

Choosing tools and methods

This chapter explores the ‘families’ of tools and methods used in the case studies: surveys; interviews, workshops and discussions; direct observation; participatory research; and case studies. Drawing on the lessons learned in our impact assessment studies, I draw some conclusions about their relevance and utility in a range of contexts and for various purposes. The final section also looks at how two of the case studies explored the relationship between costs, benefits, and impact.

Choosing appropriate tools and methods depends on the purpose and focus of the impact assessment, its context, the capacities and skills of those involved, and the resources available. One of the major skills needed by those involved in impact assessment is the ability to find an appropriate combination and sequencing of tools and methods. None of the tools presented below, *on their own*, is likely to be sufficient in determining impact, cross-checking findings, and dealing with attribution.

Surveys

A survey is simply a method of establishing a comprehensive overview of a given situation. Surveys are most suited to answering questions such as *what? how many? and how often?* Although they can be used to answer *why?* questions, they tend to be less useful in doing so. They are therefore often used to gather basic data about a group of people (their age, sex, occupation, and so on). Surveys generally use a questionnaire and seek standard, quantifiable data from a representative population. If the number of people or groups being studied is small enough (generally less than 100, depending on the number of questions) a survey can cover the entire population. Two main types of survey which were undertaken in the case studies: large-scale questionnaire surveys covering respondents across several communities, and smaller-scale mini-surveys of the populations of a limited number of communities.

Large-scale questionnaire surveys

Four of the case studies (ENDA, BRAC, Wajir, and Proshika) used large-scale questionnaire surveys varying from 200 to 1,800 respondents. Between four

and 60 enumerators administered the questionnaires face-to-face with the respondents. Although in all the studies the household was the main unit of assessment, one of the studies also targeted individual women.

Content

Each of these case studies took slightly different approaches to the survey. ENDA felt that because of their lack of baseline data they needed to use the questionnaire to identify the situation prior to project implementation and subsequent changes.

However, the team also felt that it was important to try and capture the project processes that had taken place using the questionnaire. This way it was felt that we could obtain a general perception of the project and not just the perceptions of the key informants. (ENDA case study report)

However, as a result the questionnaire was long and, possibly because it tried to collect too much information, 'the objective of trying to quantify impact still eluded the team'. They subsequently found that '[m]ost of the data that was collected using the questionnaire supported the findings of the focus-group discussion' which they undertook later. 'We feel that for the effort that was put into the questionnaire, the added value ... was not much greater than the focus-group discussions'. In retrospect, the team felt that although questionnaires may be a useful tool to assess the extent of impact it might have been better to run a shorter questionnaire after discussing possible indicators, defined by the community through the kinds of participatory methods subsequently used by the team (see section below).

The BRAC, Wajir, and Proshika studies, which had quite solid baseline data, focused their questionnaires much more tightly on presumed impacts. They also collected information to help assess whether the changes they identified could be attributed to the projects in question. The Wajir study deliberately set out to assess three levels of analysis (see Table 4.1). Level one relates to significant changes in pastoralists' welfare, as assessed by levels of food aid received; quality of diet; perceptions of changes in quality of life; child mortality; and capacity to cope with drought. Level two explores changes in livelihood which might explain the impacts on welfare described in level 1 and which relate to the outcomes of the project, such as reduced animal mortality; increased reliability of water; improved law and order; and the rate of return of new investments provided through credit and restocking. Level three examines the changes in the quality of and satisfaction with delivery of services which might explain the changes in livelihood and welfare; these can be considered as another lower level of outcomes. The areas addressed by the questionnaire are specifically chosen to trace the hierarchy of the project's objectives set out in a logical framework at the outset of the project, mainly covering outcomes and impacts.

Table 4.1: Cause-and-effect relationships between the Wajir project and pastoral welfare, in relation to the hierarchy of project objectives

Hierarchy of objectives	Indicators of objectives
<p>1 To enable pastoral communities to live in peace, without fear of starvation or ill health, while minimising their dependence on food aid and other forms of welfare support</p>	<ul style="list-style-type: none"> • Lower food aid requirements • Better quality diets • Reduced seasonal peri-urban destitution • Pastoralists perceive an improvement in their quality of life and their capacity to survive drought • Lower child-mortality rates
<p>2 To strengthen pastoral livelihoods by establishing and/ or supporting sustainable local institutions</p>	<ul style="list-style-type: none"> • Fewer animals die • Water supplies during the dry season are more reliable • Exchange rates between livestock or livestock products and essential household commodities improve • Private and public property rights are maintained • Pastoralists obtain income from a range of new sources
<p>3 To improve the delivery of essential services (health-care, education, water) by establishing sustainable institutions</p>	<ul style="list-style-type: none"> • Herders use quality drugs at correct dosage rates • Pastoralists sell animals and animal products at xx% of the price obtained in Nairobi • Prices for grain and other essential commodities do not rise more than xx% above Nairobi prices and remain stable • Pastoralists send their children to school and are satisfied with the quality of education • Water supplies are correctly maintained to avoid breakdown or siltation in the dry season • Food aid, credit, and other inputs are efficiently distributed to those most in need

The BRAC questionnaire was conducted primarily to collect quantitative data to determine the impact of their Rural Development Programme on its participants, with special reference to their material well-being, and their level of poverty in particular. However, they also included social aspects of well-being such as literacy and educational level of households, housing status, and some aspects of health, sanitation, and family planning.

Both BRAC and the Wajir impact assessment team consulted with field staff and, in BRAC's case, some village organisations in the process of developing expected impacts and indicators which became the focus of the questionnaires (see Table 4.2).

Table 4. 2: Main areas of questions asked in Wajir and BRAC household surveys

Wajir	BRAC (RDP)
<ul style="list-style-type: none">• Demographic and household characteristics• Wealth ranking: rich/ poor• Quality of life• Food aid received in past 12 months• Milk consumption per day over past 12 months• Numbers of meals per day• Child sickness and mortality• Water supplies• Income and expenditure• Conflict situation• Capacity to cope with drought• School attendance and satisfaction• Future aspirations for children• Animal health/ drug treatment• Herd profiles• Herd movements in last drought• Role of pastoral association	<ul style="list-style-type: none">• Demographic and household characteristics• Land holding and other assets• Housing status• Food stock• Health status• Family planning/ contraceptive use• Sanitation• Credit and savings• Household expenditure and consumption• Coping with crisis• Women's ownership of and control over household assets• Women's access to income-generating activities• Women's mobility• Training from BRAC• Reasons for dropping out of the programme

The Proshika survey focused on what they call two 'generic types of empowerment', economic and social. For them this emphasis on empowerment impact has 'a subtle but significant difference from the usual meaning of impact': while it assesses actual changes in the economic and social circumstances of the people it supports, it also assesses their capacities and potential to improve their circumstances in the future.

Preparation and pre-testing

No matter how well an agency knows an area or population, there is always a need to pre-test the questionnaire in order to ensure that the questions are understandable, that they are put in a suitable way, that the length of the

interview is appropriate, and that respondents will be available. The questionnaire can then be modified before it is administered to a wider group. Addressing problems and mistakes at this stage is not only much easier and cheaper, but can make the difference between a successful or unsuccessful exercise. Given the time and expense usually involved in questionnaire surveys of this type, it is well worth investing a little at the outset to get it right.

BRAC researchers tested the draft questionnaire several times to make sure that it was right for collecting information on indicators. Several revisions were made: for example, they found that it was difficult for respondents to recall income, whereas they could remember expenditure more accurately. However, they also found that although people could quite easily recall large expenditures over quite long periods (such as buying bicycles or radios) the period of recall for small purchases, particularly of non-durable food items, was much shorter. As a result, a question on consumption expenditure was repeated on three consecutive days, asking people to recall expenditure over the previous 24 hours. It was also recognised that expenditure varied greatly over the year. As a result, 'data on consumption and non-durable expenditure were collected twice — once in the lean season and again in the peak season in order to consider seasonal fluctuation in consumption and expenditure' (BRAC case study).

Respondents

The choice of respondents is clearly a critical part of the design of a survey study. The case studies used a mixture of random and purposive sampling (see Chapter 3). In the Wajir, Proshika, and BRAC cases, random samples of direct beneficiaries and non-beneficiaries were made in order to compare these groups. Care was taken to select households in similar zones and/ or with similar characteristics.

In the Wajir case study, one of the key questions was the degree to which observed changes in the Wajir area were due to the pastoral development project run by Oxfam. They therefore undertook stratified or systematic sampling, illustrated in Table 4.3. Two hundred households were selected by first dividing the area into zones. Centres in each zone were then clustered to ensure an approximate balance with respect to distance from markets and different agro-ecological zones. Five centres were selected from these, three in project areas and two in non-project areas. In each of the five centres, 20 pastoral groups (living more than 8km away from the market centre) were randomly selected for interview. Within each group two households, a richer and a poorer one, were interviewed. In project sites the sample was further stratified according to membership of pastoral associations (about 50 per cent of the sample were members). In addition, a purposive sample of 224 interviewees was selected in order to explore the impact of specific restocking and credit projects.

Table 4.3: Number of households sampled in Wajir case study

	Members of pastoral association	Non-members of pastoral association	Total interviewees
Project sites			
Khorof Harar	20	20	
Riba	20	20	
Wargadud	20	20	
			120
Non-project sites			
Duntow		40	
Batalu		40	
			80
Purposive sample			
Families with restocked herds	40		
Women involved in credit scheme	184		
			224
Total	284	140	424

BRAC, on the other hand, also had to take into account the previous impact assessment exercise (IAS 1) that they had undertaken. This had identified 'panel members', which were BRAC member households, and 'comparison members', who were not members of BRAC used as a control group. For the follow-up study, BRAC took random samples from panel members, from other BRAC members, and from comparison members. Given the wide geographic spread of the programme, BRAC also needed to get a reasonable balance across the country without spreading its resources so thinly as to make the exercise an expensive logistical nightmare. Therefore, of the 1,250 BRAC sample households which were randomly selected, 500 were from districts served by ten area offices which had also been included in the first assessment, and 750 households covered by 15 other areas offices (see Table 4.4).

As in the Wajir study, BRAC studies also deliberately included specific groups of people in their sample about whom they wished to discover more. In this case, they were curious about 200 members who had shown very high performance.

Table 4.4: Sampling frame from the BRAC case study

	Members of BRAC		Non-members
	Panel members from ten areas	Non-panel members from 15 areas	Comparison members
Randomly selected	500	750	250
Purposively selected successful cases		200	
Total	1,450	250	

Selection and training of enumerators

Undertaking a large-scale questionnaire survey which involves face-to-face interviews normally means having to hire interviewers. The numbers of enumerators or interviewers needed will depend on the scope of the survey, how long it lasts, what geographic area will be covered, and what logistical support is at the disposal of the impact assessment team. Within the studies, enumerators undertook on average between three and five interviews per day over two to four weeks. The total number of interviews per enumerator varied from about 20 to 60. If enumerators are expected to undertake repeat interviews over a period, as was the case with BRAC, the number of interviews that can be conducted will be limited. Still, BRAC managed to carry out 1,700 interviews, all of them repeated, compared to 424 interviews in Wajir and 240 in the ENDA study.

In cases where findings may vary over time (for example, because of seasonal variations) care should be taken to minimise the time-gap in the collection of data. For example, BRAC’s emphasis on collecting accurate consumption data, which vary over time, meant that many small survey teams had to work simultaneously across a geographically dispersed area. This posed some problems, notably putting high pressure on the research team to supervise the data collection effectively.

In three of the case studies which used questionnaires, the enumerators were chosen from the project areas to make sure they were accepted by the community, and to avoid language problems. Proshika contracted out their survey to a local research organisation ‘to ensure impartiality’.

At the outset, between two days and one week was spent working with the enumerators on the following areas.

- briefing about the projects being assessed;
- sampling techniques;

- developing, testing, and refining the questionnaires and data-recording sheets, and translating them into local languages;
- interview techniques, especially how to probe for apparent inconsistencies. In the Wajir study this included some training on the basic techniques of ranking and seasonal calendars (see section on participatory methods) which were included to assist in quantification and recall;
- logistics, management; and supervision.

Analysis

Simple statistical techniques were used in the Wajir and ENDA studies to analyse the data collected. This generally involved calculating average results for various categories of respondent (rich and poor, adults and children, men and women, project and non project members) and presenting this as before and after data — see the example in Table 4.5.

Table 4.5: Perceptions of the relative reliability of water supplies now compared to ten years ago in Wajir

	Ten years ago	Now
Project site	3	7
Non-project site	7	3

Note: Average scores reported by 200 pastoral families

The Wajir, Proshika, and BRAC studies assessed the statistical significance of the results obtained. They particularly explored differences between project and non-project sites in order to determine the likelihood that these could have been caused by chance, rather than the projects. Their conclusion could then be applied to wider populations. In the Wajir study it was possible to calculate, for example, what the economic impacts on the 1,788 households in the project area were, if the statistically significant findings for the 200 households interviewed were true for the whole population (see the section on costs, benefits, and impact at the end of this chapter). Table 4.6 illustrates how this analysis was done. The average animal losses per household were given a financial value, which was calculated for both project and non-project sites. This established an average difference between these areas of £265 per household. If this applied to all the estimated 1,788 households living in the area, the lower mortality rates would thus translate into an estimated district saving of £473,820.

Table 4.6: Economic impact analysis of changes in animal mortality in Wajir

	Project site		Non-project site		Difference
	Mortality rate	Annual loss per household	Mortality rate	Annual loss per household	Annual loss per household
Camels	20%	£273	31%	£402	£129
Cattle	17%	£126	32%	£249	£123
Sheep or goats	18%	£49	25%	£68	£19
Total		£454		£719	£265

The BRAC study attempted to go one step further by employing more sophisticated statistical analysis such as multiple regression to try and isolate the effects of single variables on material poverty. They had to hire trained staff for data entry and cleansing and coding data, while processing and analysis was done by staff from BRAC's Research and Evaluation Division. This sort of analysis 'is enormously demanding in terms of data requirements, technical expertise and costs. It will only be feasible on very rare occasions' (Hulme 1997, p.7). BRAC themselves note that although these tools were used to separate the contribution of external factors from that of the project, this 'might not suffice because socio-cultural aspects can not be quantified'.

Having said that, exploring correlations within the data did reveal interesting gaps between the exceptional performance of some groups and lower achievements of others, which had not previously emerged.

[I]n identifying factors which influenced the differences in the level of success between the two groups [exceptional and less well performing groups] it was found that 61 per cent of the successful household members were involved in village organisation management against only 14 per cent for the other BRAC members. ... [Other] factors contributing to significantly higher level of performance [are] involvement in village organisation management, close kinship ties in the village organisation, involvement in BRAC training and enjoyment of special loan privileges. (BRAC case study)

Cross-checking results

All the case studies cross-checked their survey findings with information gathered in other ways, for example through participatory methods. In the

Proshika study, as described in Chapter 3, a systematic attempt was made to cross-check findings with a sample of the respondents, project staff, and senior managers. Each of these groups was asked to compare their expectations of the findings with the actual results. Nearly half the findings were questioned and further participatory research was carried out on specific issues.

Mini-surveys

Large-scale surveys can be inappropriate or one may have inadequate resources. In these cases an alternative is what some have called mini-surveys, covering a limited number of people or communities. These can be particularly useful if basic information is lacking.

Summary 4.1: How to gather quality data through questionnaire surveys

- Compensate for respondents who may drop out. Some respondents will be unavailable during the survey period, so it is often worth 'over-sampling' to compensate for this. In the Proshika study, 30 per cent more households than necessary were initially selected to make sure that they would have a margin for this contingency.
- Be aware of problems observed during data collection which might bias the results. For example, in the BRAC survey it was observed when collecting consumption data that in some cases household members took their meals outside the household. In order to prevent a potential distortion of data, missing data were replaced by average figures.
- Ensure that the questionnaire focuses on key issues. Questionnaires are most useful in answering questions such as *how many? what? how often? and when?* (*Why?* and *how?* questions are often better answered by less formal or structured methods.) Discussions with project staff and beneficiaries before these key questions are framed can enhance the focus of the survey. Consider to what extent the questionnaire not only seeks to identify what change or impact has occurred, but also seeks to verify the degree to which project outcomes might plausibly have led to those changes.
- Include questions that allow comparison of issues before and after project implementation. Simple ranking of people's perceptions — for example of a change in their quality of life, relative satisfaction with different service providers or NGOs, capacity to cope with crisis or security — allow comparisons to be made across groups and simple statistical analysis to be done.

- Give careful thought to the choice of the number and type of respondents. Particular consideration must be given to how representative a sample is of a given population, and a combination of project and non-project sites should be chosen. It should be borne in mind that a scattered location of respondents may make the exercise very expensive and place additional stress on field survey teams.
 - Recognise the importance of pre-testing and refining questionnaires. Before the survey is implemented, care must be taken in the selection of enumerators and their proper training and briefing.
 - Ensure adequate cross-checking and feedback. This not only verifies findings but can also reveal problems with the survey method.
-

Three of the case studies carried out some sort of mini-survey covering a single village (Ghana), a limited number of villages or communities (CYSD), or a number of staff members (Matson). These were done in different ways and involved questionnaires and participatory exercises. The objectives were to get a relatively rapid, but comprehensive, overview of the activities, opinions, and attitudes of given groups of people.

In the Ghana study, mini-surveys were conducted in each of the three villages involved in the impact assessment. They were used to explore the degree to which various household members were involved in activities promoted by outside agencies (NGOs, government, and others), through interviewing small groups of villagers or using existing project records. This method is dependent on acquiring an accurate list of village households and determining which member was involved in what activity. This proved quite difficult because household listings tended to be incomplete for various reasons (see section below on social mapping and wealth-ranking), project records were poor or ambiguous, and it was sometimes difficult to locate individuals within specific households.

Despite this it was possible, through a process of cross-checking during follow-up visits, to develop quite a comprehensive picture in each community of who was involved in which activities, and who was not. When cross-checked with other information — relating to well-being and ethnic and clan membership — a clear pattern of inclusion and exclusion emerged. A disproportionate number of socially dominant and economically better-off households were involved in development activities when compared to poorer households. This issue of exclusion has since become an important focus for discussion when the NGOs involved in the study reviewed the preliminary findings.

In the CYSD study detailed household surveys were undertaken in three of the 21 villages covered by the study in order to cross-check and quantify changes identified previously, for example from secondary sources. This confirmed for example that a substantial increase in land encroachment for paddy cultivation had occurred and that some households were increasingly involved in income-generation activities. The mini-survey also allowed a more detailed understanding of how and why these processes were occurring.

In the case of Matson, a short self-administered questionnaire was used towards the end of the study to get a 'wider cross-section of views from the staff than what was gleaned through conversations and interviews'.

Thus, small-scale or mini-surveys can be used at different stages of the impact assessment exercise for a range of purposes. In Ghana the survey was used to generate a general picture of a whole community on a specific issue, which the researchers could then explore in more depth. The mini-survey in this case acted as a means to ensure that any further interviews and research would include the views not only of the 'included' but also the 'excluded' population.

In both the case of CYSD and Matson, mini-surveys were used to confirm findings derived from other, more qualitative, research methods and to deepen their understanding of specific issues. In the Ghana study it enabled the researchers to gauge the *scale* of a phenomenon (the exclusion of some households from NGO activities) that had been identified by other methods.

Mini-surveys offer a useful alternative to large-scale ones, but because of their very nature care should be taken when generalising from them to larger populations. Many of the lessons noted under the large-scale survey section apply to mini-surveys as well. It should be remembered that it makes little or no difference to an interviewee if they take part in a mini-survey or a large-scale one. The onus is therefore on those designing the survey to ensure that every effort is made to keep the interview as short and as focused as possible, and that the quality of the results obtained mean that people's time was not wasted.

Interviews, workshops, and discussions

Simply talking, and listening, to people is probably the most common and useful way of assessing impact. This does not mean to say that it is easy. In the case studies these discussions happened in a variety of ways: individually and in groups; formally and informally; using pre-defined questionnaires, semi-structured interviews, and workshops; and simply by chatting. This section looks at two main areas: individual interviews; and group interviews, discussions, and workshops.

Individual interviews

As mentioned in the above section, face-to-face interviews using questionnaires were undertaken in a number of case studies. Most studies also carried out less structured interviews (sometimes called semi-structured interviews) with a wide variety of informants, as well as engaging in chance opportunities to exchange views and opinions with others.

Table 4.7 outlines some of the main distinctions between surveys and semi-structured interviews, although in practice they are often not as clear-cut as the table suggests.

Table 4.7: Advantages and disadvantages of structured surveys and semi-structured interviews

Structured survey	Semi-structured interviews
Design completed, questions defined, and duration set before interviewing starts	Design continues during interviewing phase; questions are identified and modified during the course of interviews and between interviews; duration varies
All informants usually selected before interviewing starts	Informants can be identified progressively, making use of findings from previous interviews
Identical questions for all respondents	Questions or topics tailored to different informants and stages of enquiry
Informants chosen as representative sample	Informants chosen to explore a range of different types of knowledge and perspectives
Data reduced to understandable patterns using statistical analysis	Findings reduced to understandable patterns using qualitative analysis
Findings validated by calculating the likelihood of the patterns observed being due to chance rather than specific factors	Findings validated by cross-checking and triangulation
Most suited to answering what? how many? how often? questions. Can answer why? or how? questions but tends to be less effective in doing so, as these answers are often difficult to quantify and compare	Most suited to answering why? and how? questions. Can answer what? how many? how often? questions but answers will usually not be valid for a wider population

The Matson study particularly emphasised the semi-structured interview approach; it deliberately set out to discover from a wide variety of people their perceptions of change in the community. The researcher used three basic questions to structure these interviews, which otherwise were open-ended: 'What are the changes that have taken place in Matson?' 'Has the quality of life improved or not over the years?' 'What do you think caused those changes?'

In order to carry out the interviews, the researcher needed help to identify a good cross-section of the community. In this case, an advice worker at the One Stop Shop¹ on the Matson estate used his rich understanding of the community to draw up a list of categories of residents to be interviewed: older people, youth, lone parents, long-standing residents and newer residents, residents who did not use the services offered by the project, board members of the project, project staff, other service providers, local businesses, and statutory service staff.

Project staff members were then asked to identify names of people who they thought would be willing to take part, taking care to ensure a balance between men and women. The advice worker painstakingly phoned each potential interviewee and set up meetings. By the end of the study, 28 formal interviews were conducted, half of them with men and half with women. However, some groups were more difficult to engage than others as Table 4.8 reveals.

It is important to note that efforts were made to contact groups which had proved difficult to involve in the formal interviews, such as non-project users and busy officials. The researcher stressed this:

there were also innumerable conversations with the staff, board members, and the residents who 'dropped in' at the project. Notable among these conversations were ones with the Matson bus driver over a cup of coffee at the end of his shift, with an elderly woman who volunteered to help me post my letter at the post office, with members of the Phoenix Club², with a resident who had just moved into Matson and was looking for change to make a telephone call ... But a major source of getting a feel of life in Matson was from the children. Having been invited to talk to the children at Robinswood school on India, which I coupled with a few magic tricks, I made some good friends among them ... hanging around the community shop I invariably bumped into them and got talking about India, Matson and magic!! Not necessarily in that order. (Matson case study)

The Ghana study also emphasised semi-structured individual interviews, carried out repeatedly over a one-year period. Their selection was made following participatory research in the community where a wealth-ranking exercise (see section below on participatory tools) had been conducted. This allowed a sample of individuals to be chosen who represented distinct, locally determined, well-being categories. The sample selected deliberately included a greater proportion of women and individuals from poorer households (Table 4.9).

Table 4.8: Interviews conducted in the Matson study by stakeholder

Category	Number of interviews	Comments
Residents	13	Interviewees represented all the sub-groups; five of them happened to be either staff or board members.
Non-users	No formal interviews	Many casual conversations at the bus stop, post office, telephone booth, chemist, and so on.
Board members	4	
Staff	10	Interviewees included volunteers and one student placement.
Local business	1	
Other service providers	5	
Statutory service staff	No formal interviews	Opportunities were taken at various meetings and functions arranged for other purposes to engage these and other officials, including city and county councillors, the local Member of Parliament, social service managers, housing officers, and so on.

Table 4.9: Breakdown of individuals selected for repeat interviews (Ghana study)

Village	Navio		Yiziiri		Demon	
	Men	Women	Men	Women	Men	Women
A (better-off)		1		1		1
B	1		1		1	
C		1		1		1
D	1	2	1	2	1	2
E/F (worse -off)	2	2	2	2	2	2
Total	4	6	4	6	4	6

These interviews aimed to explore people's perceptions of poverty and well-being and to trace how these have changed, to examine the effectiveness of those organisations whose work affects their lives, and to record their assessments of the sources of any change identified. Individual interviews have proved to be particularly useful in drawing out:

- sensitive information on intra-household relations, for example changes in gender relations within the household or between wives in polygamous households, and sensitive information regarding income, credit, and so on;
- differences and similarities between men and women, as well as between individuals from better-off and worse-off families, in terms of their perceptions of change and what they value;
- the impact of sudden shocks such as illness or theft on vulnerable people, and how these can dramatically affect their well-being;
- the importance of seasonality and the relationship over the year between food security, income and expenditure, and illness.
- Interviews complemented and helped to cross-check information obtained in other ways, for example through community research work, as well as revealing critical issues which, although raised before, had not been emphasised strongly enough, such as child and infant mortality.

Reflections on individual interviews

The importance of interview skills and attitude

The selection, training, and management of people who regularly conduct face-to-face interviews and discussions as a central part of their work is often inadequately thought about and resourced. The quality of the information generated through interviews is dependent on the relationship that interviewers manage to develop with the respondents, and on the following attitudes, behaviour, and skill:

- sensitivity to the respondents' mood, body language, and time constraints, and to the different cultural norms that may shape these;
- understanding that questions can be asked (and answered) in diverse ways;
- ability to really listen to answers, and to probe and cross-check in a thorough but sensitive manner;
- taking notes in a discrete, non-threatening way which does not interrupt the flow of conversation;
- using humour and personal experience to bring up sensitive issues or to challenge a response.

These skills are even more important for those conducting semi-structured interviews, where the interviewer must show a lot of initiative, compared to enumerators in large-scale surveys, who have less scope to ask spontaneous questions. Rather than selecting impact assessment interviewers only according to their technical competence or sectoral expertise, candidates' inter-personal skills should be given equal weight, perhaps by conducting 'mock' interviews. Clearly, the respondents' gender, ethnic identity, and language also plays an important part in determining the sex and background of the interviewers — although there are advantages and disadvantages to involving 'insiders' and 'outsiders' in impact assessment processes (see Chapter 3 on 'Who should be involved?').

As noted above in the case of enumerators, skimming on training and briefing can be a false economy, given that the quality of data will be, in large part, determined by the interviewers' skills and behaviours. Those conducting semi-structured interviews within impact assessment processes may require training in how to ask questions that help determine what has changed over time; how significant or valued different changes might be; and why and how such change has occurred. This may need to be combined with basic interview training regarding behaviour, body language, dress, and basic listening skills; when to use open and closed questions; how to probe and cross-check; and how to record and evaluate the interview.³

Last, interviewers must be given time and space to reflect on their experience and to share it with other colleagues. Managers should help them make this space and encourage such exchange.

Overcoming assumption and bias

Everybody holds certain assumptions and biases depending on their own experiences, upbringing, and training. Interviewers are no exception, and have to try deliberately to minimise the weight of their biases. Often, these have a clear gender dimension; for example, male interviewers can be unaware that they are only communicating with other men. It should also be borne in mind that interviewees make (sometimes false) assumptions about their own communities. In the Ghana case study, discussions with the interviewers led to the following conclusions and precautions.

- Interviewers can assume that respondents are able to estimate farm size, rainfall, yield, income and expenditure, and so on. However, this data should be cross-checked through observation, visits to fields, and with other information collected during the interview. For example, since farm size, yield, consumption, and income are all related, are the responses to these questions consistent? Is the estimate of land farmed, sacks produced, amount of production sold, amount of food bought consistent with the number of months during which food is available and the number of meals eaten daily?

- Interviewers should use their own knowledge and findings about a community to understand responses better, not to confirm assumptions and prejudices. For example, some development workers assume that spending money on funerals or other social activities is wasteful. They might therefore seek to discover what proportion of household expenditure is spent on this perceived luxury and make judgements about the household's willingness to invest in 'productive' activities. In fact, discussions with communities that probe this issue further often reveal how important such expenditure is, not only in conforming to social norms and customs, but also as an investment in community and social relations. These relations help to create the bonds and trust or 'social capital' that hold families and communities together and allow them to make claims upon each other.
- In writing up interviews, interviewers should be as faithful as possible to the answers given and aim to quote an interviewee's exact words in order to minimise the degree of misinterpretation. If interviewers add their own analysis and interpretation, this should be made clear.
- If there is more than one interviewer, it may be worthwhile to conduct a number of interviews together at the beginning of the process in order to harmonise the content and style of the interviews and to offer critical feedback to each other.

Probing and cross-checking

One difficulty in assessing impact is that you quickly notice that 'facts' are not necessarily reliable. People exaggerate (or lie). Even sincere people make mistakes, or cannot remember things accurately. So you have to ask every question to different people, to judge what the truth is even about simple figures. It is worth the effort. Also, your relationship improves with people if they find you are happier to know the truth [rather than being] told something they think you want to hear. (Pakistan case study)

Individual interviews often allow confidential and sensitive information to emerge, yet also increase the possibility of interviewees exaggerating, lying, or using the interview for their own purposes. This does not mean that individual interviews should be trusted less than other kinds of interview, which also have disadvantages (see below), but as discussed in Chapter 3, the interviewer needs to be aware of bias, and ask who might be exaggerating or lying, and why. This can provide important clues to the interviewees' interests and relations with other groups or individuals, and the project being assessed. Respondents may not want to seem impolite by saying that they do not know the answer to a question, or they may see the interview as an opportunity to negotiate with those doing the survey, for example to provide more or different types of support.

In the Ghana study, contradictions seemed to emerge regarding the responses of individual men and women about women's involvement in farming activity. Probing further with the respondents, and with other stakeholders, revealed that these differences could be explained in a number of ways. In one region, among one ethnic group, the gender division of labour was indeed such that women did not farm their own plots. In others they did, but men did not consider it worth mentioning because they either considered it as 'gardening' or as such a small-scale activity that it did not warrant a mention in comparison to men's 'family farming'. For women on the other hand, this activity was a vital source of independent income, and it therefore figured very highly in their accounts of income.

Probing and cross-checking inconsistencies or exceptions is critical. If something does not fit with other findings, it usually needs to be explored further, rather than ignored.

Being opportunistic and curious

Both the Matson study and the Ghana study demonstrate the importance of seizing unexpected opportunities to chat with people whose views might otherwise be missed. Interviewers' preparedness to 'hang about' in informal settings, such as drinking sorghum beer, proved very beneficial. The CYSD study provides another example: during an informal conversation at a tea stall the researcher came across a man from a certain village who was highly critical of the project being assessed and of the project staff. The researcher inquired further, and his informant offered to respond in more detail if the researcher met him at home. There, he explained how a member of another local NGO had persuaded some villagers that his agency could do more for them than CYSD. This had led to a split among the villagers and resulted in CYSD withdrawing from the area. However, this other NGO had done nothing for the village, causing great resentment. This information led CYSD to discuss this issue with the other NGO and local leaders, and helped them understand the background to what had looked like general political interference. This sort of information was the result of a deliberate strategy of hanging around and observing (see section on 'Exploring Relationships').

Summary 4.2: Key lessons about individual interview

- Training and managing interviewers. The selection of interviewers and researchers needs to involve not only assessing technical knowledge but also candidates' behaviour, attitudes and inter-personal skills. Within impact assessment processes, special training in interview methods and recording findings may be required

- Overcoming assumption and bias. Interviewers should assume that respondents can answer certain questions and then cross-check the answers, rather than assuming ignorance. They should use their knowledge about a community to understand responses better, not to confirm assumptions and prejudices.
 - Probing and cross-checking. If a piece of information contradicts other findings it should be explored further.
 - Being opportunistic and curious. Interviewers should be prepared to 'hang about' in order to seize unexpected opportunities to chat with people whose views might otherwise be missed.
-

Group discussions, workshops, and interviews

Group processes are a commonly used method at all stages of impact assessment. Bringing people together is often a cost-effective way of eliciting several people's views at the same time, which can also generate new insights and provide a forum for questioning and cross-checking individual opinions. However, care must be taken to ensure that the views of more powerful individuals do not dominate. This section explores some of the most widely used group exercises carried out in the case studies: village meetings and community workshops; focus-group discussions; and multiple stakeholder workshops.

Village meetings and community workshops

Village meetings and community workshops were undertaken for different purposes in a number of the case studies. In some they were used early in the process to introduce the work, and to gather general information about the community, its socio-economic environment, and the changes that had occurred over time. In others, it was used after individual interviews and surveys had been completed, to cross-check and feedback information and to explore impacts on the community.

The Ghana study, for example, undertook a series of community discussions in three villages, using a number of PRA exercises over a two- or three-day period. The information gathered was used to establish a village profile and to develop a sampling frame for selecting individuals for interviews. In this case researchers believed that community discussions were required to understand likely divisions within the community, as much as to gather general information. Understanding these divisions — by wealth, gender, age and ethnic/ clan identity — clarified whose opinions the follow-up work needed to capture in order to get a rounded view of perceptions of change in these communities. During the course of the community discussions, the facilitator often divided

participants into male and female groups in order to explore particular issues and to capture likely gender differences from the outset. This is obviously easier to do than for other, less visible differences such as wealth or clan membership.

These follow-up interviews with individuals were also used to cross-check findings from community discussions and to explore key differences within households (for example, between men and women, but also related to age and marital status). This helped researchers better to distinguish which kinds of information genuinely seemed to represent a consensus about the community, and which were more contested

The CORDES study involved a series of community exercises which involved almost all the village residents. They also brought together communities to discuss their relative difference. This was interspersed with more focused discussions with specific interest groups within the villages.

In the ENDA study a concern was raised that although interviews and questionnaire surveys identified impact at the household level, they did not articulate it at the community level. The team had discovered, for example, that those people directly involved in a seeds project could clearly articulate its impact. However, members of the community who received seeds from other farmers did not attribute changes in their life to the project, although the seeds had originally been brought into the community by the project. ENDA therefore thought that community workshops could be a forum for gathering data on the broader changes that had taken place in the community, and as a means to feedback findings to the community as a whole.

The experiences in the case studies point to a number of practical lessons in carrying out community research.

Preliminary site visits: These can be a vital part of preparing a community meeting.

The preliminary site visit served two main purposes. It allowed the impact assessment team to conduct interviews with the old local field staff. The visit was also used to become familiar with the community leadership. This bonding was an important preparatory step for the participatory work that was to be conducted during the next phase of the [impact assessment] project ... The impact assessment team were also able to note the logistics of working in the areas and have an appreciation of the conditions under which the project was undertaken. Having visited the area the team felt more confident that they could adequately discuss the project with former staff and have new insight into the project. (ENDA case study)

In addition a visit like this can ensure that future meetings are conducted at times that are convenient for the majority of community members, women in particular.

Key informants and gate-keepers: These individuals provide an introduction to and entry point into communities. Know-who, rather than know-how, is particularly important if a project has ended some time ago, or if the researchers are new to a community. In ENDA's case, former staff were the key informants.

[I]t was possible to get an entry point into the community [through them, which] is an important part of getting to know the community in Zimbabwe— people are more at ease if you know people they know. People are also more willing to talk to you if they know that you have already had a similar discussion with people they know. (ENDA case study)

Useful information on community history: In nearly all cases where this was explored, community meetings provided this, particularly on events of special significance (see section on time-lines below). Community history offers important information on the role of external actors and how this was perceived, as well as on significant collective experiences.

Exclusion of women, children, and other groups: Particular groups can be excluded either because they are absent from the meeting, or because they are present but do not participate. Both the Ghana and CYSD studies report that village elders, and sometimes active male youths, tended to dominate general meetings. Researchers can divide the group into different categories, or ask groups that did not participate to review and modify the work of the larger group, but this is not always possible or desirable. How exclusion is dealt with to a great extent depends on the degree of inequality between groups and how they cope with differences of opinion and conflict. This will vary enormously from place to place. However, as Table 4.10 illustrates, this is an important issue to bear in mind when considering the appropriateness of community discussions or workshops.

Focus-group discussions

One way to avoid some of the problems outlined above, and to explore specific issues in depth, is to hold a focus-group discussion. Focus groups are small (about six to 12 people) and either made up of existing groupings of people with similar interests or identity (for example, those receiving credit from Cordes; women involved in the Wajir credit scheme; or Oxfam-supported rose growers in Pakistan) or specifically chosen to represent a variety of opinions and backgrounds. Although some of the same types of problems as in community discussions may occur — such as dominant individuals dominating the discussion, the risk of this happening is reduced if the group shares a similar background and interests. A moderator or facilitator normally helps to keep the discussion on course, to clarify areas of agreement and disagreement, and to ensure relatively equal participation from all.

Table 4.10: Advantages and disadvantages of group discussions, particularly at the community level

Advantages	Disadvantages
<ul style="list-style-type: none"> brings together a sum of knowledge greater than any individual insight, generating new insights and mutual learning 	<ul style="list-style-type: none"> can allow dominant voices to be further legitimised and the voices of the less powerful to be ignored or undermined
<ul style="list-style-type: none"> opportunity for group or peer checking and verification 	<ul style="list-style-type: none"> can make the marginal and less powerful seem to consent through silence or presence
<ul style="list-style-type: none"> can raise problems not hitherto recognised as common, and lead to discussion of quite sensitive issues 	<ul style="list-style-type: none"> can drive sensitive issues underground
<ul style="list-style-type: none"> can lead to consensus-building and conflict resolution as differences are explored and discussed 	<ul style="list-style-type: none"> may produce unprioritised 'shopping lists' may expose, but not deal with, conflicts or further polarise people
<ul style="list-style-type: none"> can create synergy and new ideas through debate and dialogue 	<ul style="list-style-type: none"> may simply recycle entrenched views
<ul style="list-style-type: none"> can be fun 	<ul style="list-style-type: none"> can be boring and frustrating methods can dominate over content
<ul style="list-style-type: none"> can cost less than individual interviews 	<ul style="list-style-type: none"> may be seen as a necessary 'ritual' that must be gone through in order to get aid or support
<ul style="list-style-type: none"> may take less time than undertaking individual interviews and then cross-checking results 	<ul style="list-style-type: none"> may waste people's time even if it saves that of the interviewer or facilitator

The BRAC case study used focus-group discussions (with six to ten key informants) to explore impacts on women's empowerment, both for individual women and their family. The group discussed women's involvement in income-generating activities, their ownership and control over assets, perceptions of their own well-being, their economic dependence on their husbands, and their mobility.⁴ The findings were used, where possible, to complement and cross-check quantitative data from the household survey.

The CYSD study also made extensive use of focus-group discussions, recognising that village meetings were unsuitable for providing personal information (say, on family-planning methods or intra-household relations) or insights into sensitive issues (village conflicts, vested interests, politics, caste or class relationships, and so on). Focus-group discussions with members of a women's self-help group were felt to be one of the most efficient method in understanding their views on changes in women's empowerment.

For example, one discussion involving 20 women from two villages (Jamena-Munda Sahi and Konkali) explored differences in daily work patterns, changes in work distribution between males and females, and overall differences between the females of 1987 and today's females (see Table 4.11). The participants were asked to identify the reasons for these changes. Participants found this easier than answering questions such as 'what difference has the project made to your lives?'. The results can be considered a mix of project outcomes and impact, as well as changes brought by wider social change in India. However, it does provide a basis for starting to understand what change is the women in question consider significant, as well as to what they attribute that change.

Table 4.11: Perceptions of changing roles of women in Jamena-Munda Sahi and Konkali, India, generated by focus-group discussion

The females of 1988	The females of today
<ul style="list-style-type: none">• no compromise on the duties assigned to females• not much care given to pregnant ladies and children• less freedom to express their views in public in comparison to today's females• had one pair of sari as their only dress• not allowed to go out alone at night• suspicious and afraid about outsiders• not enough food for women after serving other family members	<ul style="list-style-type: none">• take more care in children's education and health, are keen to immunise their children, and attend pre-natal check ups• have more daily labour available for them• more concerned with having a clean environment and good health• more concerned about safe drinking water• now wear sari and blouse• males are now mingling with females in common places

Table 4.11: Perceptions of changing roles of women in Jamena-Munda Sahi and Konkali, India, generated by focus-group discussion (continued)

The females of 1988	The females of 1988
<ul style="list-style-type: none">• female education not considered important• more dowry received by the female in comparison to the present dowry• males took decisions in family planning, expenditure, and marriage of daughters• all females were having arranged marriage in the traditional way• had faith in traditional witches and quacks• were afraid of doctors, injections, and medicines• used to practice traditional family-planning methods with herbal medicines• did not have freedom to decide the number of children	<ul style="list-style-type: none">• are now engaged in productive ventures like paddy processing and vegetable cultivation• now cultivate kitchen garden and therefore eat nutritious vegetables• increasing importance of female education• are saving money for the marriage of their children (the amount of male dowry has come down and marriage expenses have gone up)• more cases of love marriages are observed among today's tribal females• are adopting modern family-planning methods• have more freedom and participate in community work• now have unity among themselves after the formation of self-help group (SHG), and the freedom and acceptance of the community to sit together in village meetings and express their views• in some cases males are coming forward to join hand with females in household duties• increasing need among the SHG members for liaising their SHG with group ventures

These examples, and other experiences in the case studies, reveal some further interesting points related to focus groups.

Focus-group discussions can stimulate analysis which promotes critical reflection on past changes and generates ideas for the future.

In the discussion above, women expressed the view that males still consider females subordinate and that some men in the village were trying to disturb the work of the self-help group, precisely 'because it has put a check on the subordinate position of females in decision-making and freedom of expression'. The women then discussed ways of dealing with this. Although some of the women went away 'unhappy because the meeting was not able to find a solution for checking male dominance in their society', this example illustrates the potential of focus groups to produce new insights and ideas.

Focus groups do not have to be confined to community groups. Project staff, donors, other NGOs or interested parties can constitute a focus group to explore impact from their perspective. In the CYSD study, for example, a focus-group interview of project staff investigated factors that, according to them, had reduced the impact of their work. This revealed that frequent changes of project staff had severely undermined continuity, disrupted innovative schemes, and reduced staff morale.

Focus groups can help reconstruct baselines and project histories. In the ENDA and CORDES studies, focus-group discussions were found to be a very effective way of discovering what had been done in the community during project implementation. Members of the group could recall a number of events and workshops, and the names of people who worked on the project. In most cases people could recall when the first meetings were held and who introduced the project to their area. In addition, as in the CORDES study, they were often able to identify the gap between what had been planned in terms of impact, and what actually occurred.

Focus groups can help to determine and assess progress against indicators. Focus-group discussions can occur at any stage of an impact assessment process. The CYSD study used them to determine indicators for specific activities, and found them effective.

[P]articipants expressed that this [was] the first time that they [were] critically analysing ... their own development in a group. They also were convinced about some changes and impact in their life style which they had not realised before. (CYSD study)

The Pakistan study also used focus groups to explore impacts and to rank them in order of priority (see section on ranking below). They made particular efforts to ask people about negative and positive impacts, and to explore what the discussion of past impacts suggested for indicators for the future.

Focus groups can help in dealing with sensitive topics. Along with other qualitative methods such as individual in-depth interviews, focus groups have proven useful tools in addressing culturally and politically sensitive topics that people may consider threatening, such as reproductive and sexual health (Jaswal and Harpham 1997).

Care must be taken about the length of focus-group discussions and the range of topics covered. Some of the studies talk of discussions lasting up to four hours and still not completing certain tasks. Experience suggests that even in the most stimulating circumstances, attention starts to wane and people get tired after about 90 minutes. If discussions are to last more than 90 minutes, make sure that people know this before the start, and consider breaking up the session to have drink or food. Do not overload the discussions; it is better to complete a few questions in depth than to rush through many superficially, leaving participants dissatisfied with the outcome.

Table 4.12 summarises some of the main advantages and disadvantages of focus-group discussions.

Table 4.12: Advantages and disadvantages of focus-group discussions

Advantages	Disadvantages
<ul style="list-style-type: none">• relatively efficient: more people's views can be obtained than in one-on-one interviews• provides quality control through people's interaction• allow people's views to develop in exchanges with others• allows an assessment of agreement or disagreement on a given issue• can occur at any point in impact assessment process• can be used with several stakeholders• did not have freedom to decide the number of children	<ul style="list-style-type: none">• only a limited number of key questions can be asked• needs very good group-facilitation skills• requires good note-taking• may reinforce existing power relations within the group or be dominated by a small number of individuals

Multiple stakeholder workshops and meetings

A number of studies brought together several groups to reflect on the impact assessment work and exchange opinions. In most cases this was done towards the end of the assessment in order to encourage feedback, to cross-check results with a wider group, to communicate the results to individuals or groups who had not been involved up to that point, and to debate what implications the results had for the future.

In the course of the Ikafe review process in Uganda — described in more detail in the next chapter — two multiple stakeholder meetings were held, each involving more than 80 participants. The first was held at the end of the initial field work and the second four months later, following further field research. It was held in order to hear how key policy-makers reacted to initial findings and to involve groups which had previously been prevented by insecurity in the area from participating. These meetings were felt to be necessary because of the complexity of the programme and the diverse range of actors in the situation, listed in the following.

- The refugee population which had immediate practical needs as well as long-term concerns about their uncertain situation;
- The host population which, although willing to provide shelter for their 'brothers' from Sudan, wished to benefit from the infrastructure being installed for the refugees as well as preserving their own security, cultural identity, and in the longer term, their land;
- a range of elements of the Ugandan Government such as civil servants and ministry officials, teachers, local councils, and the army (whose primary interest is national security and law and order in this volatile and tense region);
- the United Nations specialised agencies which have specific mandates regarding care of refugees and food distribution;
- NGOs which, although they have more flexible mandates, in this case were supporting specific parts of the programme in the areas of health-care, education provision, water and sanitation, community and refugee organisation, agro-forestry, and land allocation. These included Oxfam's local, refugee, and international staff, local and refugee staff from Action Africa in Need, and the Jesuit Refugee Service's local and international staff.

The findings of this case study had emphasised that the problems raised by different groups were all connected and therefore needed coherent solutions, which particularly required better co-ordination between

different actors. The review process had built up this analysis over time, and had, to some extent, strengthened the capacity of both the refugee and host communities to articulate their concerns to government and aid agencies. This moment of response by policy-makers and senior agency representatives allowed some explanation as to why certain problems and misunderstandings had arisen. Refugees and the host population understood that problems which they had blamed on individual agencies (particularly NGOs) were often due to wider problems that the NGOs could not solve on their own, and gained some understanding of what might be done to solve some of the problems raised.

In the Matson and Wajir studies, workshops were also hosted for a range of stakeholders towards the end of the impact assessment process. Both proved very useful in raising important issues, particularly about how various agencies need to work together to achieve impact.

As with focus groups, this process requires very experienced facilitators with experience of dealing with potential conflict. The whole purpose of multi-stakeholder workshops is to bring to the surface differences of opinions as well as agreement. Less dominant groups may need to prepare carefully in order to put forward a clear and consistent point of view, perhaps through a chosen representative.

However, in situations where there is a degree of tension or suspicion between groups, a meeting of this type would be counter-productive and could worsen relations. A great deal of thought should go into the possible implications of multiple stakeholder meetings for the future relationships between the groups concerned. If there is strong doubt about the likely outcomes, and if adequate facilitation and conflict resolution skills are unavailable, it may be better to avoid this type of meeting.

Summary 4.3: Key lessons learned about group discussions, workshops, and interviews

- Be aware of divisions and power relations within the group which may require breaking up the group into smaller groups or complementing group work with discussions with individuals. If you think that group work is liable to reinforce existing power relations it may be better to avoid it.
- Be aware of those unlikely to participate in group meetings and seek out their views. Carefully consider the timing and location of meetings so that it is possible for key groups and particularly women to attend.

- Under the right conditions group work can encourage new insights, mutual learning, foster solidarity, and build consensus. For impact assessment teams, it can be particularly useful for reconstructing project histories and the context in which projects evolved, as well as for cross-checking findings generated by other methods.
 - Group work of this sort requires sophisticated facilitation and conflict-resolution skills.
-

Direct observation

Although it is likely that direct observation played a part in all studies, only some of them specifically refer to observation as a method of impact assessment and to the results obtained in this way. In only one study (Matson) was direct observation the primary method of assessment.

We saw earlier the Matson study's emphasis on talking to people in a variety of ways and settings. In the following, Stan Thekaekara's describes three spheres of activity in his participation/ observation methodology.

Talking to people: attending various meetings provided the opportunity for conversations with 'officials' — which included city and county councillors, the local MP, social service managers, housing officers and the like.

Being a part of whatever was happening on the project: This involved spending time at the different sites, occasionally manning a reception desk, answering the telephone (a great way to get an inkling of the kind of relationship between the project and the residents — it was obvious that most of the residents were not only on a first-name basis with the staff but invariably were very clear about who exactly they needed to speak to in order to sort out a problem).

Sitting in on whatever meetings were taking place: There were different kinds of meetings that I attended. Meetings of the Board and sub-committees of the Board; review meetings of staff; meetings of city and county council bodies; meetings of Tenant's Associations and the Tenant's Federation; meetings with other Neighbourhood Projects, meetings of the Matson Forum, and meetings of the Neighbourhood Project Network. Some activities were followed through by attending all meetings relating to it that took place during the two months — for example a new initiative of the Project to try and provide foster carers from Matson. (Matson case study)

Building trust and rapport

A distinction is usually made between participant observation, where the observer shares in at least some of the activities or discussions that are being assessed, and assessments where the observer deliberately does not become involved in the situation under assessment. In reality the distinction between participant and onlooker is often somewhat blurred, because the onlooker's mere presence can influence the proceedings. The onlooker can therefore inadvertently 'participate' in changing the nature of what is being observed, and conversely, there may be times when the participant observer will be less involved in ongoing activities. In most of the studies there were occasions when participant observation was adopted, such as joining in informal discussions as they arose, and others where onlooking was more emphasised.

In either case the studies indicate the importance of observers being present, so that people get used to them and build up a degree of trust and rapport, but also the importance of trust and rapport as a pre-requisite for making observation a useful tool for impact assessment. In this sense observation, and particularly participant observation, is a means to build trust, as well as being dependent upon its creation.

The other element stressed in the case studies is how observation and engagement offer researchers a better understanding of the context in which a given project or programme is undertaken. For example, in the ENDA study

[d]iscussions were held with community groups wherever the team encountered the people. These groups of villagers were found at boreholes, gardens, shopping centres or on the street. The team had no influence on the composition of the groups. Meeting these groups had the immediate benefit of allowing the team to become familiar with the communities, an essential element for the participatory workshops that were to be held later on.

Cross-checking

Observation can be a useful way of cross-checking information. In the CYSD study, information on agricultural practices which the project had promoted suggested that some farmers were ceasing to use these practices, despite early involvement in the project. Direct observation of farmers' practices and cropping patterns confirmed this, which in turn led to informal conversations with farmers to discover why this was the case. This helped to identify different types of farmer and the particular constraints they face in adopting new agricultural practices, as well as their attitude to the support provided by the project.

Exploring relationships

Observation is a particularly effective tool in assessing the quality of relationships between individuals or groups. Both the Matson and the NK/GSS studies specifically used observation methods to explore the relationships between community organisations and their members, between the community and other actors (NGOs, government authorities, or donors), and between community leaders and other members. Researchers participated in or observed meetings between these groups and recorded the following:

- the degree and quality of participation of individuals and groups, including who was not participating or not even invited;
- the way in which different parties treated each other, and their ideas;
- the way in which conflict or disagreement between groups was handled;
- the degree of independent decision-making in various groups;
- the body-language of the participants and the physical setup of the meeting;
- the informal interactions before, during, and after meetings and during breaks.

Such techniques of observation are also particularly useful for observing relations within the household, and between different parts of a community, for example in relation to patronage, dependency, or ethnicity.

Gaining new insights

Observation can help to gain new insights or to discover things that people may not wish to reveal in interviews, or may be not asked about in surveys. For example, some of the village meetings organised in the CYSD study

had to be stopped ... due to interference by drunkards. These people had consumed liquor (made by mixing poisonous chemicals) which is prepared in the village itself. Later through informal conversation with some females it was learned that about 30 per cent of the males in this village spend 15 to 20 rupees daily on consuming liquor.

Through spending time with villagers, the observers got to know about a particular issue which they then followed up. Although the purchase of alcohol was likely to have an effect not only on family income but also on levels of domestic violence, the issue had not been raised as a possible obstacle to improved well-being or women's empowerment in formal meetings and interviews.

Lessons learned about direct observation

Like all other methods there are advantages and disadvantages to observation as research tool, some of which are summarised in Table 4.13. What is perhaps different is that observation is normally a part of all other methods and tools. Good interviewers, for example, will carefully observe the non-verbal signals that a respondent may give and adapt their questions in light of this; they will observe the environment in which an interview is conducted to see if there are signs that confirm or contradict what the interviewee is saying. For instance, in the Ghana study respondents stated that food security this year was much worse than last year. This was brought home to the interviewers when they observed how much more time than normal small boys in the village were spending hunting bats, small birds and rodents — a sign that the season was indeed very poor.

Table 4.13: Advantages and disadvantages of direct observation

Advantages	Disadvantages
<ul style="list-style-type: none">• can help to build trust and rapport between observer and group	<ul style="list-style-type: none">• dependent on level of trust between observer and observed
<ul style="list-style-type: none">• enables an understanding of a programme's or project's context	<ul style="list-style-type: none">• resource-intensive and therefore limited in terms of sample size
<ul style="list-style-type: none">• allows observer to be open and discover things no one else has paid attention to	<ul style="list-style-type: none">• observer can affect results
<ul style="list-style-type: none">• allows observer to see things which are so obvious to people that they may not report them to others	<ul style="list-style-type: none">• dependent not only on observation skills but also on inter-personal and communication skills of observer
<ul style="list-style-type: none">• can enable observer to learn things that participants are unwilling to share in interviews	<ul style="list-style-type: none">• open to bias and observer's selective perception
<ul style="list-style-type: none">• enables cross-checking of perceptions and opinions voiced in interviews or surveys	<ul style="list-style-type: none">• can be limited to external behaviours if the onlooker approach is adopted and is not used in conjunction with other methods

For the purposes of impact assessment, observation will particularly mean keeping one's eyes open to changes that are occurring within communities

and between groups of people. It will also mean exploring the unexpected or unusual, and the unsaid. An example from the Matson case study gives a flavour of this:

In a supposedly high-crime area I was surprised to find that the Project's One Stop Shop on Matson Avenue does not pull down heavy steel shutters at night. This, in spite of valuable computers and other things within. Instead a fragile wall of glass forms the shop front, completely covered with decorations in the form of job advertisements, painstakingly stuck on every day by John Boe and his colleagues Fe Clarke and Ciaran Murphy. When I asked them why they ran such a risk their reply was: 'All the kids hang out here at night — they hardly come by during the day. If we want them to see the jobs, then night it is.' What they did not say and bears thinking about is how people react to this kind of concern coupled with trust.

Because of his knowledge of the context (high crime rate) the researcher spotted something that did not fit with what might have been expected. In observing this, and then probing further, he gained valuable insights into the relationship between the project staff and youth in the area and started to understand some of the many reasons why unemployment and crime on the patch were declining.

Summary 4.4: Lessons learned about direct observation

- Direct observation although often critical for building rapport and trust, for cross-checking results, for generating new insights, and for exploring the impact of processes on relationships between people and between groups, is often not adequately thought about or referred to as a method of assessing impact.
- Like interviewing, the importance of observation demands adequate training and preparation. Any training or preparation needs to include discussion and advice on what and who to observe; how to observe — as participant or onlooker, overtly or covertly; when to observe — at particular events, times of the day, or seasons; and how to record and communicate observations made — in notes, on tape, on video.
- Direct observation can be resource-intensive, can affect results, and is open to bias. It needs to be used in conjunction with other tools.

Above, some of the main lessons about direct observation as a tool for impact assessment are summarised.

RRA and PRA/PLA tools

As noted in Chapter 2, participatory tools and methods have been part of most agencies' impact assessment approaches for several years. They were used extensively in the case studies, during group and individual interviews, in discussions and workshops. Most teams used the more extractive RRA tools (rapid rural appraisal) rather than the more participatory PRA or PLA (participatory rural appraisal/ participatory learning and action) tools. This section explores some of the specific tools used. It should be remembered that semi-structured interviewing usually forms a critical component of participatory research and readers are referred to the section above.

Many of the problems associated with individual and group interviews, described in earlier sections, are relevant here. The tools and methods presented below should always form part of a discussion or interview, and are therefore subject to the same concerns relating to the possible exclusion of some groups or to sanctioning the views of dominant groups. This is particularly the case in group exercises that provide a public forum which some may wish to exploit for their own ends.

Time lines and historical profiles

A time line, or historical profile, used during an impact assessment can help to capture broad changes in an individual's or a community's life. It can also stimulate discussions on how those changes occurred and what effects they had on people's lives. This is normally done by asking people to recall important events in the past and then to reconstruct history by adding other events and processes of change. This can be done using stones or symbols to represent specific events or using pen and paper. For example, the ENDA study concluded that

[t]he communities readily remembered droughts, and they could articulate the year that the drought occurred. Events could therefore be related to these periods. People in Zimbabwe also remember that independence was in 1980 and [this] is used as a standard reference point.

An example of the historical profile and contextual analysis produced at a workshop in Chivi is presented in Table 4.14. It indicates that 'the communities were able to distinguish clearly between the activities of the different NGOs and also indicate the years in which these NGOs were most active'.

Time lines have several uses during impact assessment. First, they allow a project to be situated in a longer historical time-frame. They can indicate the role and significance of other actors (NGOs, civil servants) and factors such as rainfall or important political events. Third they can give an indication of

positive and negative impacts and how these have changed over time. For the purposes of impact assessment, these issues are particularly important as they can help with a lack of baseline data and the problem of attribution (as discussed in Chapter 3). In addition they can allow any given project to compare how its impact is perceived in relation to others, and how its activities combined with those of others to achieve certain impacts.

Time lines are usually, but not always, a good way of getting to know a group or community in a relatively unthreatening way, indicating an interest on the part of the researcher in the community's own history. This may not be the case where mixed groups hotly contest historical events and how they are interpreted. In zones of recent conflict, and in communities where there is liable to be more than one interpretation of history, care should be taken in resurfacing potentially conflicting views. In such cases it may be preferable to undertake this exercise with particular sub-groups and to then compare results.

Table 4.14: Historical profile and contextual analysis for Madamombe (Chivi North)

Year	Event	Change	Impact (change in lifestyle)
1984	Danida assistance	Built a dam	Gardens sprang up and more water for cattle
	Madamombe clinic built	Health-care services easily available/ accessible	Better health for the community
	Secondary school built	More education for the people	Increase in knowledge base
1985	Tapped water available	Clean water	Decrease in the incidence of disease
	Toilets available	Better health	Conservation awareness increased
1987	Introduction of ENDA	Lessons on how to grow indigenous trees	Fewer malnutrition cases reported
	Drought	Yellow maize introduced by DANIDA Food for Work	Less privileged were cared for

Table 4.14: Historical profile and contextual analysis for Madamombe (Chivi North) (continued)

Year	Event	Change	Impact (change in lifestyle)
1989	Surveyors pegged out the road		
	Donors came to the community	People given items of clothing	Less privileged were cared for
1990	Zvishavane water project	Dams, wells, and fish ponds built	Clean water easily accessible
1991	Drought	Cattle, crops, and trees died	Shortage of drought power, money and water; increased diseases and soil erosion
	Drought relief	Food crisis eased	No people died
1992	AIDS epidemic	Increase in people dying and number of orphans	People are now more morally conscious
1993	Tarred road built	Increase in road traffic	Transport load reduced
1995	Government grain scheme		Sustenance for community
1996	Floods, ground frost	Crops destroyed, trees damaged	Food shortage; fewer trees
	Post office and suggestion box	Banking available	Life made easier
	Public call box	Means of arresting thieves, easy communication	
	Hospital expansion	Better health-care facilities	

Ranking

There are many ways of ranking or ordering data. In the case studies, some of the most commonly used techniques, or the most interesting for the purposes of impact assessment, were wealth ranking, problem ranking, impact ranking and performance ranking.

Wealth or well-being ranking

Ranking by wealth or well-being seeks to discover the criteria by which individuals or groups describe a given individual or household as more or less wealthy, or well-off, than another individual or household. It aims to gain a picture of who is better-off, or worse-off in a given group, and why. This is usually done by gathering the names of the individuals or households to be ranked on separate cards and then asking individuals or groups to place each card in piles with others of similar well-being. Questions are asked during, or at the end of, the process to try and determine more precisely the common characteristics of the various piles that emerge. The example from the Ghana case study in Table 4.15 indicates the kind of analysis that can emerge from this exercise.

The results of wealth ranking can be used in several ways for impact assessment purposes. It can allow a direct assessment of the degree to which different strata of a community are involved in a given project and thus what likely impact may have occurred. In the Ghana study, this revealed that particular groups and families seemed to be involved in all activities and others, usually the less well-off, were not. It therefore also provided a sampling frame from which to select individuals for interview who represented different levels of well-being in the communities concerned.

Second, by understanding the key criteria that people use to determine well-being, researchers can gain insights into the problems faced by people with different levels of well-being. It can point to a project's negative effects; for example, the promotion of a particular agricultural activity which demands high labour input may bypass families which see lack of labour capacity or a large number of dependants as a critical reason for their poverty.

Third, asking what changes have occurred in the situation of individuals or households over the past few years (which can be done by asking people to re-sort the piles to the situation ten years ago), can point out changing patterns of well-being and the reasons for increasing well-being or impoverishment. In the Ghana study it became clear that even for 'richer' households a sudden illness or theft of cattle could have a dramatic impact on levels of well-being.

Table 4.15: Well-being ranking in Navio, Upper East region of Ghana

Ranking Criteria	Better-off				Worse-off
	A	B	C	D	E
Income?	Satisfactory (salary)	Yes Satisfactory (farm income)	Insignificant	No	No
Food availability	All year round, plus surplus for sale	All year round, no surplus	Available for a greater part of the year, about eight months	High proportion of dependants	Food deficit
Land	Larger plots, fertile land		Smaller plots, less fertile	Not fertile	
Own fruit trees?	Yes	Yes	No	No	No
Own livestock?	More cattle, sheep, and goats	Fewer cattle, sheep, and goats	Sheep, goats and poultry	Goats and poultry	Poultry
Number of wives	More than one		One	One	One
Number of households in category	11	8	8	13	19

Dated April 1997

Fourth, wealth ranking can provide an important baseline for the future and offer a basis for prioritising areas of desired future change. Given that the lack of baselines was a particular problem for some of the case studies, well-being ranking provided some of them with much need information. Last, the mechanics of wealth ranking, in particular pile-sorting, can be used for many purposes such as comparing different villages, groups, or types of organisation. In Ghana, this technique was used to distinguish between the different members of the Northern Ghana Development Network. In other studies, such as the Action Aid Impact Assessment⁷ work in Vietnam, wealth ranking has been used with government officials. However, the studies faced a number of difficulties with this method.

Getting the initial list of individuals or households. It can sometimes be difficult to gather an up-to-date list of individuals or households in a community. Even where such lists exist from censuses or local government offices, they may be inaccurate and are rarely up-to-date. In the Ghana study, where such lists did not exist, social maps⁶ were drawn from which lists were then developed. However, in some cases further research revealed that some households had been deliberately excluded by those drawing the map because they were not considered part of the community. Once again, cross-checking and keen observation can help to minimise such problems. Experience from other parts of Africa⁷ suggests that involving women, who may be better informed about the community, is critical, as is being aware that certain categories of people often tend to get 'forgotten': older people, disabled people, and the families of women who are second or third wives in polygamous households.

Unwillingness of people to categorise or label neighbours and friends. In some, but not all, cases people were unwilling to classify their neighbours by wealth or well-being. In some studies (CYSD) the issue was therefore depersonalised, so groups were asked to describe the characteristics for example of a 'wealthy man'. This is consistent with other experiences of participatory impact assessment, such as the evaluation of Calcutta Slum Improvement Project⁸, which asked respondents to think of the slum population according to categories of well-being and then to assess the approximate income range for each category as well as the proportion of the population that fell into each.

Problems of defining households and associated gender relations. The household has usually been the main unit of analysis in wealth ranking, which creates several problems. Definitions of what constitutes a household vary between places and groups. Moreover, the household is not a static entity: its membership may change suddenly or over time; at times people may include relatives who have migrated or children who are being temporarily looked after, at other times they may not. But perhaps the most problematic aspect is that ranking households tells us little or nothing about well-being within the household. Women often prioritise different aspects of well-being than men and, if care is not taken, their views may easily be ignored or subordinated to those of males. The result can be a wealth ranking of households which reflects only men's criteria of well-being, and which conceals differences in well-being within those households. Getting men and women to rank households separately is one way of overcoming the first problem and can reveal important differences and similarities — this may be preferable to having one mixed group doing the ranking, which was the case in most of the case studies.

In order to understand differences in well-being within the household, other tools and methods such as individual or focus-group discussions are

probably more appropriate. Although one of the studies planned to undertake well-being ranking of some individuals within households, this proved to be extremely difficult, because the problems getting an initial list of individuals were even greater than getting one of households, and people were even less willing to label individuals' well-being than that of households.

Problem or preference ranking

Problem or preference ranking is simply a means of understanding people's priorities. This technique usually involves asking people to list their problems or preferences for change, which are then ordered in some way. This can be done by simply ranking them from most important to least important. A more sophisticated method is to weight problems by assigning them a score – for example, by asking participants to distribute a certain number of stones, beans, goat droppings or points (whatever is handy or understandable) between the items listed. This gives a better understanding of the relative priority between different problems. Table 3.13 provides an example from an Action Aid study in Andhra Pradesh in India, in which the ranking was done by men, women, and young people, separately and then together. This indicated how important it is not to assume that these groups have the same problems or priorities.

Impact ranking

This is a variant of problem ranking developed in the Pakistan study. In this case, groups of project beneficiaries were brought together to identify a full range of positive and negative impacts, using a checklist of possible dimensions of impact (see Table 3.2 for details). Then, either individually or as a group, people were asked to rank those impacts by order of importance and explain their decisions.

In one of the projects in the study — a micro-credit project supporting women's goat-rearing activities which had been assessed as successful in terms of its economic impact⁹ — the women involved felt that the impact of the project on other aspects of their lives was as strong, or stronger. For instance, women valued the new decision-making powers they had gained and their ability to offer a goat at the festival of Kurbanī Eid.

When we meet together, we talk about the project and about our goats. We discuss when to sell them and how we can get a good price. Other money in this house goes into man's hands. We use this money to buy jewellery, dishes, and clothes. Men don't ask us how we spend our money.

Before we did not sacrifice goats; now we sacrifice them. We give them, and now God will give us something. Before, we worried that God would be angry at us for not sacrificing. There is a benefit at Ramadan times too. Before, we cooked black tea. Now there is milk, so, to open our fast, we can cook tea with milk. (Pakistan case study)

In addition to these specific examples of impact, the Pakistan study recorded how individual group members ranked wider dimension of change. Table 4.16 shows the results from one exercise in Ravat Goth, Pakistan (for an explanation of the categories see Table 3.2).

Table 4.16: Individual impact rankings in Ravat Goth, Pakistan

Name	First	Second	Third
Amnud	Religion	Relationships	Participation
Qaim Khatun	Knowledge	Participation	Relationships
Hakim Zaid	Relationships	Knowledge	Art
Kareemat	Life	Knowledge	Participation
Shahzadi	Knowledge	Participation	Relationships
Mumtaz	Participation	Knowledge	Life
Sabahi	Participation	Knowledge	Life
Zadi	Knowledge	Religion	Participation
Sabi	Religion	Relationships	Participation
Sakina	Life	Religion	Relationships

Table 4.16 indicates that even among a small group of women from the same community, receiving the same support for the same activity, different value is assigned to the impacts achieved. So although we can say that most of the women considered the knowledge they gained from the project either the most or the second most important impact, for three participants (Amnud, Sabi and Sakina) the impacts on their spiritual life, relationships, participation, and life (in the sense of health and material security) were all valued more than the knowledge gained.

In this case participants had to rank impacts according to predetermined categories or dimensions. While in theory this makes comparison easier, it also groups dimensions in ways which may not be evident to participants. In addition, the very fact that the data can then be averaged runs the risk of creating scores which squeeze out the diversity and difference which the individual scores retain. If both are presented in the report, the reader at least has the chance to interpret what had been lost in the aggregation process.

Performance ranking and satisfaction matrices

The performance of development organisations or government services, and the degree of satisfaction with a given project or agency can also be ranked or scored. As others have noted:

When hogs, figure skaters or bodybuilders compete, judges assign cardinal scores to subjective criteria: quality of coat for hogs; artistic impressions for figure skaters; and muscle tone for bodybuilders. Grades for academic papers are another familiar example: a professor's subjective evaluation of a humanities paper is given a cardinal score. In each case these subjectively assigned scores are added, averaged and tabulated in ways only appropriate to cardinal data. This means that judging requires training to achieve this level of inter-subjective agreement. For instance judges of livestock contests are occasionally judged on the degree to which their subjective judgements conform to those of established judges. (Insham, Narayan and Pritchett, quoted in Alkire 1999)

As this quote suggests, there are agreed standards upon which subjective judgements are based. This agreement has often been lacking in qualitative assessments undertaken by NGOs. Oxfam staff tried to overcome this by discussing what was meant by, for example, empowerment or participation, and how it was observed, before judgement was made about the impacts of a number of different projects. In the Pakistan case study, as we saw in Chapter 3, Oxfam programme officers agreed a scale for scoring participation which ranged from '0', indicating no awareness of project partner's activities by beneficiary community, to '5', where the community is working in more than one sector, taking initiatives and is able to mobilise its own resources. First individually, then together, they scored the projects and organisations involved in the impact ranking process, based on agreed criteria for each variable.

However, people will have their own (often differing) views about projects and organisations with which their communities are involved. It is therefore usually not possible or desirable to achieve the same sort of agreement. However, it is possible to compare how satisfied a range of groups are with a given situation. The example below comes from the evaluation of the Calcutta Slum Improvement Project mentioned above. In this case, residents from six slum areas separately identified key criteria for a satisfactory environment and then scored their level of satisfaction with each from 1 ('not at all satisfied') to 5 ('very satisfied').

The groups were asked to list criteria by which they defined the slum environment. These were written on cards which were placed along a vertical axis. Afterwards, a horizontal axis was drawn with the numbers 1 to 5 along it ... [t]he group then placed seeds against each criteria. After completion, a discussion was facilitated where reasons for decisions were noted ... Criteria

related to project interventions could then be assessed for how well they had been met, whilst new criteria gave an idea of possible future interventions or areas for community action. (Kar et al. 1997, p.85)

The results from the different slums was compiled into Table 4.17 which indicates overall levels of satisfaction across the slum communities, indicating areas of high satisfaction, for example with the electricity situation, mixed satisfaction with the provision of children's sports and games, and low satisfaction in areas such as unity and women's income.

Table 4.17: Satisfaction matrix from a participatory impact assessment exercise in Calcutta

Criteria	Extent of satisfaction				
	1	2	3	4	5
Road			2	3	1
Drainage	1		1	1	3
Tap water supply	1		2	2	1
Latrine		1	1	3	
Electricity				1	4
Primary health-care		1	1	2	1
Children's education	1			3	2
Children sport and games	1		1		1
Reduction in women abuse			1		
Reduction in alcoholism	1		1		
Cleanliness		1			
Women's income	1	1			
Women's awareness/ rights		1		1	
Economic earnings			1		
Political conflict		1			
Unity		2			
Children's safety		1			

Definition: 1= not at all satisfied, 5 = very satisfied

Numbers in boxes indicate frequency of response in different slums

Source: Kar et al. 1997

What is useful about this sort of approach is that it allows the reader not only to get an overview of the situation, but also to see where differences of opinion exist and which criteria are important in different slum areas. Community groups' expression of dissatisfaction — with children's safety, political conflict, or women's income — is made visible rather than subsumed into a picture of averages. This, along with other elements of the study, led the researchers to question whether existing project indicators should be changed to reflect more closely people's own satisfaction criteria. Further, it indicated that although most slums were relatively happy with the outputs that the project had delivered — such as roads, drainage, electricity, health-care and education services — there were also a number of other areas which affected people's overall quality of life (an improvement in which was one of the project's desired impacts) which they felt less good about — such as alcoholism, unity, and women's rights. This helps us to understand what change can be attributed to the project, and also what changes the project might seek to promote in the future.

Lessons about ranking

Overall, the types of ranking described above are powerful means of comparing a wide range of topics — well-being, problems, organisations, projects, impacts, performance, and satisfaction — not just by ranking these hierarchically, but by describing their relative importance. As we have seen, there are also ways of retaining the diversity of opinions that are gathered through ranking, and comparing and contrasting these to understand where different people and groups seem to agree or disagree.

However, I have also noted potential problems associated with ranking, particularly wealth or well-being ranking, and emphasised the danger that ranking data by its very nature allows the use of the standard array of statistical tools, which may be inappropriate or misleading. Ranking exercises are subject to the same problems (and benefits) outlined in the section on group interviews, notably the danger of manipulation by dominant groups, particularly men.

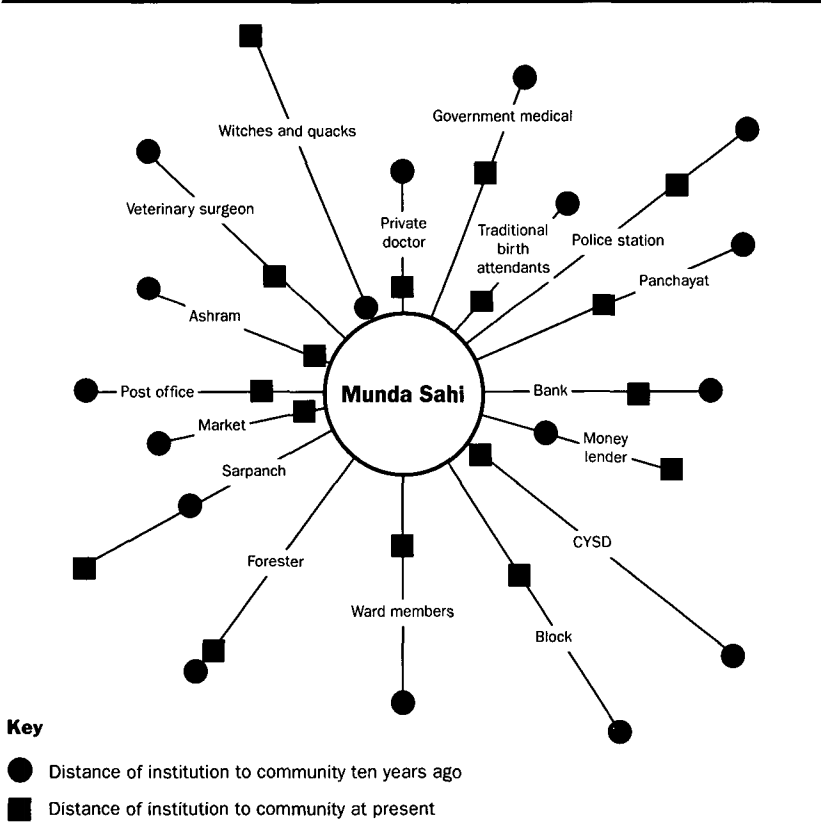
Chapati or Venn diagrams

Venn diagrams were used in the case studies to capture respondents' perceptions of the perceived importance of various institutions and of their relationship with them. They are sometimes called chapati diagrams after the circular Indian flat bread.

In the CYSD study respondents were asked to draw circles, the size of which represented the importance of each institution to them. The relationship with each institution was reflected by the distance between circles and the centre, which represented the respondents' group. So a circle placed close to the centre indicated a close relationship with that institution. The respondents were asked

to do this for the current situation, and as it was ten years before, but the complex nature of the task made it quite difficult. The compromise was to specify the 'before' and 'after' relationship but not to include the relative importance of the institutions.

Figure 4.1: Chapati diagram tracing 'before and after' relationship of key institutions/ stakeholders with village Munda Sahi (CYSD case study)



The results are illustrated in Figure 4.1. For instance, the exercise revealed that in the health sector, traditional birth attendants and private doctors now had a closer relationship with the community; government medical centres remained quite distant; and 'witches' and 'quacks' now had a much more distant relationship. In the economic sector it seems that the market and banks are assuming greater importance in people's lives, while the relationship with money lenders is declining. This exercise and the discussion it generated allowed CYSD to explore with the communities not only the outputs of its

health and credit activities, but also how these might have contributed to changing relationships and patterns of behaviour at the community level.

These diagrams can play a useful role in stimulating debate about relationships between groups, and the CYSD study indicates how this can include attempts to explore change in relationships over time — a crucial aspect of impact assessment. However, in other cases, such as Ghana, chapati diagram exercises were found to be confusing for people.

Impact flow charts

Flow charts depict the flow or direction of a particular activity or process. They typically start with an event or an action, such as planting a seedling or setting up a group, or a problem, such as lack of rain, and then explore the consequences. Usually this is done by asking ‘what happened next?’ or ‘and what did this lead to?’ or ‘what effects did this have?’. The impact flow chart is the opposite of the ‘problem tree’ method which seeks to discover the causes or roots of a problem in that it investigates the effects of a problem or action.

The Ghana case study effectively used impact flow charts by asking men and women, separately, to cite the most important change event in their community in the past few years. The facilitator then encouraged them to explain the knock-on effects of this events through the use of a flow chart. The results are presented in Figures 4.2 and 4.3.

Figure 4.2: Impact flow chart of most important intervention in Yiziiri (men’s opinion)

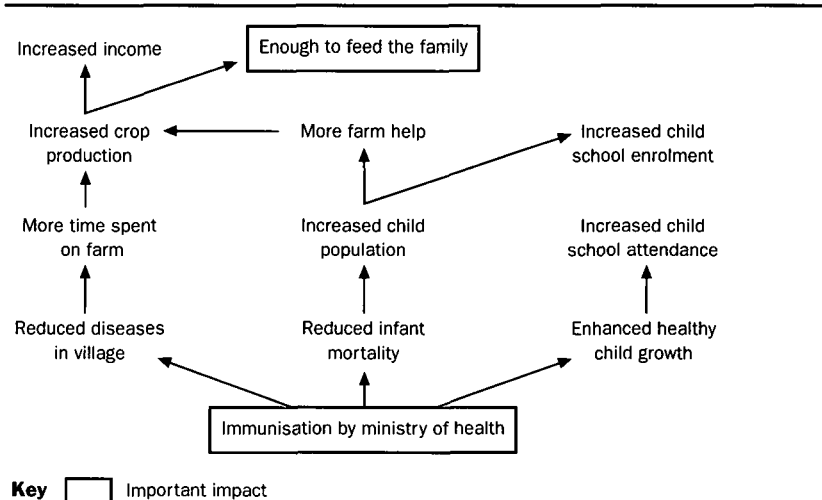
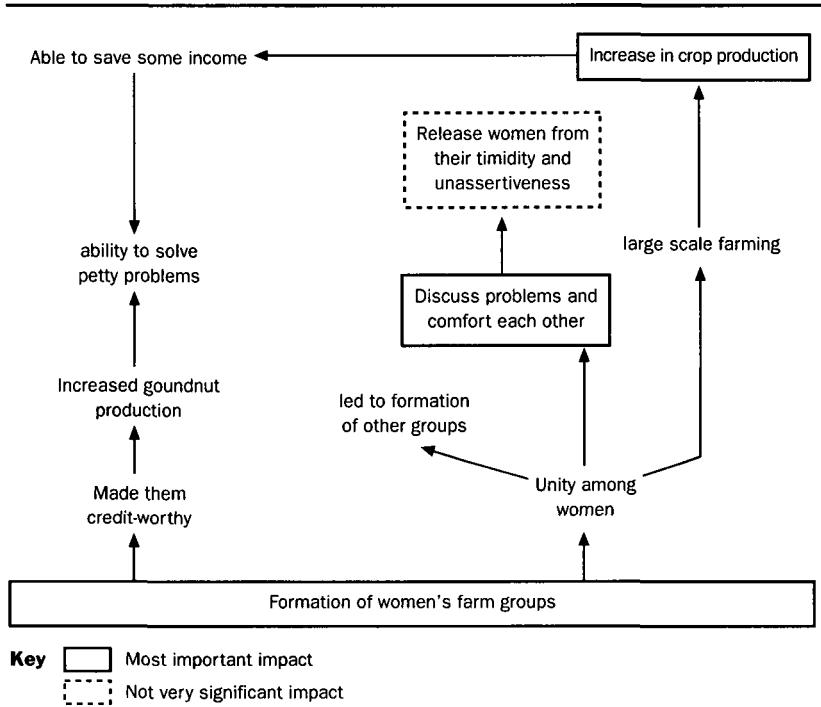


Figure 4.3: Impact flow chart of most important intervention in Yiziiri (women's opinion)



There are several interesting things to observe about this exercise, not least the differences between men and women's views of what is important. Women emphasised that the formation of their own group had provided them with a degree of unity and a forum to discuss problems, and that this allowed them to achieve things together, which had individual benefits in terms of economic gain and increased self-confidence. Men on the other hand stressed how the immunisation campaign by the Ministry of Health had improved child health. This had positive effects on education through higher school enrolment, but more significantly for the men, by increasing the availability of farm labour it contributed to higher production, income, and ultimately food security.

The impact flow charts also proved helpful in understanding how people saw the process and sequence of change, rather than how the project envisaged it, and where these two perceptions overlapped. For example, the women saw their unity as a key stepping stone to several other things, whereas the project placed much greater emphasis on the provision of credit. Being aware of such potential differences is crucial for impact assessment

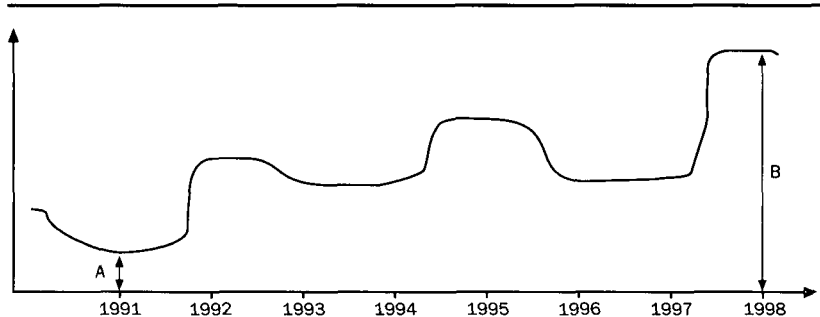
exercises: communities and NGOs may have agreed to various projects for different reasons, which may not be clear at the outset. This may become clearer if impact assessments include opportunities for various groups to explain the importance to them of the effects of a given intervention, in a way that is not 'led' or predetermined by the intervention's declared objectives. If an activity's potential unplanned consequences are not explored in this way, there is a danger of mistakenly identifying the obvious project inputs, such as credit in this case, as the critical success factor behind a given impact, when in-depth research makes it quite clear that the combination of credit and group formation was the determining factor.

Trend analysis

The analysis of trends and change over time lies at the heart of impact assessment. Some of the simplest examples of these tools are the 'before and after' or 'ten years ago and now' exercises illustrated in Table 4.5 and 4.11. These show perceptions of the changing reliability of water supplies in Wajir and perceptions of changing gender roles in Jamena-Munda Sahi and Konkali, India. Respondents were either asked to simply describe changes or asked to score them using simple 1-5 or 1-10 scales. Such exercises, used in several case studies, have proved simple yet effective in adding a dynamic element into the analysis.

However, comparing only two points in time in a 'before and after' exercise runs the risk of fluctuations between those points being concealed (see Figure 4.4). This can lead to exaggerating or underestimating changes that have occurred, and/ or being unable to distinguish between changes brought about by the project and changes caused by other factors. Other exercises used in the case studies therefore tried to capture trends using data over a number of years.

Figure 4.4: Showing how measurement at only two points can conceal large fluctuations between those years



The example in Table 4.18 from Ghana illustrates how this can be done. Respondents assessed the relative improvement or deterioration of certain factors over time by assigning scores from 1 to 10 (the factors had been defined in the well-being ranking described earlier in this chapter, and in other exercises such as an historical profile). They assigned scores for every other year over the past eight years. What emerged from this exercise, despite the rather idealised characterisation of 1989, was the devastating effect of a serious conflict with neighbouring groups in 1993/4, which not only led to large-scale loss of life, but also seriously disrupted economic and social life. In 1997, NGOs and government made strong reconstruction efforts particularly in the health and education sectors, but very poor rainfall affected food security and water availability. Clearly if an exercise had simply looked at 1989 and 1997, the significance of the conflict and the recovery from it may not have been adequately taken into account.

Table 4.18: Trend analysis in Demon village, Northern Ghana

1989	1991	1993	1995	1997	Issues
10	8	6	4	1	Water
1	4	2	4	10	Health
10	4	1	1	1	Food security
10	8	1	2	8	Peace
3	4	6	6	10	Rainstorm [high winds]
10	6	1	1	2	Income
10	8	1	2	5	Sheanut processing
10	5	1	5	2	Fishing
10	6	1	2	3	Livestock
2	2	1	7	10	External interventions

Key: 10 = good, 1 = poor

April 1997

For the purposes of impact assessment, trend analysis can give an insight into change over time and suggest causes for fluctuations. This in turn allows researchers to include factors outside the project or programme in the assessment which have had a bearing on people's lives (in this case, conflict

and poor rainfall). Situating a project in a wider context helps to minimise any exaggeration of, as well as underestimating, project impacts.

Trend analysis thus provides a useful tool in looking at change and attribution. However, it is clearly better suited to some situations, issues, and groups than others. Apart from obvious points — such as farmers being more likely to remember trends in weather or crop production over time than city dwellers — it is wise to remember the following.

People tend to remember best those things that are most important to them. Trend analysis should probably focus on these. For example, the Ghana study looked at trends related to the criteria which had emerged during the wealth-ranking exercise and which respondents themselves deemed critical to well-being.

People tend to be able to recollect particular memories when they can place them in the original context. Memories are often triggered when people recall the place they happened, the people who were present, or the political or environmental climate at the time. As the historical profile in the ENDA study indicates, drought years or the date of independence are common reference points. It is sometimes better to start trend analysis from these points and work forwards, or backwards, from them.

People tend to remember relative change rather than absolute change. They may also be more willing to disclose relative change. Rather than asking people what their income was this year and what it was last year, it may be more useful to inquire whether their income has increased compared to last year. Another method is to ask people to assign a score or weight to a given year's income and then to do the same for the following or preceding year, and so on. In this way, as in the Ghana example above, an insight can be gained into not only if a given year was better or worse for someone, but also relatively how much better or worse.

People may deliberately or accidentally not tell the truth or omit information. As noted in the section on interviewing, they may have very good reasons for doing so! The notion that people cannot or will not undertake this sort of analysis, or that their opinions are 'subjective' and therefore less valid, has meant that this often invaluable source of information has not been tapped. But its flaws do not invalidate the tool; they merely underline the importance of cross-checking (see Chapter 3) and seeking to understand why people might lie or exaggerate, and the need to build an atmosphere of trust and mutual understanding in which might truths can be shared openly.

Other participatory tools and methods

Other tools such as mapping, transect walks, pie charts and other diagrams, were also used in the case studies, although not as commonly as those detailed above. Some of the case studies used video material – CYSD, for example, looked at tapes which had been filmed in previous years to learn about changes in the density of forest cover.

Indeed the use of participatory video, role play, and drama in evaluation is becoming more commonplace (Braden 1998, Mavro 1997). Initial experience suggests that both can be useful in helping people tell their story, in ways that they wish it to be told. So in theory participatory video allows people to prepare, view, and adapt their representation of themselves and then communicate that to others, uncensored by intermediaries. Drama can allow people to express opinions and judgements about projects or organisations that they could or would not otherwise reveal. This is an area where much more can be done and learned.

Participatory tools — some issues and lessons learned

Participatory tools are increasingly being used in evaluation and impact assessments. At the same time a number of observers and practitioners are expressing doubts about the quality of some of this work, its potential to exclude certain groups, the way that it has become an orthodoxy or ritual to be followed, and the way that it can become dominated by tools, methods, and exercises at the expense of the content (Mosse 1994). The experience of the case studies suggests that there are some critical issues for ensuring the credibility of participatory research.

Recognising power and social relations. As the section on group interviews in this chapter suggests, participatory exercises in groups can neglect some people's views (for instance, women's or children's) and, moreover, validate and legitimate the views of dominant groups, thus increasing their power vis-à-vis others. Some of the studies revealed that minority clans or ethnic groups were excluded from meetings and even 'forgotten' during exercises such as wealth ranking. However, the studies also revealed that where careful attention is paid to this issue, participatory tools and methods can be used to reflect differences of power, opinion, and perception, as well as analyse the relations between different groups. This must be a deliberate strategy and part of the research from the outset.

Respecting people's time. Participatory exercises and especially 'community PRAs' can take up a lot more of people's time than individual interviews. This is particularly the case if a whole range of exercises is to be

undertaken in a short intensive period, because this suits the researcher. The danger of this happening is particularly acute when the tools and methods start to dominate — the very diversity of techniques and exercises that have been developed compounds the problem. It seems, as a recent Action Aid confirmed (Goyder et al. 1998), that more thought needs to go into not only prioritising what information is particularly important, but also into choosing tools which are relatively efficient in providing it. The experience of the case studies suggest that one should focus much more clearly on those aspects of change which different groups perceive as most important, to compare and contrast these, and to explore the causes of those changes.

Combining participatory tools with other methods and sources of information. The dissatisfaction with conventional research and evaluation approaches, and the appeal of participatory methods have sometimes led to proponents of participatory methods failing to build on and utilise the many years of experience accumulated in the study of economics, science, sociology, and anthropology. This is in some cases due to the theoretical debates which lie at the root of the different disciplinary traditions, but may also be due to ignorance and prejudice. From a practical point of view, it is a waste not to be prepared to use other methods, to learn from them, or to use information collected by other methods. Such information provides important opportunities to cross-check and verify findings, to avoid wasting resources and time, and, equally importantly, to question the validity of data collected in other ways.

Developing clear standards and criteria for processes of participatory inquiry. A number of 'criteria of trustworthiness' have been suggested for participatory processes (Pretty 1994) which include length and depth of engagement of actors; persistent and parallel observation; cross-checking of sources, methods, and investigators as well as participant checking; analysis and expression of difference and negative case analysis; evidence of searching out different views and explanations; journals and peer review; and impact on stakeholders' capacities to know and act. The case studies covered in this book and recent findings from Action Aid's participatory impact assessment work recommend that one should also think about balancing the 'length and depth of engagement' with a process or schedule agreeable to both community groups and researchers, and make more efficient use of existing sources of information. Methods should evolve and be adapted based on a mutual analysis of their strengths and weaknesses; and one should monitor the extent to which the information gathered actually has an impact — for instance, whether it leads to change in the policies or practice of the project or organisation being assessed and whether it is meaningful to all parties.

Summary 4.5: Key lessons about participatory tools

- Participatory exercises in groups can ignore some people's views, and legitimate the views of dominant groups, thus increasing their power. However, if used with care, participatory tools and methods can reflect differences of power, opinion, and perception as well as analyse the relations between different groups.
 - Participatory exercises can take up too much of people's time. Focus on those aspects of change that different groups perceive as most important, comparing these, and exploring their causes.
 - Participatory tools need to be combined with other methods and sources of information. By using secondary sources and key informants time can be saved and findings cross-checked and verified.
 - There is a need to develop clear standards for processes of participatory inquiry which balance the length and depth of engagement with a schedule agreed by both community groups and researchers; efficient use of existing sources of information; the development of methods based on a joint analysis of their strengths and weaknesses. One should also consider the extent to which the information gathered will have an impact, such as changing policy or practice of the project or organisation being assessed, and the extent to which it is meaningful to all parties.
-

Case studies

Case studies are not really a tool or a method, but rather an approach to gathering comprehensive, systematic, and in-depth information about a case of interest. They typically involve multiple research tools and methods, many of which are described above. Case studies can examine, among other things, individuals, communities, events, time periods, programmes, or organisations. Regardless of the unit of assessment, case studies generally seek holistically to describe that unit in depth and detail, to set it in context.

Case studies are particularly valuable where broad, complex questions have to be addressed in complex circumstances. In these cases the number of variables will usually be far greater than can be controlled for, so that experimental approaches, which seek to isolate the effects of different variables, are not appropriate. This is particularly pertinent for development and policy work, given the many parallel and interrelated factors which affect people's lives.

Case studies are also relevant where individual, rather than standardised, outcomes are sought. They are increasingly used in assessing health and education services and policies, and development agencies such as the World Bank and USAID are carrying out more case studies and synthesis (cross-case) studies. The following factors may make a case study approach more or less appropriate: if there is a high number of interrelated variables within an unpredictable and uncertain environment, and more individual outcomes, then case studies will be more appropriate. If there are few variables which are clearly related, a predictable and stable environment, and if you expect more standardised outcomes, then standard quantitative measures are probably more appropriate.

Case studies therefore, are generally used when one needs to understand a specific group of people, a particular problem, or a unique situation in great depth; where one can identify cases rich in information — rich in the sense that a good deal can be learned from a few examples — and where the complexity of the issues makes standardised approaches less suitable. Like many other approaches, case studies usually involve exploring several different, and competing, interest groups. As they may have different interpretations of events this often again requires cross-checking of findings. In many situations it will be necessary to combine case-study work with broader, less in-depth, surveys. This was the general approach adopted in the impact studies reviewed in this book.

In the BRAC study, for example, case studies were used to explore the qualitative changes associated with women's empowerment, and to find out more about households which had dropped out of the programme as well as those which had demonstrated exceptional success. For assessing women's empowerment, a random sample of 25 village organisations was selected for case-study work, while the dropouts and successes were chosen from relevant data collected through the household survey. This use of case studies allowed BRAC to understand the factors which influenced the different levels of households' success, such as close kinship ties and involvement in the management of the village organisation; higher levels of training; and enjoyment of special loan privileges. In collecting women's views of impact, qualitative case-study methods provided more in-depth information than the quantitative survey. This was particularly the case in deepening the understanding of how the programme had contributed to women's positive self-perceptions, increased their self-confidence as their dependency on male household members decreased, changed decision making at the family level, and increased women's mobility which enabled them to communicate better with the outside world.

CYSD also used a case-study approach to explore critical incidents which had influenced the impact of their programme. These included conflict

between groups in a particular village, and two projects which had failed. In addition, they found that past case studies of particular individuals, villages, groups, or organisations proved to be among the most valuable documentation for reconstructing baseline data, because they provided detailed information, about the position of females, as well as about health practices and coping strategies before the project intervention. The study of the two project failures – which tried to introduce new, inappropriate, technology – revealed a high speed of project identification and implementation (due in part to donors' pressure), as well as a poor understanding of the technical and labour demands of the new technology.

Project staff attributed some of the causes of this problem to the fact that fulfilling their output target had become their primary concern, but also to poor advice from external consultants. These examples demonstrate how an in-depth case study can bring up facts relevant for the workings of an entire organisation, in this case suggesting the need to clarify relations with donors and consultants; to remedy weaknesses in project identification and appraisal; and to adopt and adapt targets and indicators in ways that discourage the uncritical fulfilment of inappropriate targets.

The Ghana study used a 'layered' approach to case studies. At the first level we have case studies of three local organisations. At the second level, case studies were made up of three villages where these organisations worked. Third, ten individuals were selected in each of the three villages. This approach tried to ensure that case studies reflected the wide variety of given groups, so that any similarities that emerged were likely to reflect core impacts. In this case this meant capturing the varied nature of the members of the Northern Ghana Development Network (NGDN); the agricultural and climatic diversity of the areas in which the Network operates; and the varied levels of well-being in the villages. Thus at each level (network, organisation, and village) information was collected in order to select case studies. This 'snowballing' technique recognises that the information needed to choose case studies may not be available at the outset of the research, but may need to be collected during the course of the assessment.

Table 4.19 builds on these examples from the case studies to illustrate the usefulness of different types of case study.

One should not forget how useful case studies can be in demonstrating and communicating impact, as well as for training processes related to impact assessment. Stories and examples usually make up an important part of an organisation's formal and informal communication as well as its decision-making processes. Case studies are a great device for explaining and illustrating complex and sometimes dull programmes. They can act as a

Table 4.19: Types of case study and their usefulness

Types of case	Usefulness
Unusual, extreme, or deviant cases (programme dropouts, failures, or successes)	Useful in understanding puzzling cases which seem to break the rules, and why certain people or organisations seem to achieve particularly good or bad results. Useful in understanding the reasons for exceptionally good or bad performance.
Typical or average cases	Useful in understanding the situation of most people, communities, and organisations. Findings may be replicable in other 'normal' situations.
Homogenous or similar cases (for example, looking at impact on a group of women of the same age, or looking at a number of credit projects)	Useful in looking at particular sub-groups in depth, which may be important when many different types of people or activities are involved.
Varied or heterogeneous cases (deliberately seeking out different groups of people, organisations, or types of programmes)	Useful in exploring common or distinct patterns across great variance. Common patterns in such cases are likely to indicate core and central impacts of wider relevance, precisely because they occur across diverse groups.
Critical cases (may have wider relevance; can be used for broader purposes, such as innovative work or work with new groups; or may produce results which have high political impact)	Useful when a single case study can dramatically make a point; statements such as 'if it happens here it can happen anywhere' or 'if it doesn't work here it won't work anywhere' indicate that a case is critical.
Snowballing cases (one starts with a few cases and then selects others on the basis of the findings)	Useful when the information to select all case studies is not available or are dependent on a greater understanding of the situation.
Convenience cases (possibly the most common type where case studies are chosen solely because it is easy - the information already exists, the site is very close, and so on)	Generally a bad idea if these are the only or most important reasons for choosing case studies.

Adapted from Patton (1990)

shared starting point from which managers, trainees, and others can debate wider issues. This is vital if staff do not share common experiences or if there is a tendency for debates to become very abstract and theoretical. Good case studies can ground debates in reality, as well as providing empirical material to determine wider policy questions. Summary 4.6 makes some key points about the use of case studies for impact assessment purposes.

Summary 4.6: Lessons learned about case studies

- Case studies proved particularly useful to explore qualitative impact such as women's empowerment and self-confidence, and relationships within the household.
 - Case studies were also useful in following up some of the results generated by other methods, for example in order to analyse particularly successful or unsuccessful projects or groups, and exploring why they had succeeded or failed.
 - Case-study material, because it tends to cover single examples in depth, is particularly useful for illustrating and communicating impact to others; if chosen carefully, the study can have much wider relevance.
-

Costs, benefits, and impact

In recent years NGOs have been criticised for not doing enough to explore the relationship between the costs involved in what they do, and the benefits or impact that is achieved (Riddell 1997). The reasons normally given for this are that NGOs are resistant to this sort of quantification of complex processes, or that this type of analysis is too complex and difficult to undertake.

Two of the case studies (Pakistan and the Wajir programme in Kenya) attempted to explore the relationship between costs and benefits in some depth. Both divided project benefits into those that yielded clearly measurable economic gains, and those that were less easy to quantify but nonetheless made significant contributions to a secure livelihood and human development. In Wajir, quantifiable economic benefits included changes in livestock mortality; benefits brought about by restocking destitute families with animals; women's credit; savings brought about by the provision of animal and human drugs; and savings made through greater security in the area. In the case of the Pakistan study, three micro-projects were studied and compared: a goat loan project, a literacy project, and a rose cultivation project.

The Wajir study

In the Wajir study, most of the information for assessing economic benefits was collected by surveys of representative samples of project participants and non-project respondents, as described at the beginning of this chapter. This data was then used to compare families living in project and non-project areas and to calculate average savings per household in relation to differences in livestock mortality and expenditure on drugs. These average figures were then extrapolated to the total number of targeted families in the project areas, which permitted an overall calculation of project benefits.

For more targeted activities, such as women's credit and animal restocking, more specific surveys and focus-group discussions were undertaken with a sample of beneficiaries to determine economic costs and benefits. Average benefits of the credit scheme were calculated in terms of the return on loans disbursed and of savings accumulated; the economic benefits of restocking were calculated in terms of increase in the value of the stock of animals and of the additional income achieved through milk sales. A financial value was also calculated for livestock stolen in the area which the Wajir Peace and Development Committee had recovered. Table 4.20 describes the steps undertaken for each project component in order to arrive at the overall cost-benefit figures.

Table 4.20: Steps undertaken in calculating economic benefits of different components of the Wajir programme

Livestock mortality	1	Calculated average difference in mortality for 200 households in project and non-project areas.
	2	Calculated average monetary saving per surviving animal and thus per household per year.
	3	Extrapolated findings to all the 1,788 families targeted by the project.
	4	Calculated overall reduction in livestock mortality and the estimated annual saving to pastoral communities over the lifetime of the project (£473,582).
Restocking	1	Found out number and value of animals distributed per family.
	2	Calculated average number of animals per family at the end of years one and two of the project, as well as the number and value of animals at the time of the assessment.

Table 4.20: Steps undertaken in calculating economic benefits of different components of the Wajir programme (continued)

Restocking	3	Calculated average monetary benefit per family over the first two years of the project.
	4	Calculated the value of average additional milk sales from a family's flock.
	5	Calculated total benefits per family.
	6	Compared total benefits with initial input costs (income over and above the initial value of the animals purchased by the project).
	7	Calculated the total benefit of the project's restocking component based on the 500 families which have been restocked (an increase in livestock capital, over and above the original value of the donated animals, and milk production valued at £101,000).

Women's credit	1	Calculated average return on loans made to a sample of women
	2	Calculated average rate of return on the loans disbursed (102%).
	3	Compared this to initial capital disbursed (KES 4,662,000 disbursed to date generated KES 4,755,240 in one year).
	4	Calculated overall return (£44,000).

Drug savings	1	Discovered average expenditure on drugs per person per household in project and non-project areas through questionnaire survey.
	2	Calculated difference between the two groups (on average, households in project sites spend 42% less on drugs per person compared to households in non-project sites).
	3	Calculated average annual savings per person.
	4	Calculated overall reduction in drug expenditure in the project area (£9,726).

Wajir peace	1	Discovered amounts of stolen goods recovered by the Wajir Peace and Development Committee.
	2	Calculated value of these goods: records show that two land cruisers and one trailer, livestock and 2,000 guns have been recovered through the initiative.
	3	Calculated overall reduction in theft losses (£30,000).

Table 4.20: Steps undertaken in calculating economic benefits of different components of the Wajir programme (continued)

Overall analysis	<ol style="list-style-type: none"> 1 Made assumptions about the likely future costs and benefits of the programme as a whole — based on current performance for years one and two, the programme plans, and likely inflation rates. 2 Calculated future annual benefit of each project component, and the annual overall costs of the programme, for years three to ten of the programme. 3 Predicted and added up the total present value of the economic benefits of each of the components, taking into account the effect of inflation for future benefits. 4 Predicted the total present value for the economic costs of each of the components, taking into account the effect of inflation on future costs. 5 Calculated a number of ratios of present cost-to-benefit value (total or net present value — total present value of benefits minus total present value of costs — was £1,305,663), a simple cost-benefit ratio (1:2.2), and the internal rate of return (the rate of inflation that would be required to reduce overall benefits to the same level as costs — in this case, the IRR was 55%).
-------------------------	---

This process of calculating overall figures for those parts of the programme that lent themselves to this sort of analysis permitted a comparison of the key economic costs and benefits of the entire programme. The project team could then derive some commonly used ratios such as overall net present value, cost-benefit ratios and internal rates of return¹⁰. As these ratios were calculated for the lifetime of the project (ten years) and the assessment was undertaken in year three of the project, this demanded, like most assessments of this type, some rather 'heroic' assumptions be made:

- that the benefits measured in year two would be maintained at the same level over the next seven years;
- that as a result, there would be no absolute increase in benefits over the project's lifetime;
- that the capital available for credit would remain available at the same level;
- that the operational costs would fall in line with the programme plans.

In addition these sorts of calculations require predictions, among other things, about future inflation and how this will affect costs and benefits.

Despite these uncertainties the cost-benefit analysis helped to determine the relative importance of elements of the programme in terms of economic benefits (see Table 4.21). Combined with findings from the qualitative assessment undertaken, this can provide useful material for future planning.

Table 4.21: Proportion of economic benefits derived from different parts of the Wajir programme

Elements of the Wajir programme	Percentage of total economic benefits
Animal health	74.6%
Restocking	12.0%
Credit	8.2%
Drug savings	1.5%
Wajir Peace Committee	3.7%

Bringing together economic impact analysis, cost information, and qualitative information, also allowed the Wajir team to draw meaningful conclusions about those elements of the programme which were extremely difficult to measure, particularly the institutional development efforts.

The relatively high benefit:cost ratio, even under the some of the most conservative assumptions, point to an overall success of the project. Institutional capacity development, especially with respect to delivery of animal health services to pastoralists, makes the greatest contribution to the package of economic benefits that accrue to the pastoralists.

It is important, however, to note that the pastoralists only started to realise economic benefits during the second year of the project and that the greatest benefit will only be realised from the sixth year of the project. This brings to the fore a very important consideration in development programmes in pastoral areas. It is often necessary to consider the amount of capital investment needed to establish and/ or strengthen local institutions before they can be able to deliver the necessary services. It takes a lot of financial resources to remunerate personnel involved in capacity development, gain the community's confidence by responding promptly to their most immediate needs, travel to the project sites

in order to get the institutions functioning, train the members of these institutions, invest in capital and monitor and evaluate all on-going activities. This initial investment in capacity development, though costly and less efficient, is often necessary if one considers future stream of benefits from the project even in the absence of the donors. (Wajir case study)

The Pakistan study¹¹

The assessment team in Pakistan undertook similar analyses of three micro-projects. These included a goat loan project and a rose cultivation project, as well as a literacy project, where the economic returns were less obvious. The information collected on economic impact was recalculated according to different assumptions — about how women’s work is valued and whether Oxfam’s support costs are included in the calculation or not — and compared to participants’ own ranking of benefits, to explore how they compared economic benefits to others (see the section on ranking above for more on the method used in this study). Some of the key economic results are found in Table 4.22.

Table 4.22: Comparative results of cost-benefit analysis in three projects in Pakistan

	Goat-rearing	Khoj literacy	Rose cultivation
Total Oxfam grant	63,400R	506,329R	16,764R
Participants’ socio-economic status	poor women	poor women	poor women and men
Number of direct beneficiaries	140	66	10
Number of years of operation	5	3	2
Grant total per participant	453R	7,671R	1,676R
Annual income per participant	1,102R	190R	2,286R
Market viability?	Yes	No	Possibly
Plausible internal rate of return¹² including Oxfam costs	20%	-6%	-52%
Plausible internal rate of return excluding Oxfam costs	20%	-6%	36%

All costs in constant 1992 Rupees

Adapted from Alkire 1999

The conclusions for the goat-rearing project are relatively clear-cut: it is economically viable and seems a worthwhile investment. Interestingly, the women involved ranked non-monetary benefits higher than the impact on their income. They particularly ranked their self-confidence, the solidarity that came from working with other women on the project, and the knowledge they had gained in terms of goat breeding and marketing, higher than the economic benefits. Clearly there are linkages between these different benefits and untangling them is difficult — women's increased knowledge about goats may be reflected in the sale prices they receive. In this case at least some of the women's increased knowledge can, in theory, be given an economic value. However, the women themselves feel that there are other benefits which cannot be 'priced' in this way.

In the case of the Khoj literacy project, the economic and non-economic impacts go in opposite directions. From an economic point of view the project is not sustainable without continued support, but the non-economic assessment makes clear that the project has had a 'fundamental and transformative impact' on the women students. This was particularly the case in terms of their ability to speak out and to deal with physical abuse — the following quote from a graduate student is a typical example.

My mother and father did not allow girls to speak; they would beat us. Now I have learned to trust my own talk, and ability to judge that this is good and this is bad ... My heart has become strong. I can speak about my rights, can even slowly, politely tell my parents that they have done something wrong. We have begun now to talk with them about the difference between sons and daughters, that we are equal. Literate people can solve their own problems. (quote from a literacy student in Alkire 1999)

One of the main reasons why the economic returns to the project are low is that even literate women find it hard to get jobs, because overall female participation in the labour market is so low (7.15 per cent in Pakistan's urban areas in 1993/4). The primary barrier to their participation is not so much low productivity, or lack of skills and entrepreneurship, which literacy might address, as their inability to work outside the home or to undertake lucrative employment in the home. The barriers to female employment are therefore largely due to social constraints and conventions — which participants are starting to question as a result of the method of literacy used in the project.

With this understanding one can make more sense of the fact that Oxfam's grant to this project was eight times larger than the grant to the goat-rearing project, and reached less than half the number of women. Indeed, it may be a much more strategic investment because it, and projects like it, can arguably contribute to changing social constraints which will allow many more

women, including those who are already literate, to realise their potential in the labour market.

In the case of the rose cultivation project, the main issue in determining the project's economic viability is the cost of Oxfam's support. As we see from Table 4.22 the inclusion or exclusion of Oxfam's support in the calculations makes a critical difference. The conclusion of the cost-benefit analysis alone would be that

while the income generation activity for roses would be viable and indeed lucrative as an economic investment, and while this income would have significant social premium because it would have accrued to the poorest households in the village, the high cost of Oxfam support turned what would have been a highly positive social cost-benefit into a strongly negative one. (Alkire 1999)

However, while this analysis makes it clear that Oxfam should be increasingly devolving responsibility for the project to the women's group it does not capture why these costs were perhaps necessary in the first instance. The project history and qualitative assessments make this much clearer as they reveal the catalytic role of the female Sindhi-speaking Oxfam project officer (for example, in ensuring adequate technical feasibility) and her efforts to ensure that the group's organisational development was taken into account. The fruits of this labour are evident in the participants' perceptions who once again rank non-monetary aspects of the project extremely highly.

Conclusion on costs, benefits, and impact

As a result of these sorts of analysis NGOs may often be faced with a choice between efficiency and empowerment. If economic impacts could be improved without compromising the other impacts, as in the Khoj project, this would be the preferred path for the future. However, if the option was to support a successful income-generation project that did not have high impacts on knowledge, relationships, or empowerment, the choice would become starker and more difficult. The choice can be made arbitrarily or on the basis of 'gut-feeling'; an agency can decide to prefer one objective (income) over another (empowerment) as a matter of policy, and develop appropriate expertise in this area (this may in part also be determined by what other actors are doing). Lastly, a judgement or policy decision can be made to work with and support particular groups of people, for example, the poorest, and to identify with these groups the activities and impacts they wish to achieve, and develop a broad expertise (and range of contacts with other specialist groups) that would allow a flexible response.

Clearly while economic analyses of costs and benefits can be useful in making these sorts of decisions, there is nothing in economics which relieves

us of the obligation to choose' (Robbins 1932, pp. 135-136, quoted in Alkire 1999). And indeed, as the Pakistan and Wajir case studies make plain, it is the *combination* of the economic analysis with a qualitative assessment of dimensions of change, which is likely to make those choices at least better informed, if not easier. Summary 4.7 outlines some of the key lessons learned in the case studies about trying to undertake a more systematic analysis of the relationship between costs, benefits, and impact.

Summary 4.7: Key points about economic impact assessment

- There is a need to distinguish between those elements of a project or programme for which an economic analysis makes sense and those for which it does not.
- Collecting data for project participants and non-project participants with similar characteristics can allow differences between the two to be calculated; care must be taken to disaggregate this data, especially with regards to gender. Remember to 'price' people's labour and time.
- When assessing economic benefits, look not only at income gained but also at savings in expenditure.
- Where it is not possible to survey or talk to all participants, talk to a sample to arrive at an average value and then extrapolate this to all participants. The statistical validity of this will depend on your sampling procedure (see Chapter 3) on the variance between the values discovered among the sample. Care must be taken to ensure that gender differences are not subsumed in averages.
- Compare the costs and benefits for different components of a project or programme if possible, as well as looking at the project or programme as a whole. Remember that the value of money changes over time; always use constant prices. If constant prices are not used, assessments of a project's future must take into account the effects of inflation on future costs and benefits (constant price figures, and inflation and discount indices are usually available from multilateral agencies or government statistical departments or finance ministries).
- When making assumptions about the future (or for example, about the opportunity cost of someone's time) it is often better to include a range of values or scenarios, for comparison. Don't forget to spell out that these are assumptions.

- Always remember that economic analysis on its own is incomplete and can be misleading. It needs to be combined with qualitative and non-economic assessments. Particular care is needed in comparing projects with very different economic and social elements.
-

Both economic and non-economic assessments are necessary for taking informed decisions. They do not, however, remove the need to make judgements and to develop policies which guide the choices that are made between rival objectives.

The previous chapter explored issues related to the design and preparation of impact assessment processes. This chapter has looked in some depth at the different tools and methods used to undertake impact assessment and drawn a number of lessons from the case studies. In the following chapters, I shall discuss how these lessons and findings apply to specific aspects of NGO work: emergencies, advocacy, and organisational self-assessment.

5

Impact assessment and emergencies

This chapter describes some of the specific characteristics of emergency situations and their implications for impact assessment. In particular, I explore the similarities and differences with the lessons learned in Chapters 3 and 4. The first section analyses some of the specific characteristics of emergency situations and the implications for impact assessment work. The second part looks more specifically at the design of the impact assessment process in emergencies, and the choice of tools and methods. Although none of the case-study projects were involved in large-scale acute, or first-phase, emergency work, some — notably Wajir, Ikafe, much of Cordes' work, and one village in the Ghana study — either developed out of an emergency situation and were medium-term responses to it or, in the case of Ghana, involved assessing impact in a community ravaged by a local conflict. The chapter therefore refers to relevant parts of these studies as well as other material on evaluation and impact assessment in emergency contexts. In particular, it draws upon an evaluation of Oxfam GB's work in Africa's Great Lakes region (Collins et al. 1998), other work undertaken on the subject by Oxfam (Dawson 1998), recent work attempting to draw together lessons about evaluating humanitarian assistance from across the sector,¹ and the work of the Sphere project.² The chapter concludes by drawing together the lessons on key indicators to be examined in emergency situations, on tools and methods, and some of the questions regarding the ethics and practices of humanitarian organisations which can affect their impact.

Recent years have seen greater scrutiny and a growing critique of humanitarian operations, following events in Somalia, Rwanda, and Bosnia. Rightly or wrongly, Northern NGOs have borne the brunt of much of this: they stand accused of being ineffective and dishonest about their achievements; of lacking accountability and the ability to learn; of acting as sub-contractors for major donors in a 'humanitarian industry'; of being a fig leaf for government inaction, or worse the vanguards for the creation of western-style civil society and liberal democracy; of having a poor understanding of the political economy in which they are engaged; and of having a vested interest in their own survival (African Rights 1994, Duffield 1996, 1997). This is compounded

by a growing 'contract culture' where a focus on 'on the fulfilment of contracted inputs and outputs rather than on actual humanitarian outcomes' allows 'the industry to demonstrate contractual success even within spectacularly unfulfilled mandates' (Stockton 1995).

This barrage of criticism has largely focused on the 'dark side' of humanitarianism, as Hugo Slim points out.

[M]ore academic attention is being paid to the negative effects at present and it would seem somewhat urgent for NGOs to be able to assess and advocate the positive effects more clearly in the years ahead. Such assessment of the 'bright side' of humanitarianism must move beyond traditional quantitative indications of output (food and blankets given etc.) to a more subtle analysis of impact and outcome which can stand up to the increasingly rigorous analysis of the dark side of relief in war. (Slim 1998)

Characteristics of emergency situations

Assessing the impact of interventions in emergency situations demands addressing many of the points that have been made in previous chapters, such as being clear about the model of change being pursued, setting clear objectives and indicators, and collecting adequate baseline data. Many of the challenges described, relating for example to attribution and aggregation, are equally present in emergency situations.

However, emergency situations also have very distinct characteristics. If a humanitarian emergency is defined as any situation in which there is an exceptional and widespread threat to life, health, or basic subsistence beyond the individual's or the community's capacity to cope, the following characteristics usually apply:

- a high level of immediate needs in terms of security, health-care, or subsistence;
- increased vulnerability of some groups who are less able to express their specific needs (particularly women, children, refugees, and minority groups);
- weak local capacity to respond;
- a volatile, turbulent environment which may also be insecure;
- possibly the presence of multiple actors/ agencies, often deliberately or inadvertently pursuing different goals;
- a poor level of information on precisely what is going on and why.

Hallam and Borton³, in their review of 'Good Practice in Evaluating Humanitarian Aid', also suggest that complex emergencies or conflict situations have an additional number of characteristics. In many civil wars — which make up the majority of conflict situations — the objective of opposition or rebel groups is to overthrow the government or to secede. The degree of recognition and status accorded to opposition groups is therefore a highly sensitive issue with implications for the role of international agencies, the legal basis for their operations, and the legal rights of affected populations. Civilians may be deliberately attacked and their way of life and means of livelihood undermined; rape and sexual violence are often common. Such attacks on communities often have the objective of instilling such fear in groups of similar background that they will seek refuge elsewhere, releasing territory for use by the faction responsible for the atrocities. Attacks may also be carried out on targets which play a special role in the cultural identity of particular groups, such as places of religious worship. Women are often regarded as bearers of a group's cultural identity, which affects how they are treated during and after war, and how they view themselves. They can become targets for abuse as a way of degrading a whole society or ethnic group.

Complex emergencies frequently last for several years, but within this the situation can be highly fluid. Fighting is increasingly causing civilian as well as military casualties, and threatened populations flee, creating more or less localised displacement crises. Chronic problems are therefore frequently intensified by acute situations requiring urgent responses. In both cases the degree of freedom enjoyed by individuals, and institutions such as the press and the judiciary is either severely constrained or eliminated. Actors involved in the conflict and those involved in trying to provide humanitarian assistance therefore operate in a context of severely weakened national accountability mechanisms.

This in turn means that conflict often results in the development of economic activities which would normally be classified as illegal or semi-legal, such as the exploitation of natural resources, drug-running, money-laundering, and trading in arms. Often these activities are controlled by, or taxed by, leaders of the warring factions who use them for personal gain or to continue the conflict. Commercial organisations based outside the affected countries may be involved. Such 'war economies' have enabled conflicts to be prolonged.

Humanitarian assistance can be diverted from the intended beneficiaries and controlled and taxed in the same way as 'war economy' activities. In many recent conflicts it has been difficult to distinguish between civilians and combatants. In many cases, the intermingling of combatants and civilians is a deliberate strategy of the warring parties. In this situation, humanitarian agencies are often unable to prevent assistance distributed for use by civilians and vulnerable populations being used by combatants and warring factions. We lack reliable evidence of

what proportion of humanitarian assistance is diverted, but questions arise about the degree to which external assistance can fuel and prolong conflicts.

A large number of actors can be involved in complex emergencies, ranging from national relief structures, local NGOs, international NGOs, UN agencies, the ICRC, the IFRC, and the national Red Cross and Red Crescent societies. Increasingly human-rights agencies and monitors are also deployed in ongoing conflicts, as are organisations seeking to resolve the conflict. Humanitarian activities may run parallel to, or in concert with, peacekeeping operations, which may involve troop contingents from a variety of countries. The conflict may involve neighbouring states. The coordination of such a multiplicity of actors is a major challenge both to local authorities and external agencies. With such a wide range of actors, it is almost guaranteed that they will not share the same goals or even be at odds with each other.

Much of the recent literature on evaluation in emergency situations, because of events in Bosnia, Kosovo, Somalia, and Rwanda, has tended to focus on complex emergencies and conflict situations. However, it needs to be recognised that there are also important distinctions between various types of emergency. The distinction between 'man-made' and 'natural' disasters is often challenged, but it is clear that dealing with earthquakes, floods, hurricanes, or drought in areas free of conflict presents specific challenges. It is different from humanitarian work in conflict zones, and therefore holds different challenges for impact assessment. In addition one can distinguish between varying capacities within the world's regions to deal with disaster. Much of the literature on impact assessment and evaluation presumes an important, if not predetermined, role for international agencies and actors in delivering humanitarian relief and in developing policy. There seems to be much less literature and evaluation experience about assessing the impact of local response, local preparedness, and local capacity-building.

One can further distinguish between those emergencies that are periodic in nature and to that extent predictable, and those which are not. Floods in Bangladesh are a regular event and it is their severity, rather than the event itself, which is less predictable. In such circumstances, a risk analysis identifies those likely to be most vulnerable in advance of the emergency, and allows preparedness measures and plans for response to be put in place. In theory, the experience of previous disasters can be disseminated after the event and fed into the next response. In Bangladesh it is reported⁴ that as a result of lessons learned, some agencies have recognised that in order to assist women and children, women staff must be included in assessment teams and provision for women made in places of safety. It is believed that this has reduced the number of women who have died, or suffered serious losses, in recent floods in Bangladesh.

Implications for impact assessment

The characteristics outlined above have several implications for impact assessment. The scale of need and the weakness of local capacity may mean, at least in the short term, that direct intervention 'high up the impact chain' is necessary, for example by providing food or water in order to prevent immediate loss of life. In the short term, building the capacity of a third party to provide food and water (in other words, providing support lower down the impact chain) may not be an option. Generally, the fewer the links in the chain the easier it is to assess whether a given input achieves an impact.

Discontinuous and rapid change makes planning and setting objectives more difficult. The need for a timely response can also limit the possibilities for adequate collection of baseline data and for participation, although both can be added to an intervention over time.

Different groups of people have different capacities to cope with crises as well as different vulnerabilities, particularly depending on their gender and age. It is essential to understand these when establishing impact indicators and assessing results. Recognising these differences and collecting baseline data for future assessments might necessitate diverting resources from ongoing operations, which might be considered an unacceptable luxury when there are lives to be saved.

The complexity of the emergency and the interconnectedness of the actions of multiple agencies may make attribution of impact particularly difficult. In this case, an impact assessment of a single agency's work will be less useful than an analysis of the combined actions of several agencies. On the other hand, the short impact chain in an emergency may, in some cases, make the attribution of impact easier than in long-term activities. For example, an agency's provision of clean water in a refugee camp can have an obvious, immediate effect on people's health and life chances, whereas it will be much more difficult to assess whether training a local team to provide counselling for refugees actually leads to reduced stress levels.

The insecurity, fear, and trauma of people affected by emergency situations may make 'telling the truth' not only dangerous but deeply disturbing. Especially in conflict situations, various groups will also have an interest in ensuring that their version of the 'truth' predominates. This has implications for participatory processes.

Local responses, preparedness, and capacity are often undervalued and sometimes undermined by external agencies. A comprehensive approach to impact assessment in emergency situations has to include the relative merits and impact of local actors vis-à-vis external actors, their impact on each other and the implications of this for a longer-term response, as well as the degree of predictability of the emergency in question.

These points have consequences for the design, choice of methods, and implementation of impact assessment in emergency situations, which I discuss in the following sections.

Design

Preparation

As stated in Chapter 3, being clear about the purpose of an impact assessment is the first step in designing an effective process. Being clear about who the assessment is for, what its boundaries and its focus are, what financial and human resources are available, and what timescale is envisaged are obvious, but vital questions. However, experience also suggests that the very nature of emergency situations makes answering these questions very difficult. Some therefore, have suggested that a preliminary scoping study or design phase may be appropriate, 'which allows the evaluator to identify potentially significant issues and prepare a more accurate schedule and budget for the remainder of the evaluation' (Hallam and Borton 1998). Again, as previously discussed, preliminary workshops involving a range of stakeholders may be an effective means not only of determining the purpose of a given study but also of designing one that is realistic.

Clarifying assumptions

All interventions make certain assumptions about how change happens, why it has occurred, and how it might occur in the future. Making these assumptions explicit and being clear about the 'logic' behind them is usually the first step in designing an assessment. This permits the kind of institutional and policy analysis which is increasingly recognised as an important part of impact assessment (see particularly Chapters 6 and 7).

What types of impact and indicators are to be assessed?

Impact assessment studies, particularly in complex emergencies, must recognise that objectives and strategies may well change in response to changing circumstances, that some objectives are often not stated or agreed by all parties in advance, and that unanticipated effects are usual. As the Wajir study concluded,

the impact on project beneficiaries' livelihoods and welfare often calls for the continuous integration of (new) impact indicators during the project implementation period. Only in this way can we know whether the project is having any impact on the beneficiaries. The constant integration of impact indicators ... is also necessary because of the uncertain and unpredictable nature of the project environment. (Wajir Economic Impact Assessment)

What this suggests, and as indicated in Chapter 3, the process of assessment and indicator development must:

- emphasise monitoring and 'impact tracking' in addition to adequate planning and evaluation;
- recognise the potential for differential impact on specific groups, particularly relating to gender and age;
- update and reformulate existing indicators, as well as introducing new indicators and dropping others, in the light of changing circumstances;
- explore significant changes which occurred as a result of the projects but which lay outside the initial indicators, and use this information to develop indicators for the future, as well as deliberately setting out to capture negative change and to seek out those who might provide this information.

Choice of Tools and Methods

Surveys

In Chapter 4 we saw how the use of large-scale household surveys can provide useful impact data if executed properly and if used in conjunction with other methods of data collection. The Wajir programme in north-eastern Kenya illustrates how surveys of this type can be used in what is essentially a long-term rehabilitation programme. However in large-scale emergency operations with a high degree of population movements, such as the Great Lakes crisis, it is unlikely that this sort of approach will be feasible after the operation, given that making contact with refugees and displaced groups and returnees will prove almost impossible.

However, this does not mean that impact assessment surveys cannot play a role during the crisis as a form of impact monitoring, or in post-emergency situations which do not involve large population movements. In Goma in the Democratic Republic of Congo, this role was performed at least partially by a number of epidemiologists who gathered data from a broad perspective, assessing potential risks and lapses in coverage and identifying emerging problems. This was identified as a major factor in improving humanitarian response in Goma compared to other sites where this role was largely absent (Borton et al. 1995, Collins et al. 1998).

In this case, impact data tended to focus on mortality and morbidity rates while individual agencies examined the quality and quantity of their outputs. In Oxfam GB's case, because of its role in water provision, the focus was on the quantity, quality, and accessibility of water. In some cases, where more

qualitative feedback from refugees was gained, this proved important in helping the programme to evolve in different directions. For example, although there was a 95 per cent coverage of public latrines in Chabilisa camp in Karagwe, the population felt that public latrines were very dirty and unattractive to users. This was particularly acute because the population were unwilling to take responsibility to clean latrines that so many people used. As a result, teams of cleaners were employed, and the introduction of family latrines in some camps proved very popular and encouraged families to build and maintain them (Collins et al. 1998).

Clearly many of the lessons which relate to the management and implementation of surveys for impact assessment of long-term development work which are highlighted in Chapter 4 are equally relevant in emergency situations:

- focusing on a small number of critical issues;
- getting the balance right between exploring impact and verifying the degree to which project outcomes might have led to it;
- collecting data which allows comparison of issues over time;
- checking how representative the sample chosen is, and being aware of the need for disaggregated data;
- pre-testing and refining questionnaires before the survey is implemented;
- taking care in the selection of enumerators and ensuring that they are properly trained and briefed.

Interviews, workshops, and discussions

Since documented information is likely to be unavailable, interviews, workshops, and discussions are an important part of assessing impact in emergency situations. These will probably need to involve a wide variety of stakeholders in order to reconstruct the context and project history, and to cross-check findings from a number of perspectives. Many of the advantages and disadvantages of group discussions or interviews compared with individual interviews highlighted in Chapter 4 are also relevant to emergency situations. For example, while bringing together people with different experiences and perspectives can help paint a picture of impact, in a conflict situation the ability of less powerful individuals or groups to express their views openly in the presence of others may be extremely limited.

On the other hand, the experience from the Ikafe case study in northern Uganda suggests that it is possible to strengthen the voices of some groups (in this case the host population and refugees) vis-à-vis others (aid agencies and

local government authorities) through adequate preparation and a carefully designed and facilitated process. The exchanges in Ikafe provided rich food for thought for the agencies involved as their own perception of performance was contrasted with that of the refugees and the host population. These different perceptions are summarised in Table 5.1 which outlines the positive and negative impacts as described by Sudanese refugees, the Ugandan host population, and the aid agencies.

As the Ikafe experience demonstrates, different stakeholders may not share common objectives and may in fact seek contradictory impacts. They may use the political opportunity that an evaluation offers to advance their views and opinions, and to downplay those of other groups, leading to exaggerated responses and half-truths pre-dominating. An impact assessment team must be aware of that bias and ask who might be exaggerating and whose views are not represented and why, in order to understand the interests of different groups and their relations with others.

The Ikafe case study also brought to the fore the high levels of tension, trauma and stress that are present within stakeholder groups and between them. Embarking on a process which asks people to relive painful events and involves them in discussions with others who may have a different perspective requires extreme care and thoughtful preparation.

If impact assessments need to reconstruct history, often through interviews, there is a growing consensus that this should take place relatively soon after the event. However, it may also be necessary to make provision for the assessment of longer-term impact, for example in relation to psycho-social trauma, which may have long-term effects that take time to address.

Direct observation

Chapter 4 elucidated the role of direct and participant observation in building trust and rapport, cross-checking information, exploring relationships, discovering new insights, and in enabling a better understanding of an intervention's context. However, we also saw that the success of the method is dependent on existing levels of trust between observer and observed; that observers can affect the results they are observing; that it requires not only good observation skills but also interpersonal and communication skills; that it can be open to bias and selective perception on the part of the observer; and that it can be limited to external behaviours if the onlooker approach is adopted and it is not used exclusively. It is also a resource-intensive method.

In an emergency situation, it is clearly impossible to be a participant observer in the true sense, without being a victim of that disaster. Yet, as the evaluators of the Oxfam programme in the Great Lakes pointed out, boundaries between victims and onlookers can become blurred in complex situations.

Table 5.1: Different perceptions of project performance and impact in Ikafe

Group Perception	Sudanese refugee	Ugandan host population	Agencies including Oxfam
Positive	Situation improved since moving from Koboko; infant mortality decreased.	Local economy showed some improvement. Infrastructural improvements benefited some locals.	Infrastructure developed; food and water delivery, health-care and educational services set up which benefit all refugees and many locals, leading to improved nutritional status of refugees and increased food security. Seeds and tools delivered to most refugees who have been assisted with cultivation. Land allocation in process: strong boost for local economy.
Negative	Water and food provision irregular. Land often of poor quality. Seeds and tools delivered late. Alternative livelihood opportunities limited, and markets far away. Young not attending school. Lack of involvement of refugee council.	Not enough local people employed or contracts given to local contractors. More infrastructural development should have been located in the main local town instead of in the centre of the refugee settlement. Lack of consultation on land allocation to refugees, and violations of some sacred sites.	Lack of land and security problems hinder development of refugees' self-reliance.

During many of the Great Lakes emergency operations circumstances were often extremely harrowing and threatening. High levels of stress were, as a consequence, often unavoidable. The extent of psychological trauma suffered by many of the local staff in the Kigali office was so extreme as to be beyond the comprehension of the evaluation team ... One senior office manager lost 40 members of her family during the genocide. (Collins et al. 1998)

Agency staff, particularly local staff, do share, at least partially, the horror and the risks inherent in an emergency situation (whereas external assessors usually come in after the height of a crisis). Thus staff are observers whose views and experiences may be very pertinent in understanding impact and the process by which it occurred. By the same token, however, their very close involvement in possibly traumatic circumstances also means that they may find it difficult to distance themselves from the events. As noted in Chapter 4, observation is a skill that needs to be developed and rewarded, a theme I shall return to in Chapter 7.

RRA and PRA/ PLA tools

Many of the participatory research methods covered in Chapter 4 can be useful for impact assessment work in emergency situations. Some examples are given below of tools which would seem to be of particular relevance, given the characteristics of emergency situations.

Time lines and historical profiles

Time lines and historical profiles, or matrices, can be useful means of reconstructing the context and programme operations that occurred. They can also be used to gain an overview of the actions of various agencies, enabling assessors to understand overall impact. The example from northern Ghana in Table 5.2 illustrates the perceived relationship between the outbreak of a bout of ethnic conflict and the actions of the government and international and local NGOs. It traces the activities of local and external agencies and also records a number of key events over the past 40 years. This was done by asking a group in the community to recall key events and local and external activities over time. As with other techniques of this kind, time lines are a means to provoke and focus discussion. The actual output or diagram is usually much less important than the dialogue and insights that it stimulates.

In this case the method helped to reveal the limited involvement of government, and local and international NGOs in the area until the conflict of 1994/95. It also provokes a number of questions as to appropriateness of the overall response in meeting emergency needs and in terms of the likely long-term response to the structural issues which provoked the conflict.

Table 5.2: Time line of Demon village, northern Ghana (April 1997)

Year	Key events	Activities supported by local organisations	Activities supported by external agencies		
			Government NGOs	Ghanaian	International NGOs
1960			Set up primary school		
1988	Famine Windstorm				
1989					Two boreholes
1990		Formed welfare association			Built day nursery
1991				Set up women's group	
1992		Organised youth association		Distributed cashew seeds Started credit scheme	
1993		Organised men's <i>susu</i> group			
1994	Ethnic conflict	Formed farmers' association	Started night schools		Distributed corn seeds Renovated primary school Provided fertiliser
1995	Ethnic conflict Windstorm Flooding		Provided fertiliser	Peace and reconciliation work Started revolving fund Set up women's groups	Provided farm inputs Supported resettlement Started revolving fund Set up feeding schemes for school
1996	Windstorm		Delivered Food Aid KVIP	Started revolving fund Distributed cotton seed	Provided roofing sheets for chief's place
1997				Set up structure for livestock	Financed blocks for teachers' building

Scoring and ranking

Although in an emergency it is important to quantify absolute numbers of for example, the displaced, the dying, or the prevalence of disease, one must also understand changes in knowledge, attitudes, and perceptions. Although these may be difficult to quantify, it is not impossible.

The Wajir case study managed to quantify and compare perceptions of relative change. Two hundred pastoral families were simply asked whether they felt the conflict situation was better or worse than ten years ago and to score the situation ten years ago and now, using up to ten stones (with one signifying 'bad' and ten signifying 'good'). In both project and non-project sites, the situation ten years ago was assigned three marks, compared to an improvement to seven marks now. So the perception in both areas was that the situation had improved, whereas other parts of the impact assessment had revealed important differences in the welfare of respondents in project and non-project sites. Comparing these different findings allowed the assessors to propose that the greater perception of security, which was felt in both areas, could not explain the other differences observed in the project and non-project sites. This strengthened the argument that the observed changes were due to the programme they were assessing. This example illustrates how this sort of scoring can provide important insights into people's perceptions of change over time. It also shows how, by comparing the perceptions of different groups, questions of attribution can be addressed.

However, practitioners should always remember that this sort of scoring provides an understanding of relative, not absolute, change. In the above example, people generally feel that the situation has improved, but we cannot be certain that there is any agreement between respondents about what the score of one or ten means as these were not defined in advance. Average figures may conceal great variations between respondents' answers, particularly if some groups are likely to face different threats to their well-being and security than others. For example, women and children's perception of their own personal security in emergency situations may differ dramatically from that of men.

One way of ensuring that the views of different groups are represented is to undertake a well-being ranking, where a number of key informants help determine the characteristics of a range of groups (see Chapter 4 for more details). Although these types of ranking normally focus on material well-being and assets (for example, as illustrated in Table 5.3) it is possible to include political vulnerabilities related to, say, ethnicity or religion, as well as other aspects of identity such as gender. These issues will be of particular relevance in monitoring the situation of those most vulnerable, and in assessments of differential impact.

Table 5.3 is based on a well-being ranking exercise carried out in four refugee communities in Ikafe, Uganda, with a range of informants, most of whom were women. It indicates that people who were employed by aid agencies or who owned enough capital to start small businesses were considered among the most well off, while those without assets or sufficient labour to farm, or who are physically disabled are considered the least well-off. This sort of information can be very useful in setting a sampling frame for continued monitoring of the programme's impact on different groups.

Table 5.3: Well-being ranking from Ikafe refugee settlement, northern Uganda

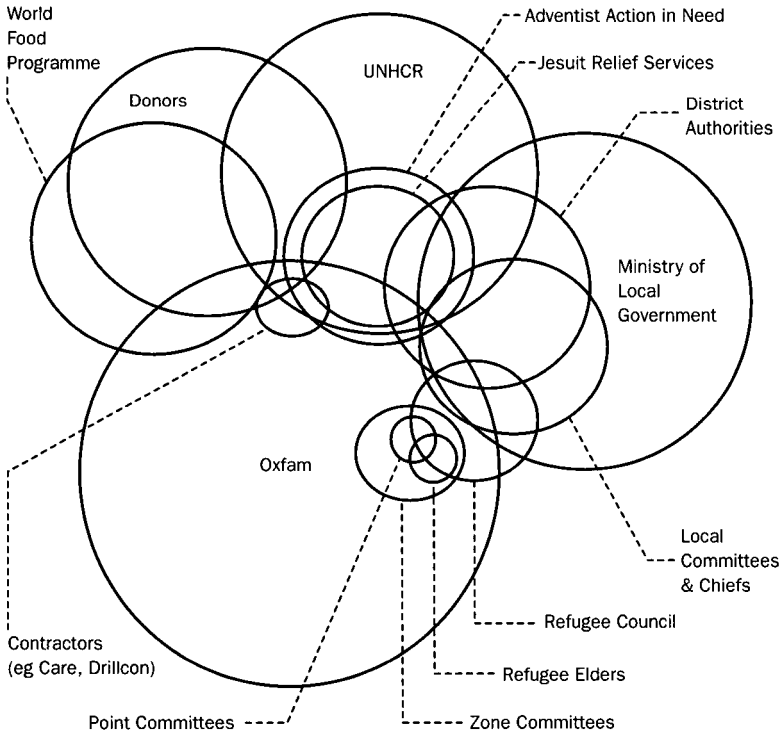
High well-being	Average well-being	Low well-being	Very low well-being
People employed by an international agency in Ikafe/Imvepi	Families with enough labour to farm the land that has been given to them	Those with no capiatla to engage in any petty trading activities	Those who have sold all their assets Those who consumed seeds and
People with enough capital to run a small business (a hotel, restaurant, and so on) in Ikafe or Imvepi	Those who have a small amount of capital to trade (selling beer, fish, tea) in the refugee settlement Those who have a technical skill (mechanic, bicycle repairer, carpenter, teacher) and can earn money from this activity Those who own some livestock (goats, a cow, some chickens) Those who are strong (able to construct their own house, latrine, and shower shelter) and can sell their labour to Ugandan nationals	Those families who are short of labour (due to ill health, disability, death) Those who have been transferred to areas where land is infertile	sell all non-food items in order to buy food Those who are physically disabled and can not sell their labour Those families who don't have labour even to farm the land they have been given access to (for example, unaccompanied orphans, weak widows, disabled people)

Chapati or Venn diagram

Emergency situations can involve a multiplicity of stakeholders pursuing different interests, and understanding who they are, how they relate to each other, and how they are perceived, is an essential part of the impact assessment process. One way of gathering this sort of information relatively quickly, and stimulating discussion on the issue, is through the use of chapati or Venn diagrams, as illustrated in Chapter 4.

Figure 5.1: Venn diagram of relationships in the Ikafe programme according to Oxfam managers

Dated April 1996



Note: The diameter of the circle indicates relative (perceived) influence or importance; the overlap indicates (perceived) intensity of contacts

The example in Figure 5.1 from Ikafe illustrates Oxfam managers' perceptions of the relative influence and inter-relationship between the substantial number of agencies and groups involved in the programme. It was part of the early work done during the impact assessment process; along with other methods it helped establish the key actors who needed to be involved in the assessment, as well as Oxfam staff members' perception of them. The refugees' relative lack of power compared to Oxfam, UNHCR, the donors, and the local government is clearly brought out in this diagram, and was a major concern for the assessors who had designed an assessment process that deliberately sought refugees' views.

Flow charts or mind-maps

Chapter 4 showed how flow charts, and the comparison of flow charts prepared by different groups (by men and women in Figures 4.2 and 4.3) can help deepen understanding of cause-effect relationships and reveal secondary impacts that may not have been intended. Given the turbulence of emergency situations, the high likelihood of unexpected impacts, and the fact that hard evidence may be lacking, flow charts can be a useful tool in promoting discussion about impact. (Like other participatory tools, they need to be used in conjunction with other methods.)

The example in Figure 5.2 from Ikafe reveals the complexity of the host population's perceptions of change brought about by the influx of refugees, and the project designed to support them. In addition it provides a holistic understanding of the context which can become a starting point for distinguishing between change that can be directly attributed to the project, such as increased employment, and change that can generally be attributed to the refugee influx, such as greater insecurity. It also provides space for exploring unexpected or unintended impacts — in this case, these included the violation of cultural sites and the perceived decline in moral standards — to be considered alongside those that were expected, such as better social infrastructure.

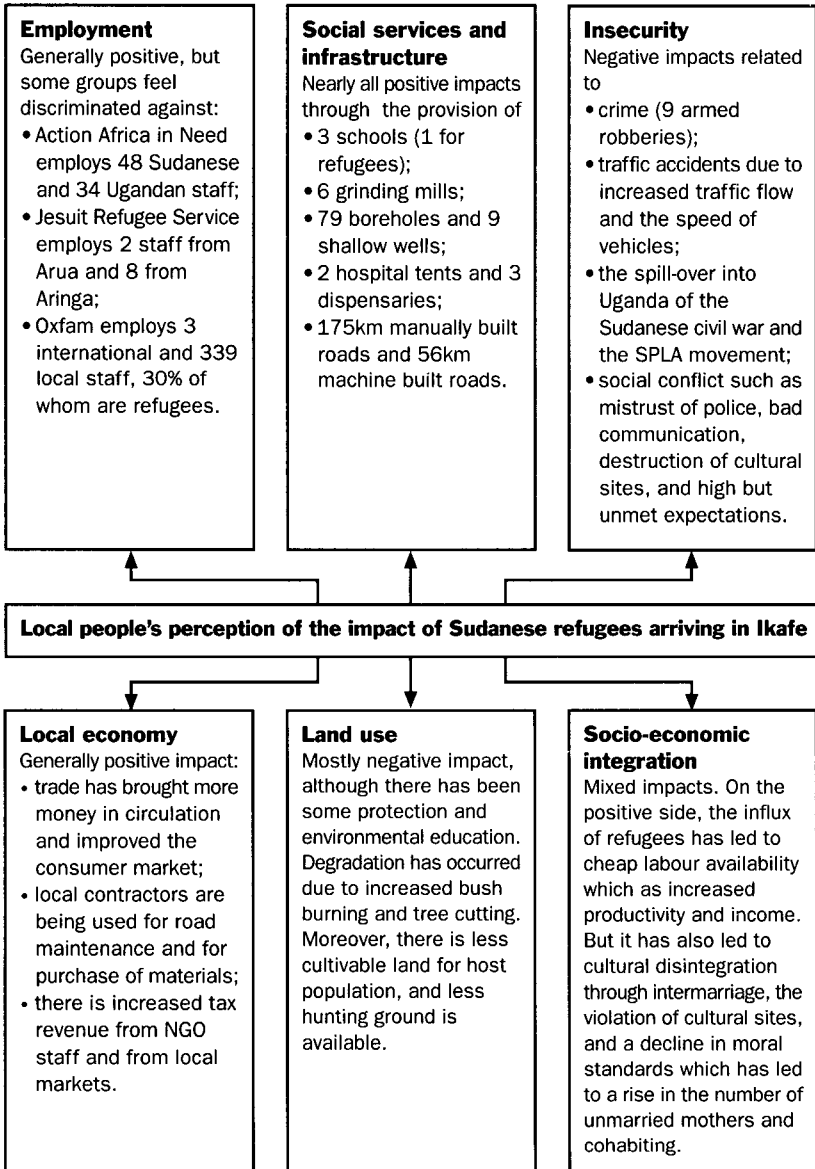
It also illustrates the degree to which local people and refugees in Ikafe attributed many of the things happening in their environment to the project. As one of the team involved in the study stated,

... the Sudanese refugees and local Ugandan people seemed (themselves) to want to attribute everything to the project intervention. This was mainly because they thought that their lives had been negatively affected since they had been moved from Kiboko to Ikafe (rather than the Oxfam activities per se). (Ros David, personal communication)

This in turn can become a bone of contention with staff who feel that the views of the refugees, and host population, about the project do not refer to those issues over which the project, and project staff, have any control.

Figure 5.2: Ugandan host population's perception of impact of refugee arrival and programme responses by aid agencies

Adapted from Ikafe case study



The Ugandan host population in Ikafe identified six main areas of impact in the diagram, most of which are seen to have both positive and negative elements. The first relates to employment by the aid agencies, in terms of absolute numbers and in terms of the employees' origins: Were those employed refugees or local Ugandans? Which district (for example, Arua) and group did they belong to (Aringa or Terego)? The second area of impact refers to the generally positive changes in the local economy, notably the general increase in the amount of money circulating because of the aid agencies' presence. The third category outlines the mainly negative perceptions of changes in land use brought about as a result of the settlement, including increased environmental degradation and reduced land for hunting. Fourth, the host population identified changes in socio-economic integration, in particular negative aspects such as declining morality, cultural disintegration, and violation of sacred sites. They saw some positive elements in the availability of cheap labour. The fifth element indicates the wholly negative perceptions of declining security through increases in crime, road accidents, social conflicts, and the movement of rebel troops. Lastly, people recognise impact related to better social services and infrastructure including roads, dispensaries, wells, grinding mills, and schools.

Trend analysis

A 'before and after' analysis referred to in the example of perceptions of relative change in security in Wajir, can be of limited value in rapidly changing emergency situations, because there can be great fluctuations. In order to understand rapidly changing circumstances impact tracking or monitoring needs to be carried out, and programmes adapted as a result.

This can also be done retrospectively, as in the time line from Northern Ghana in Chapter 4, where villagers reviewed change over a number of years, and assigned scores relating to key elements of their livelihoods. This kind of exercise can provide an insight into perceived changes in people's quality of life over time, and suggest causes for these. This then allows for factors outside the project or programme to be explored, situating a project in its wider context and helping to minimise any exaggeration, or underestimation, of project impacts.

Lessons relating to participatory tools

The call for greater involvement of 'affected populations' — local actors and organisations — in assessing impact is increasingly made, complemented by a recognition of the need to assess the different impacts that projects can have on groups such as women, men, children, and older people. In addition, the views of those who did not benefit from the assistance should be deliberately

sought out as they may reveal problems with the agencies' targeting of beneficiaries. The case studies, and Oxfam's experience in the Great Lakes region, reveal that the potential for involvement varies at different stages of the emergency. In the Great Lakes crisis, obstacles included being obliged to work through camp authorities, many of whom were implicated in the genocide as members of the pre-war Rwanda government, and working in fragmented and divided communities in Rwanda itself (Collins et al. 1998). As Christopolos points out, a naïve approach to participation in situations of great abuse of power can be counterproductive.

Agencies that claim to follow a pluralistic view may also end up acting in a relativist manner due to a naïve borrowing of concepts and working methods from liberal democratic structures and transposing them to contexts where crude and unobstructed abuse of power abounds. There is no effective technique to empower someone who has a gun pointed at his head. Simple populist rhetoric — let the people decide — creates easy prey for those immoral characters who wish to manipulate situations to their personal benefit ... Accountability to beneficiaries and local institutions must be based on listening to the voices of the people in the camps, the burnt out villages, the ruins. Participatory approaches have a major role to play here. Caution is called for, however, to ensure that faith in methods does not blind us to those who will abuse them. There is no special technique that we can apply to vaccinate ourselves against the manipulation of participation to immoral ends. (Christopolos 1998)

People affected by trauma, political and sexual violence, threats and social dislocation may not only be scared to talk, it may be positively dangerous for them to do so. Therefore approaches to participative research and evaluation need to be modified and adapted. This may mean among other things:

- ensuring that a range of stakeholders is involved in the process of sampling,
- being aware of the political ramifications and risks of the assessment, and how these affect the credibility of the findings in the eyes of various stakeholders;
- seeking advice about who to talk to, and how to go about it, from trained trauma counsellors as well as those who genuinely represent these people;
- ensuring the confidentiality of interviews and interviewees;
- avoiding particularly traumatic and distressing areas of discussion;
- cross-checking information by using a variety of sources and methods (see below).

Despite these real difficulties, it is usually possible to increase the involvement of local people. In terms of impact assessment, this often brings out critical information. For example, Oxfam's water provision programmes in the Great Lakes region usually monitored the quality of water at source, rather than at point of use. By monitoring at point of use and talking to users, assessors found out that in some camps there were major variations in the quality of water consumed due to how it was carried and stored. This led to improved practices, such as adding chlorine to water tankered to site, and a jerry-can cleaning campaign (Collins et al. 1998). In Tanzania, focus-group discussions with men and women in two camps revealed that long queues at water-collection points contributed to attacks on women.

These examples illustrate that — even if opportunities for the full involvement of local people may be limited due to the political and safety situation — a better understanding of their situation and constraints can provide important feedback for project design and management. In turn, these limited initiatives can expand over time as and when it becomes possible. Given the fact that key groups are consistently *not* involved in programme design and implementation — particularly women, older people, and children — it is even more crucial that critical attention is paid to their involvement in impact assessment processes. This means being constantly aware of, and exploiting, the changing opportunities that exist for their involvement, and devoting adequate resources and time to it at the outset.

Case studies

As we saw in Chapter 4, case studies are not really a tool or method but more an approach that is particularly useful in complex situations where a large number of variables are inter-related. This makes the case study approach especially suited to emergencies. They necessarily involve many of the tools and methods described above.

Case studies in emergency situations often fall under the critical case type described in Chapter 4. Although emergency situations are usually unique they often produce intense managerial and organisational stresses that test the robustness of existing systems to the limit. In this sense emergency cases often have wider relevance and can throw light on critical organisational and policy issues which have high political impact. For this reason undertaking analysis across a number of case studies can be particularly important in drawing out generic lessons, and in supporting processes of institutional learning.

The importance of cross-checking and feedback

As explained in other chapters, cross-checking or 'triangulating' findings is of paramount importance. In emergency situations, which may be even more

dependent on qualitative methods of data collection, and where the manipulation of information may be even more in evidence, it becomes even more vital. As detailed in Chapter 3, cross-checking can be done by deliberately using a variety of types and sequences of research methods, researchers, respondents, or sources of information.

Since contradictions in findings may be even greater in emergency situations, the impact assessment design may have to be adapted in order to cross-check them as they emerge. Those managing impact assessment processes must therefore incorporate a high degree of flexibility in their design and be open to share their findings and give stakeholders the opportunity to respond:

[D]espite every effort to obtain all relevant information, there will be information which the team did not find, think to find or gain access to. There will inevitably be information that has been misinterpreted by the evaluators. The process of distributing the draft reports for comments and discussion is an integral part of evaluation. (Dabelstein 1996, p.291)

Criteria and indicators

The experience of the case studies and recent literature on the evaluation of humanitarian work underline the importance of adapting and reformulating indicators in the light of changing circumstances. However, four key area of change require evaluation in all contexts. These are:

Mortality and morbidity rates. These are particularly useful indicators in the first phases of a crisis or in refugee camps and other settings where the risks of widespread disease and contagion are high.

Coverage and differential impact. Agencies need to know who has and who has not been covered by a particular intervention, and what the impact is on different groups particularly men and women, young and old, displaced and host populations.

Protection and security. It remains a major challenge to resolve the issue of intervening, or persuading others to intervene, in ways which can enhance peoples' right to security in emergency situations. Changes in the levels of sexual violence and rape, destruction of homes and property, and population movements are likely indicators of the general security situation. Again, particular attention must be paid to human-rights violations of vulnerable groups which may not only include women, older people, children, and ethnic minorities but also males of fighting age. Faced by violent conflict and human-rights violations, various sections of the population will have different vulnerabilities. Women may be targeted for

rape to undermine a community (Uganda, Bosnia), men may be raped as a tactic of terror (Sierra Leone), young girls may be taken as sexual slaves by fighting forces, and young boys coerced into fighting and abused as part of the initiation (Liberia), males of fighting age may have to move from house to house never sleeping in the same place twice for fear of being press-ganged (Bosnia), ethnicity may become a cause of violence (former Yugoslavia, the Great Lakes region), and the elderly may no longer receive the customary social respect and care.

Sustainability or 'connectedness' with longer-term issues. This describes the extent to which humanitarian work is undertaken in ways that support durable recovery; several areas of assessment within this have been suggested. These include

- net benefit — an assessment of the overall short-term and long-term cost and benefits;
- capacities and vulnerabilities — the degree to which long-standing individual, community, organisational, or even national vulnerabilities have been reduced and existing capacities increased;
- changes in gender dynamics — the degree to which positive changes that may have occurred have been maintained and reinforced or negative changes countered or reduced;
- peace, conflict resolution, and justice — the degree to which the programme has contributed to building understanding and social cohesion between conflicting groups, and reduced the tendency to resolve conflict through violence.

There is currently some controversy over the degree to which NGOs, in particular, can or should deliberately seek to build capacities, change gender dynamics, or promote peace during conflicts, or whether they should stick to a minimalist humanitarian relief role (Slim 1998, Duffield 1996, Macrae 1996). It seems unlikely that a categorical position either way is very helpful, although this does not negate the need for NGOs to understand the broader and longer-term consequences of their work, *whether they are intended or not*. However, it is often only possible to see broader impact with hindsight — predicting long-term or secondary effects in volatile situations is notoriously difficult. In the humanitarian arena, not responding to immediate threats to life on the — often hypothetical — grounds that it may contribute to prolonging a conflict is usually not ethically acceptable (Slim 1997). The provision of humanitarian assistance does not absolve warring parties and their supporters of their moral and legal obligations under International Humanitarian Law and the Geneva

protocols. Nor does it absolve the international community from its duty to help bring conflicts to an end and to resolve the root causes that brought them about. In fact, it makes it even more important to gather evidence about the broader impacts of humanitarian assistance in situations where other actors do not comply with their moral and legal obligations. Not only will this help humanitarian agencies adjust their operations in the light of the findings, but, more importantly, it will contribute evidence to any potential process of legal redress. This sort of analysis can also illustrate the cost of the international community's failure to create a more 'rounded' political and economic, as well as humanitarian, response to complex emergencies.

Tools and methods

Emergencies in general, and complex political emergencies in particular, present severe challenges to orthodox approaches to choosing tools and methods. The examples cited in this chapter indicate how methodological innovation and adaptation to the specific contexts of emergencies can occur and result in important insights being gained. This is significant, given the recent findings by Hallam and Borton that

[c]urrent monitoring and reporting systems for humanitarian assistance programmes often do not take full account of the needs of ex-post evaluations and the effectiveness of the evaluation process is considerably hampered. (Hallam and Borton 1998)

Overcoming this requires improving monitoring systems, developing indicators and standards, and better information management. It should also be remembered that impact assessments may in themselves have little impact on organisational practices and policies if they are not followed up. It is necessary to set up a clear and systematic process, with a timetable and allocation of responsibilities for discussing and agreeing findings, implementing recommendations and changes, and adapting policies and procedures. An agreed monitoring process for ensuring that this happens is also sometimes required: for example, in the case of the multi-donor evaluation of the crisis in Africa's Great Lakes Region, the Joint Evaluation Follow-up Monitoring and Facilitation Network (JEFF) was set up for post-evaluation monitoring.

If NGOs and other humanitarian agencies are to respond to the challenges thrown down by recent critiques of their work they need to be clear about their values and ethics and how these relate to international humanitarian law and the Human Rights and Refugee Conventions, as the evaluation of Oxfam's Great Lakes programme suggests:

In the case of humanitarian organisations the base line for determining humanitarian intervention should derive from International Humanitarian Law (IHL), Human Rights Conventions (HRC) and the Refugee Convention (RC).⁵ These provide basic, but absolute, rights to protection and assistance for people affected by conflict. They constitute an internationally agreed general reference point for humanitarian intervention— an obligation on and a right of the international community ... IHL, HRC and RC provide the internationally accepted norms which justify intervention and determine accountability for those who intervene. (Zetter 1998)

This must then be complemented by becoming more vigilant, at an operational level, about setting consistent objectives and monitoring both changes in the context and in the impact of an intervention, and how these weigh up in the light of the growing body of standards that NGOs are starting to set for themselves.

Ethics and organisational practices

If impact assessment has to incorporate considerations of ethics and organisational practice, how would this work? Hugo Slim has set out a possible framework for ethical analysis (see Summary 5.1). Slim reminds us that relief agencies in conflict situations are always responding to the violence of others, and that the difficult moral choices they face are normally due to the immoral choices of others. He outlines the important characteristics of a moral dilemma (a choice between two evils) and distinguishes it from other 'tough choices' that can masquerade as moral dilemmas. In impact assessment, being clear about the types of choices and decisions that are or were available to agencies is a first step in analysing its actions and responsibilities in a given situation.

The next step is to determine what code of ethics the organisation adheres to and whether it is essentially 'duty-bound', like most of the medical profession or the ICRC, which undertake certain actions because they believe they are intrinsically right. The responsibility for the consequences of those actions is more dispersed; for instance, a doctor cannot refuse to treat a patient because of what he or she may do in the future. The alternative to being duty-bound is to be 'goal-bound' like many development NGOs, which involves a complicated and uncertain process of anticipating the wider outcomes of one's actions and holding oneself, at least partially, responsible for their long-term consequences. Once this basic point is addressed, other questions (relating to an organisation's motivations, prior knowledge, capacity, deliberation, mitigation, and organisational conscience) may be relevant. Answering these will allow a systematic review of an agency's ethical standards.

Slim's framework provides a useful if untested tool for undertaking an element of institutional and policy analysis that is usually lacking. When combined with other tools and approaches for looking at organisations and impact assessment, such as those described in Chapter 7, perhaps it will help to make the link between significant changes in individuals' lives and the organisational context in which they occur. An organisation's values and ethics (or lack of them) are an important indicator not only of how decisions are taken but also of which projects or policies will be delivered, and what value is placed upon the impacts that are achieved.

Summary 5.1: A framework for ethical analysis adapted from Slim (1997)

What code of ethics?

Is the organisation secure in the belief that its actions are always good in themselves? Or does it believe that it needs to have a sure grasp of the wider consequences of its work to be certain of the goodness of its work? Answers to this will determine how much moral responsibility an agency assumes for conditions surrounding its work.

How are different intentions and motivations balanced?

Is the agency acting out of the best motives and intentions or doing the right thing for the wrong reasons? How does it balance different organisational motives such as compassion, income, publicity, and influence; which tend to win out?

What knowledge is sought?

To what degree has the agency made every effort to gather all information relevant to its decision-making? If the agency did not have all knowledge at its disposal could it have reasonably been expected to collect it?

What capacity is available?

Did the agency have the capacity (resources, influence, power) to do anything about it? An agency can only be held responsible for not doing something it could have done, but chose not to do.

Is deliberation valued?

Did the agency make every effort to counsel, consult, debate, and weigh carefully the various aspects of the problem? A commitment to ethical deliberation should be a part of responsible programme design.

Did the agency seek to mitigate any negative impacts?

Did the agency seek to minimise any likelihood of negative impacts? In a choice between 'two evils', did it seek to limit the damage caused?

Is there an organisational conscience?

Do senior managers promote a moral ethos and ensure that staff have the ethical skills and a knowledge of past dilemmas to increase the likelihood that they will make sound decisions in difficult moments?

As far as other organisational practices are concerned, recent work by the Sphere project recognises that 'without clearly defined, systematic organisational policies and procedures the implementation of high quality programmes will not be possible' (The Sphere Project 1998). This approach complements Slim's ethical framework and the project has developed a set of organisational best practices which provide the means to achieving agreed minimum standards for humanitarian assistance. These are described in Summary 5.2.

Summary 5.2: Organisational best practices for humanitarian agencies (Sphere project 1998)

Philosophy

The agency's goals reflect universal humanitarian principles as outlined in the humanitarian charter and the Code of Conduct.

Accountability

The agency is accountable to the recipients of its services, to staff, and to donors.

Management and support of staff

The agency recognises that the effectiveness and success of humanitarian programmes depends on all the people working for them. Human resources issues are therefore central to strategic plans.

Analysis

Interventions are based on an understanding of the context and a clear analysis of need. Any response is based on the culture(s), needs, and rights of those it is intended to assist.

Co-ordination

There is commitment to a co-ordinated response, which takes into account the views of all the stakeholders and where local authorities take a lead if appropriate.

Participative programme planning

Programming decisions are consistent with the agency's philosophy and involve all stakeholders in the process. Programme characteristics are determined through application of the planning cycle which facilitates monitoring and accountability.

Gender integration

The agency considers gender aspects in all programming decisions.

The special needs of children

Programmes address the special needs unique to children affected by calamity or conflict.

Environment

The agency takes a proactive approach to environmental concerns in all programming decisions. Prevention of environmental damage should be the norm rather than the exception.

Protection

The agency demonstrates concern for the physical safety of disaster-affected persons. Although protection is the responsibility of states, agencies have a responsibility to be knowledgeable and aware of the implications of interventions.

Security

The agency demonstrates concern for the security of local and expatriate staff.

In each of these areas, Sphere gives detailed information which attempts to further define best practice. This is cross-referenced to thematic minimum standards (for example for water supply, nutrition, food aid, shelter, health-care, and site planning) so that each thematic area has standards of analysis for assessment and monitoring and evaluation processes; minimum standards related to the thematic area in question; and some related capacity and competence standards on the part of the agencies involved. Although many of the standards and associated indicators are not impact indicators in the strict sense, they are qualitative and quantitative output measures which, if

reached, could reasonably be expected to make a significant difference to people's lives in emergency situations.

These sorts of standards and guidelines for best practice can, therefore, make an important contribution to assessing performance and in the development of clearer accountability. However, NGOs are increasingly combining operational relief efforts in humanitarian situations with lobbying and advocacy work, so any assessment of overall impact has to also explore this other aspect of their work. The next chapter attempts to do this in some depth.