more detailed understanding of the environment in which the programme was operating. Moreover, it became possible to identify the differences between what people reported (i.e. what they wanted us to hear) and the reality of the actual events they experienced. Compared to the single visit approach the revisits enabled a high degree of confidence in the data set to be developed. They also generated data that would not have been captured in a single round about the ways in which businesses stopped and started in relation to women's own skills, experience, marital status, domestic situations, seasonality, and loan programme design.

3.5 An example from Peru

The last example is of doctoral research, conducted over a period of eighteen months, to review critically the extent to which microcredit could be empowering for women in low-income rural and urban locations of Peru (Wright, 2001). In this case, the emphasis was less on how representative data was, and more on understanding the social processes affecting microcredit outcomes. It also 'gave voice' to a diverse group of women and their husbands or partners, and provided source material on how microcredit agencies interacted with other institutions at the community level. An important feminist theoretical standpoint was to articulate the perspectives of multiple participants (Henwood & Pidgeon, 1995, 23).

The sample comprised over one hundred and fifty women participating in a range of microcredit schemes operating in Lima and Cajamarca. Care

was taken to ensure that the sample included households at different stages of the life cycle; including single, separated and married women, heads of households and widows. Interviews were structured around a standard list of questions, with responses transcribed from a tape-recording or noted down by the researcher. Data collected included the following: (i) age, marital status, education, dependants, migration and residence history of respondent; (ii) microenterprise activities and motivation for taking out credit; (iii) changes in the family and microenterprise; (iv) socio-economic profile of the household: income and expenditure, structure of the family, decision making, male and female roles in the household; (v) political and economic context; (vi) design of the microcredit scheme; and (vii) perceptions and aspirations for the future.

Life history methods were considered the most appropriate technique for gleaning this kind of information. As Davidson (1999) has observed, documenting personal histories and struggles can illuminate both an individual's courses of action (or inaction) as well as the effects of constraints and barriers which evolve over the life-course at different times and places. Detailed interviews in both rural and urban areas took the form of guided conversations (Lofland, 1971). The majority of interviews in both Cajamarca and Lima were conducted in women's homes (which often doubled up as their workplaces) and in markets. The advantage of interviewing women in their homes was that it allowed the observation of relationships and interactions between family members. Although men were not interviewed systematically, there were many opportunities for conversations with men about gender relations too. Where possible, second and third interviews were conducted with the same informants in order for them to have the chance to think about the issues over a longer period of time.

Being fieldwork for a doctoral thesis less importance was attached to timeliness or cost-effectiveness of data collection and analysis. However, the slower and more open-ended fieldwork strategy highlights important issues affecting the quality of data that any protocol for qualitative impact assessment would also have to address. For example, during the interviewing stage, much emphasis was placed on the construction of questions and techniques outlined by Foddy (1993) such as the importance of the ordering and wording of questions. The questions were then carefully piloted. At this point in the research process, the need to rephrase the questions drawn up meant that the researchers prejudices and pre-conceptions as a Western woman were revealed and her identity was questioned. For example, the question: "What problems did you experience?" did not get a response. Women in rural areas were sensitive to the suggestion that "they had problems" and so the question was rephrased to "What makes you angry?" This rewording unlocked the key to the real issues and proved to be the turning point in this research (Wright, 2002).

4. Towards a standard protocol for the use of in-depth interviews

4.1 Introduction

The previous section served to illustrate the scope for using the in-depth interview as a flexible method for obtaining rich data about impact. This section goes one step further by suggesting what a standard protocol for such an approach might look like. For ease of reference this prototype protocol is called the "Qualitative impact protocol" or *QUIP*, the word quip being defined as a 'pointed saying.'

4.2 Objectives, sample selection and scope

A key consideration is that a *QUIP* should be able to provide information on impact quickly, cost-effectively and reliably enough to facilitate organisational learning. This can be done in two ways, though these are not mutually exclusive. First, specific one-off studies may be commissioned to examine particular aspects of impact in relation to specific concerns about performance that the organisation wishes to address. Second, qualitative interview material may be accumulated in a database over time. Indeed, one important advantage of the qualitative interview is that benefits (in terms of useful data) accrue continuously in relation to the number of interviews carried out, whereas meaningful positivist analysis requires a critical minimum sample size, as well as inclusion of a control group.[8] Hence *QUIP* should be useful for managers who are interested in understanding problems with specific sub-groups of clients for which they can commission a quota sample of interviews. At the same time, adoption of a standard protocol should permit findings from batches of interviews to be integrated cumulatively into a standard database for further and wider analysis.

As with any piece of research the first step is to identify the questions that the research seeks to answer. While the overall question of impact may be clear, specific questions about impact pathways, outcomes arising from specific aspects of programme design, or additional questions about wider impact will vary between organisations and over time.

This raises the question of sample selection and how representative findings can be from smaller but richer data sets. The important point is to emphasise that there are a range of choices between the purely anecdotal case study and the fully statistically representative sample. Qualitative data is most useful where it can be combined with a clear typology for stratification of the MFOs client portfolio.[9] Researchers can then use quota sampling to update the profile of the different client

---

[8] This difference arises from the contrasting epistemologies underpinning the two approaches. Qualitative enquiry addresses the attribution problem by assessing the consistency and coherence of respondents' narratives. Positivist enquiry seeks (often with partial success to problems of selection bias) to solve the attribution problem through statistical inference.

[9] See Copestake (2001)for a full discussion of the idea of client portfolio classification.

groups.[10] The aim is <u>not</u> to estimate average statistics for all clients, but to ensure that management and other stakeholders improve their understanding of the <u>diversity</u> of impact across the client portfolio.

Turning to consider the scope of data collected, a major potential strength of a *QUIP* is that it would offer managers a broad or holistic understanding of the impact of their services - at business, household, individual, community and market levels. Clients are likely to talk first about the direct impact of participation on their own business activities and income. But with appropriate probing, they can then be encouraged to talk about direct non-material impact on their understanding (e.g. of business management), perception/attitudes (e.g. self confidence), and relationships (with MFO staff, other group members and kin (Chen, 1997). More experienced clients can also be used as key informants for discussing wider or indirect impact, on overall access to financial services, market opportunities, supply constraints and community activities[11]. The open-ended scope of the approach is important, because reliability depends in part upon the consistency of respondents' explanations for why certain kinds of impact are more marked than others.

This broad scope of interviews contrasts strongly with impact assessment based on pre-coded questionnaire based approaches, which have to be more precise and selective about research hypotheses. On the other hand, efficiency of data collection and analysis does require that interviews are clearly planned and structured around broad domains of impact. There is also scope for combining in-depth interviews with quantitative sample survey based work. The case studies in the previous section illustrate the range of possibilities. In Kenya, open and closed questions were combined sequentially in a single interview schedule. The main drawback of this was the resulting length of the interview. In Zambia, the in-depth interview was a repeat interview, building on base-line findings about the same clients from a closed questionnaire a year earlier, which also served as the sample frame. In Malawi repeat in-depth interviews were undertaken in parallel with an impact survey, but with separate clients, thereby ruling out the possibility of being able to combine and then jointly analyse quantitative and qualitative data. A further option is to supplement in-depth interviews with data routinely collected by the MFO, for example at the loan appraisal stage.

4.3 Preparation for field work and the conduct of interviews

Given the more open-ended nature of the data collection process, adequate training of field researchers is particularly important. One set of issues concerns how interviews are set-up and initiated. Another is the

---

[10] Quota sampling starts with a clear definition of sub-categories of the population. For example, this might be based on client loyalty. A decision is then made to select an initial quota of respondents from each sub-category. A random mechanism for doing this must then be identified. If the data generated by an initial quota does not reveal a clear pattern then additional quotas may be sampled. The method is inefficient in enabling statistical inferences to be made about mean values for the whole population. But it is effective in building on prior knowledge of the <u>diversity</u> of the population.

[11] See for example Copestake (2002). This reports on information collected by visiting a quota sample of 'three year old' clients on the Zambian Copperbelt who had already been interviewed twice before, as described in Section 3.2.

ethical dimension. Many essential principles of good ethical practice have been outlined in the literature and applied by researchers in the field. However, proponents of participatory methods go further, suggesting that data collection and analysis should empower informants (Chambers, 2002). Not all microfinance organisations adhere to such a participatory ethos. Some microfinance organisations are designed for service delivery rather than empowerment. Nevertheless, guidelines for ethical practice must be incorporated into the protocol. In order to build trust and minimise response bias respondents need to understand who researchers are, what the data is being collected for, and how it will be used. There is also the fraught issue of whether respondents should receive any material incentive or 'token of thanks' for agreeing to participate. The best approach will clearly be culture and context specific, limiting the scope for a standard, prescriptive protocol. However, there may still be scope for preparing some kind of decision tree to help guide researchers through a range of options.

A second set of issues concerns design of the interviewing schedule, including framing and phrasing of questions. A trade-off exists between questions that act as prompts and open-ended questions. In the former, the researcher introduces new elements or invites responses to a specific topic. Such questions risk not allowing the informant to answer using their own categories, but rather imposing categories based on the prejudices of the researcher and thus tainting the nature of the response. But without some structure, on the other hand, interviews may becoming unacceptably time consuming or irrelevant.

There is no substitute for the skills of the researcher in sensitively addressing this dilemma. However, less experienced researchers can be assisted not only through training but also through development of a flexibly structured interview schedule. The first piece of structure is to identify a broad agenda of topics to be covered, starting with the most general ones, and moving on to those that relate directly to the service under review. Each topic can be introduced with a predetermined and tested open-ended 'generative' question, leaving the informant free to talk in their own terms. Researchers may then also be given a set of optional supplementary questions, phrased to raise issues but not suggest answers. Even if they are not used these can serve as a useful checklist of what is covered by unstructured narrative. In the final analysis, some trade off is necessary. Pragmatists have to learn to be more patient, while academic purists have to be realistic. The pay-off is that such an approach should improve understanding of issues, such as group dynamics and power relations in the wider community, which can negatively affect the operation of development projects but which organisations currently neglect.

4.4 Data entry and analysis

Having discussed the need for attention to question design and data collection methods, we now return to the major task which is to devise a simple, replicable, timely and cost-effective way of analysing the data collected. Feedback from the first year of the *Imp-Act* programme

suggests that it is lack of clarity about how to go about systematically analysing qualitative data that inhibits MFOs from using this approach more.

Broadly, we suggest that there are two ways of analysing the data, and that these can either be used in isolation or to complement each other. The first is to produce a written report that summarises the findings from a set of individual in-depth interviews. Starting with typed written summaries, this can be done by sorting excerpts from each by a list of topics. This work can be done using a qualitative data analysis package such as nudist, by using cards, or more simply by cutting and pasting text in a word processing package. Either way the analysis is illustrated by Table 3. Reading the table downwards by column we have excerpts from the fictional narrative summaries.[12] This can then be turned into a report that moves systematically through the topics in the left-hand column. Thus under "wider market" there is reasonably strong evidence that increased access to credit resulted in the market for fish becoming more competitive. Note that the fuller and more explicit the links are from narrative summaries to the general report then the more rigorous are the findings in the sense that peer reviewers can trace the evidence on the basis of which generalisations are made. Reliability is increased even further by randomly auditing the quality of narrative summaries against archived tape recordings of the original interviews.

Table 3. Illustration of qualitative analysis of narrative summaries

Case No. Topic/level	Respondent 1	Respondent 2	Respondent 3 etc
Business activity	Selling fish	Selling fish	Selling fish
Material impact on business	She was able to purchase stock in bulk (+)	She took on a second stall in the market (+)	Her capital was used up so she stopped trading (-)
Material impact on household	She bought more food for the family (+)	Her children had to help her in the market (-)	They had to skip meals in order to save money for repayments (-)
Material impact on individual	She was able to buy better clothes (+)	She felt tired all the time (-)	She went hungry more often (-)
Cognitive impact on individual	She learnt how to manage money (+)	She was forced to learn how to manage her time better (+)	She decided never to borrow money again (0)
Perceptual impact on individual	She started to make more plans (+)	She became less lazy (+)	She lost confidence (-)
Positional impact on	She became less dependent	She made friends who helped her	She returned to ask for more from her

_____
[12] For more details see the Appendix.

individual	upon her brother for cash (+)	out on the stall (+)	husband (0)
Financial market impact	She opened a bank account too (+)	She shifted her savings from the cooperative (0)	She started saving small amounts of money (+)
Wider market impact	More women used loan money to enter the business ())	More women used loan money to enter the business (0)	Customers insisted on buying on credit (0)
Wider social impact	She became more confident public speaking (+)	She learnt about the market by talking to other members (+)	She fell out with old friends who had to cover for her arrears (-)

The second way of analysing narrative summary data is by using simple scoring methods. In the above case, for example, impact has been classified by the interviewer on a three point scale (+, 0, -). Adding these scores horizontally, facilitates analysis of overall impact in relation to each level. Adding the scores vertically produces a single impact score for each respondent (+8, +3, -4). These scores can then be analysed by cross-tabulation against statistics about the respondent, their business, location and so on.

Of course such scoring is an imperfect and subjective process.[13] Yet, the pay-off is that complex findings from numerous interviews can be summarised, analysed and presented to others much more easily. The quality will depend upon the reliability and level of detail of the narrative summary, and the skill of the researchers. It may also be affected by choice of scale and weights ascribed to different categories of impact, and there is considerable scope for experimentation with different ways of scoring. In the Zambia case, for example, scores were elicited not only by subjective scoring by the researcher on the basis of the narrative summary, but also by asking respondents themselves to make an assessment immediately after concluding the open narrative section of the interview that raised the related topic. Comparable findings based on respondents' own questions and on simple scoring for the same sample of sixteen interviews are reproduced in the Appendix.

4.5 Research management, staffing and funding

In the context of the *Imp-Act* programme, most MFOs are interested in impact findings that can be used to feed into the development of new products or refinement of existing ones. The thrust of *Imp-Act* is therefore to experiment with approaches to impact assessment that are within the means of MFOs themselves and meet the needs of their stakeholders. In most of the projects it is MFO staff themselves who are implementing the research, albeit after being trained by external specialists, and with their

---

[13] Where respondents are asked to score or rank data, this may have cultural or symbolic meaning of which the researcher is unaware. Nevertheless, it has been argued that subjective values are amenable to mathematical analysis and quantification (Hanton, 2002).

ongoing supervision and guidance. This approach has been facilitated by the existence of pre-packaged tools, particularly those developed by SEEP/AIMS and MicroSave.

It is tempting to infer from the above that with clearer guidelines, in the form of QUIP, MFOs can similarly be helped to make more use of qualitative methods using their own staff or directly recruited social science graduates. However, some caution is needed. Not only do more open-ended methods of interviewing and analysis require more specialised researchers (Copestake, 2000). Their value also resides in opening up more sensitive issues, such as relations with field staff and group leaders. But this potential is less likely to be realised by researchers who belong to the organisation itself. Larger MFOs may be able to afford to keep such specialists in-house and allow them sufficient autonomy from operations to avoid bias. But for most MFOs there is a strong case for contracting out such work to specialist research consultants. MFO network and technical support agencies may in some instances also be able to carve out a useful role in provision, or facilitation of such work.

5. Conclusion

This paper has been concerned with highlighting the need for a practical qualitative impact assessment protocol (referred to as QUIP). It has raised issues that would need to be addressed in devising such a protocol and has tentatively suggested the form it might take. Production of such a protocol would hardly represent an earth-shattering methodological breakthrough. Rather, it would represent a negotiated compromise between the canons of social science and the requirements of MFOs for methods that are cost-effective and replicable, yet also more rigorous and reliable than most current practice in this area.

By developing guidelines on data collection and analysis it is hoped that MFOs will be able to make use of qualitative methods with more confidence and to better effect. However, this paper represents only a first step towards production of such guidelines. To build credibility and acceptance it is recognised that further design and testing is required, and it is hoped that the Imp-Act programme will provide a framework for this. In particular, this work will require a continued balancing act between standardisation and flexibility: the former for consistency and ease of use, the latter to allow adaptation to specific conditions and ends. Methodological guidelines by themselves will not go far towards guaranteeing quality. Rather the protocol must also contain guidelines on how the process of using the methodology should be documented, so as to facilitate peer review.

Appendix: Quantification of impact using direct questions and scoring methods.[14]

Table 4. Responses to attitudinal questions

S8. "Overall, how has your economic situation changed since the last interview?"	Better	Worse	No change	Not sure
	8	6	1	1
S9. "How does this compare with other people living in your area?"	Better	Worse	No diff	Not sure
	7	3	1	5
T7 "Overall, how profitable are your business activities now compared to a year ago?"	More	Less	No change	Not sure
	7	6	1	2
T8 "How does this compare with other people (but not in your TB) doing this business?	Better	Worse	No diff	Not sure
	6	5	0	5
V4 "Overall, how has belonging to a trust bank affected you as a person"	Better	Worse	No change	Not sure
	11	1	2	2
W5 "How has the welfare of your household been affected by the respondent joining a TB?"	Better	Worse	Same	Not sure
	7	5	1	3

Table 5. Summary of scores derived from narrative summaries (n=16)

	Hypothesis	+2	+1	0	-1	-2
1	Trust Banks provide better services to continuing members during their 2nd year	4	3	2	5	2
2	Sustained TB membership strengthens clients' business activities	5	3	2	5	1
3	It leads to material benefits to clients and to other household members	3	6	1	2	4
4	It resulted in increased (+2) or more stable (+1) employment	2	6	5	2	1
5	It improves clients' personal capabilities, self-confidence and capacity to plan	1	7	5	3	0
6	It strengthens their role in decision making within the household and beyond	4	2	5	4	1
7	Other household members view TB membership positively	4	3	4	4	1
Key. +2 indicates strong support for the hypothesis, +1 indicates support, -2 indicates strong support for the counter hypothesis, -1 indicates some support for the counter hypothesis, 0 indicates no support either way.						

---

[14] This data is fully analysed and discussed in Copestake and Mlotshwa (2000).

References

Atkinson, R. (1998). *The life story interview*. London: Sage.

Bernard, H. (1994). *Research methods in anthropology: qualitative and quantitative approaches*. London: Sage.

Chambers, R. (1997). Whose reality counts? Putting the first last. , Longman.

Chen, M. A. (1997). *A guide for assessing the impact of microenterprise services at the individual level. *. Washington D.C.: USAID Office of Microenterprise Development. AIMS Project.

Chirban, T. (1996). *Interviewing in depth: The interactive-relational approach*. London: Sage.

Cohen, M. (1999). *Microfinance Impact Evaluation: Going Down Market* (Paper Presented for Conference on Evaluation and Poverty Reduction, World Bank). Washington, DC: USAID.

Copestake. (2002). Poverty, inequality and the polarising impact of microcredit: evidence from Zambia's Copperbelt. *Journal of International Development, 14*(6).

Copestake, J. (2001). *Towards a general framework for client monitoring within microfinance organisation* (Learning note No.5). Brighton: Imp-Act: Improving the impact of microfinance on poverty: an action research project. Available under new publications from www.imp-act.org.

Copestake, J., Gosnell, H., Musona, D., Masumbu, G., & Mlotshwa, W. (2001a). *Impact monitoring and assessment of microfinance services provided by CETZAM on the Zambian Copperbelt: 1991-2001. Final Report* (Available on www/bath.ac.uk/cds). Bath: Centre for Development Studies, University of Bath.

Copestake, J. G. (2000). Integrating impact monitoring and assessment of microfinance. *Journal of Development Practice, 10*(4).

Copestake, J. G., Bhalotra, S., & Johnson, S. (2001b). Assessing the impact of microcredit on poverty: a Zambian case study. *Journal of Development Studies, 37*(4).

Copestake, J. G., & Mlotshwa, W. M. (2000). *Impact Monitoring and Assessment of the Christian Enterprise Trust of Zambia: report on the first round of repeat client interviews* (3). Bath, UK: Centre For Development Studies, University of Bath.

Davidson, G. (1999). *Poverty amidst plenty: a study of disadvantage, vulnerability and social exclusion in Singapore. *, University of Liverpool, Liverpool.

Denzin, N., & Lincoln, Y. (1994). *Handbook of qualitative research*. London: Sage.

Flick, U. (1998). *An introduction to qualitative research*. London: Sage.

Foddy, W. (1993). *Constructing questions for interviews and questionnaires: Theory and practice in social science research*. Cambridge: CUP.

Hammersley, M. (Ed.). (1992). *Social research*. London: Sage.

Hanton, D., (2002). From words to numbers: A basis for translating ethnographic description. *Unpublished article.*

Henwood, K., & Pidgeon, N. (1995). Remaking the link: Qualitative research and feminist standpoint theory. *Feminism and Psychology, 5*(1), 17-30.

Hudson, J., & Jones, P. (2002). *International trade in quality goods: signalling problems for developing countries.* (Unpublished): Department of Economics and International Development, University of Bath.

Hulme, D. (2000). Impact assessment methodologies for microfinance: theory, experience and better practice. *World Development, 28*(1), 79-98.

Hulme, D., & Mosley, P. (1996). *Finance against poverty*. (Vol. 1). London: Routledge.

Johnson, S., Copestake, J., & Consultants, w. K. (2002). *FINCA-Malawi Impact Assessment Research: final report.* . Bath: Centre for Development Studies, University of Bath.

Kanbur, R. (2001). *Qualitative and quantitative poverty appraisal: complementarities, tensions and the way forward.* . Ithaca: Cornell University.

Marr, A. (2002). Studying group dynamics: an alternative analytical framework for the study of microfinance impacts on poverty reduction. *Journal of International Development, 14*(4), 511-34.

Maynard, M., & Purvis, J. (Eds.). (1994). *Researching women's lives from a feminist perspective*. London: Taylor and Francis.

McCracken. (1998). *The long interview*. (Vol. 13). London: Sage.

Moris, J., & Copestake, J. (1993). *Qualitative enquiry for rural development: A review*. London: IT Publications.

Onselen, V. (1993). The reconstruction of rural life from oral testimony: Critical notes on the methodology employed in the study of a Black

South African sharecropper. *Journal of Peasant Studies, 20*(3), 494-514.

REME. (1999). *Synthesis Report on BASE Programme* .

Research International. (2000). *Financial Services Project Exploratory Study: case study reports.* : Research International (E.A.) Ltd. and Department for International Development (Nairobi).

Room, G. (2000). Globalisation, social policy and international standard-setting: the case of higher education credentials. *International Journal of Social Welfare, 9*(2), 103-119.

Sebstad, J., & Cohen, M. (2001). *Microfinance, risk management and poverty* . Washington D C: World Bank.

Silverman, D. (1997). *Qualitative research: theory, method and practice*. London: Sage.

Simanowitz, A. (2001). *A review of impact assessment tools* : *Imp-Act*: Improving the Impact of Microfinance on Poverty: an action research programme: www.Imp-Act.org.

Skevington, S. (2002). Advancing cross-cultural research on quality of life: observations drawn from the WHOQOL development. *Quality of Life Research, 11*, 134-144.

Woller, G. (2002). From market failure to marketing failure: market orientation as the key to deep outreach in microfinance. *Journal of International Development, 14*(3), 305-324.

Wright, K. (2001). *Women's participation in microcredit schemes: evidence from Cajamarca and Lima, Peru.* , University of Liverpool, Liverpool.

Wright, K. (2002). *"Problems, madam? We have none at all!" Qualitative data collection for impact assessment: getting the questions right*. Bath: Centre for Development Studies.