How to generate statistics and influence policy using participatory methods in research

Statistical Services Centre

Working Paper

Carlos Barahona (Statistical Services Centre, The University of Reading)
Sarah Levy (Calibre Consultants, Reading)

November 2002

Statistical Services Centre
The University of Reading
PO Box 240
Reading RG6 6FN
United Kingdom
www.reading.ac.uk/ssc

Summary

This paper aims to show that it is possible to generate statistics which will be taken seriously by policymakers from research using participatory methods. A key requirement is to produce results from a representative sample, which can be generalised in order to reach conclusions for the population of interest. This implies working in a larger number of sites than is common for most studies which use participatory methods. However, the Malawi experience presented in this paper shows that it is possible to organise such studies at a reasonable cost.

Key requirements are for the study design to incorporate statistical principles; and for PRA tools to be adapted to meet the demands of standardisation and comparability of data produced across sites. We argue that if research studies using participatory methods follow this approach, the data generated will be suitable for standard statistical analysis. The statistics produced by such studies should be capable of informing policy at national level. However, there are some concerns about empowerment and ethical issues, which present challenges for the future.

Acknowledgements

We would like to thank Robert Chambers of the Institute of Development Studies at The University of Sussex for his interest in our Malawi experience and for encouraging us to write this paper; Harry Potter of the Department for International Development (DFID)-Malawi for commissioning the studies with the aim of promoting evidence-based policy making; Ian Wilson of the Statistical Services Centre (SSC) at The University of Reading for his contribution to the design of the studies, his thoughts on sampling and his comments on the first draft; Roger Stern of the SSC for his comments on the first draft; the Malawi and UK-Malawi research teams for their dedication to excellence; and the people of Malawi for the time that they invested in participatory discussions, with the hope that they would inform better policies for the future.

Table of contents

Summary ... 1
Acknowledgements ... 2
Table of contents ... 3
List of tables and figures ... 5
1 Introduction ... 7
1.1 The approach ... 7
1.2 Our studies ... 7
1.3 What this paper sets out to do ... 8
1.4 When to use participatory methods for research ... 9
1.5 Our experience and future challenges ... 11
2 Statistical principles and generalisation ... 12
2.1 The challenge of generalisation ... 12
2.2 A matter of definition ... 13
2.3 An issue of fairness and transparency ... 13
2.3.1 What is the magic of probability-based selection? ... 14
2.4 Managing uncertainty ... 14
2.5 Conclusion ... 14
3 Sampling ... 15
3.1 Sampling in research using participatory methods ... 15
3.2 Population definition and the basic unit of study ... 16
3.3 Structure ... 16
3.4 Stratification ... 17
3.5 Sampling strategy – a Malawi example ... 17
3.6 When participatory methods help in sampling ... 18
3.7 Within the community ... 20
3.7.1 Voluntary versus random approaches ... 21
3.7.2 When community mapping is not applicable ... 22
3.7.3 When sampling is not possible ... 23
4 Sample size ... 24
4.1 Precision and variability ... 24
4.2 Population size is irrelevant ... 27
4.3 Other factors ... 27
4.3.1 Structure of the units of study ... 28
4.3.2 Level of aggregation ... 29
4.3.3 Availability of resources ... 29
5 Adapting PRA tools ... 30
5.1 Building bridges ... 30
5.2 Standardisation ... 31
5.3 Comparability ... 32
5.3.1 Trend analysis ... 32
5.3.2 Scoring versus ranking ... 33
5.4 Community mapping with cards ... 34

5.4.1 An example of community mapping with cards 36
6 Planning, organisation and cost 43
6.1 Planning 44
6.2 Preliminary design 45
6.3 Main fieldwork phase 46
6.4 Analysis and presentation of results 47
6.5 Cost considerations 48
6.5.1 Comparison of costs 48
6.5.2 Cost control 50
7 Data analysis 50
7.1 Data management 50
7.2 Statistical analysis 51
7.3 Scaling up 53
8 Empowerment 54
9 Ethical issues 55
9.1 Ignoring the evidence 55
9.2 Transparency and consent 56
9.3 Flexibility 57
9.4 Sharing information 58
9.5 Confidentiality 58
10 Conclusion 59
List of references 61
Appendix 1: Malawi study outlines 63

List of tables and figures

Figure 3.1: Hierarchy and sampling approach (SP2 Module 4) 18
Figure 4.1: District A..... 25
Figure 4.2: District B 26
Figure 5.1: Integration of PRA and statistical design of experiments 30
Picture 1: A community map (TIP2 Module 2)..... 34
Picture 2: Household cards (TIP2 Module 2) 37
Table 5.3: Relationship between receipt of TIP and food security status 39
Table 5.4: Relationship between receipt of TIP and sex of household head..... 40
Table 5.5: TIP recipients – criteria used by non-recipient focus groups 41
Table 5.6: Code 10 analysis at national level..... 41
Table 5.7: Variability in the proportions of deserving TIP recipients across sites..... 42
Figure 6.1: Stages and activities 43
Table 6.1: Hypothetical research study using participatory methods..... 49
Table 6.2: Hypothetical survey of 3,000 households in 135 sites 49

1 Introduction

1.1 The approach

The focus of this paper is an approach that we will refer to as 'research using participatory methods'. The studies on which it is based worked with participants from rural communities, but they were not community-led in the sense of having an agenda determined by the participants and evolving in response to local development needs. The emphasis was on research outputs which would inform national government policies. With the sort of studies that we describe, there is often an input by the communities into moulding the agenda in a preliminary phase of research, but the key questions to be researched are usually set by policymakers, donors, researchers and other stakeholders. Nevertheless, the researchers try to maintain the main principles of participation (Chambers, 1994a).

1.2 Our studies[1]

The six research studies using participatory methods that we co-ordinated in the 1999-2000, 2000-01 and 2001-02 agricultural seasons in Malawi (see Appendix 1) were part of an evaluation of Malawi Government programmes that provided 'Starter Packs' of free inputs (fertiliser and seed) to smallholder farmers in rural areas[2]. The evaluations were funded by the UK Department for International Development (DFID). Four of them were carried out by Malawian teams, and two were conducted by UK-Malawian partnerships.

The studies aimed to answer policy questions about the free inputs programmes and about safety nets, as Malawi's National Safety Nets Strategy intended to incorporate a

[1] In this paper, these studies are referred to as 'our studies' because they were designed by ourselves together with the Malawi or UK-Malawi research teams. When referring to a specific study we use the name of the team leader for that study, e.g. Cromwell (2000), or the year and module reference, e.g. 1999-2000 Starter Pack (SP2) Module 4 (see Appendix 1).

[2] "The 2000-01 and 2001-02 Targeted Inputs Programmes (TIPs) provided rural smallholder households with one Starter Pack containing 0.1 ha-worth of fertiliser, maize seed and legume seed. The TIPs followed on from the Starter Pack campaigns in 1998-99 and 1999-2000. A key objective of these campaigns was to increase household food security amongst rural smallholders in Malawi. The 1998-99 and 1999-2000 Starter Pack campaigns were designed to cover all rural smallholder households, providing 2.86 million packs each year. The 2000-01 TIP (TIP1) was enough for roughly half this number of beneficiaries, while the 2001-02 TIP (TIP2) was further scaled down to 1 million beneficiaries". Levy and Barahona (2002).

free inputs element. These questions required answers at a national level. The key issues to be researched included:

- What was the impact of the free inputs on the beneficiaries, in particular the poorest and most vulnerable members of rural communities?
- The relationship between free inputs and rural livelihoods, gender, diet and intra-household distribution of food.
- The impact of free inputs on sustainable farming practices.
- How many smallholder households are there in the rural areas of Malawi, and what proportion deserves to receive free inputs?
- Does community poverty targeting work for free inputs and for other Safety Net interventions, such as direct welfare transfers?
- How much of the benefit is likely to be redirected to relatively wealthy farmers?
- How should beneficiaries be selected, and how should the benefits be distributed?
- What was the impact of the agricultural extension messages accompanying the free inputs programme?

The six research studies using participatory methods which looked at these questions formed part of a wider programme of research comprising twelve studies, including five nationwide surveys and one set of case studies[3]. The 'modular' design of the research programme, in which different types of research complemented each other, was based on our view that the research method should be a function of the type of information that is required (see Section 1.4). For example, we used surveys to collect information about household food security and individual farmers' use of the free inputs, as these questions could best be addressed using questionnaires.

1.3 What this paper sets out to do

This paper attempts to reconcile some of the differences between survey-based research (often referred to as 'quantitative') and research using participatory methods (often labelled 'qualitative'). Survey-based research can generate statistics that are 'representative' of a population, and, as such, tends to be seen by policymakers as more useful than research using participatory methods, which generates results that are valid at the local level but usually cannot be generalised in order to reach conclusions for a population of interest.

Our experience in Malawi over the last three years suggests that the dichotomy is a false one. It is possible to apply statistical principles to research using participatory methods and to generate both text- and numbers-based analysis that is

[3] A CD containing the final reports of all the studies and the data collected by them can be obtained from the DFID offices in Lilongwe or from c.e.barahona@reading.ac.uk

'representative' of a population. There are major differences between survey-based and research using participatory methods, but these should not be because one approach is representative while the other is a collection of 'case studies'. By adopting certain statistical principles and making some adaptations to the PRA tools, this difference disappears in most cases.

The key difference that remains is the type of information that can be collected. Surveys collect simple pieces of data using questionnaires. Research using participatory methods studies uses discussions to explore deeper matters, often with tailor-made 'tools' or 'techniques'. Even if they generate numbers, these are the result of discussion and exchange of views rather than an on-the-spot reaction to a question.

It is important to recognise that different types of data, fulfilling different objectives, require different approaches to information collection (see Section 1.4). It is not the intention of the authors of this paper to argue that the type of research using participatory methods developed in Malawi could replace survey work. Rather, we believe that research using participatory methods complements survey work by collecting information that surveys cannot collect efficiently. The reverse is also true. It would be inefficient to attempt to collect the sort of data that questionnaires capture using participatory methods. In many research exercises, both types of information have a role to play. The challenge is to ensure that policymakers give equal weight to the findings of research using participatory methods by making it representative of the populations of interest.

1.4 When to use participatory methods for research

From the researcher's point of view, the questions that we want to explore and the type of information in which we are interested are the most important factors influencing the decision about what research method to use – a survey, participatory methods based on group discussion, semi-structured interviews or another approach. Questionnaires used in large-scale surveys can deal effectively with information that can be answered on the spot by closed-ended questions that provide a limited choice of answers[4], while complex matters requiring thought and benefiting from discussion can more appropriately be captured by research using participatory methods. For example, Module 4 of the TIP1 evaluation (Chinsinga, 2001) was interested in researching whether the work-constrained poor should receive cash or benefits in

[4] A closed-ended question is one which has a comprehensive set of possible answers, established by prior research. An 'other, please specify…' option may be included in case of doubt, but the aim is to avoid unforeseen answers which have to be post-coded. Semi-structured interviews, by contrast, use mainly 'open-ended' questions. These are not suitable for large-scale survey work, because of the laborious nature of post-coding.

kind (goods and food) as direct welfare transfers. This topic required research using participatory methods owing to its complexity (see Box 1.1).

Box 1.1: TIP1 Module 4 – Direct Welfare Transfers

For the research on whether people would prefer to receive cash or benefits in kind (goods and food) as direct welfare transfers, two focus groups were organised in each study site, one looking at cash transfers and one looking at in-kind transfers. Various aspects were considered in each focus group, including who would receive the transfers (using a card game), how they would be managed (including scoring of alternatives) and the timing of delivery (using seasonal calendars). The two focus groups then came together, and presented the case for the type of transfer that they had been discussing to the other group. The advantages and disadvantages of each type of transfer were listed, and then participants were given the chance to 'vote' for each type of transfer. The researchers were able to find out peoples' preferences for cash or in-kind transfers based on informed choices, and the reasons behind these choices. They were also able to find out whether the communities could target and manage the benefit, and when it should be delivered. A survey would not have been capable of providing reliable information about these matters. At most, a questionnaire would have been capable of asking whether people preferred cash or in-kind transfers. But there would have been no guarantee that they understood the difference between them, let alone the implications of choosing one type or the other.

For further information on the TIP1 Module 4 study, see Appendix 1.

A related issue when deciding what methods to use – although not as critical as the 'questions to be explored and type of information required' issue – is that of the unit of study. It is often suggested that if you want to find out about individuals or households, you need a survey or semi-structured interviews, while if you want to find out about communities or groups within them, you need to use research using participatory methods. However, in our experience, participatory methods can be adapted to collect some basic information (including numbers) about individuals or households in addition to community or group information, while semi-structured interviews are often useful for asking questions about the community or group (e.g. when the interviewee is a key informant, who can be expected to provide information not only about him/herself but about the people who he/she works or lives with). In a developing country context, surveys are not efficient at providing information about

communities or groups. Thus, if the unit of study is the community or group, key informant interviews or research using participatory methods are essential.

On the other hand, research using participatory methods is not efficient at providing detailed information at household or individual level or 'lots of numbers'. Individual, household and numerical information needs to be kept relatively simple in research using participatory methods. For example, in TIP2 Module 2, a mapping exercise was used to place the households in each site in 'food secure', 'food insecure' and 'extremely food insecure' categories. This was highly effective in establishing what proportion of the population fell into each category (see Section 5.4.1). However, it would not have been possible to find out from the mapping exercise participants what strategies were adopted by each household to cope with food insecurity. It would have been more practical to collect this type of information using a questionnaire. Therefore in some cases we attach a small-scale survey to a participatory methods exercise. For example, to find out about the coping strategies of households identified as extremely food insecure in a mapping exercise, we could survey a sample of these households.

Thus, questions that need to be posed by the researcher at the beginning are:

- What questions do I need to explore/what type of information do I require?
- What is my unit of study (individual, household, village, community, group)?

It is important to emphasise that we do not consider the choice between survey and research using participatory methods to be determined by whether we wish to convince policymakers that the findings are representative. If the approach set out in this paper is adopted, the findings from research using participatory methods have the potential to be representative of the population of interest.

1.5 Our experience and future challenges

Sections 2-7 of this paper discuss the theory and practice of the 'research using participatory methods' studies that we have developed in the past three years in Malawi. The common aim of the six studies undertaken during this period has been to collect information from smallholder farmers in a way that policymakers and donors can accept as 'representative' of the smallholder farmer population as a whole.

Sections 8-9 discuss some of the issues for which we do not have answers, and which represent challenges for the future. Firstly, if much of the agenda is externally-determined, is this type of research compatible with the principles of 'empowerment' that practitioners of participation uphold? And secondly, what are the ethical issues raised by this type of research?

2 Statistical principles and generalisation

2.1 The challenge of generalisation

Participation has become a generally accepted philosophy of work for those interested in development. It is probably due to the success that participatory methods have had in facilitating empowerment and action at the local level that participation ideas have been exported to multiple areas of application including research[5]. The adoption of principles and methods of participation makes sense for research but at the same time produces new challenges. One of these new challenges is that of generating results that can be generalised and used to influence decision-making.

What do people mean when they say that a result is generalisable? They usually mean that information collected from a sub-section of a population, processed and analysed by the researcher can be used to reach conclusions that apply to the population as a whole. This process is called 'inference'. Information acquired[6] through participatory methods is often seen as context-dependent and not suitable for making inferences.

Results from studies based on statistical methods such as surveys are accepted as potentially suitable for making generalisations[7]. We start this discussion by exploring what are the elements that allow statistical methods to claim that their findings can be generalised. Later, we discuss how some of these elements can be adopted in research using participatory methods to allow a similar claim, and what are the consequences of this.

---

[5] In this paper we will use the word research to refer to the process collection and analysis of information with the purpose of learning and acquiring a better understanding of a situation and facilitate evidence-based decision-making.

[6] We prefer the term 'acquire' to 'extract' as in the process of research using participatory methods the researcher has the explicit and transparent intention of acquiring information that will help in the fulfilment of the research objectives. This contrasts with the negative connotation attached to the 'extraction' of information from participatory discussions and which, if secretive or hidden, contradicts the notion of sharing that underpins participatory practice (see Section 9).

[7] However, the user of such results should take the precaution of checking whether the methodology used gives the results the property of generalisability. Too often survey work is carried out with insufficient care to ensure that the results are reliable, let alone generalisable.

2.2 A matter of definition

To what population do the results apply? Statistical methods solve this problem right at the beginning of the research process. The definition of the population of interest or 'study population' is one of the first steps in the design of the research.

The population is often defined in terms of the properties of its elements, their location or a combination of both. These populations go from real, easy to identify groups such as 'smallholder farmers in rural Malawi' to constructs that define scattered units, whose identification is not trivial, and that only form a coherent body in so far as sharing certain properties. An example could be 'the vulnerable poor' of the country. Whatever definition of population is agreed and used for selecting the subset of members about whom information is collected, determines the population to which the inferences made will apply.

2.3 An issue of fairness and transparency

In most cases the population of interest is too large to allow the researcher to seek information on each and every one of its members. In consequence, most studies reach only a small part of the population: a sample. It is important to be able to convince the users of the information that the sample is a fair representation of the whole population – at least insofar as the objectives of the study are concerned. In other words, that the sample is 'representative'.

Studies based on statistical methods claim that a sample is more likely to be representative if the selection is carried out by:

First Using an objective method of selection where all members of the study population have a definable chance of being part of the sample[8]. This is called a 'probability-based' method of selection.

Second Taking a sample that is large enough to capture the main features of the population as well as the divergence from the main features or 'variability'[9].

[8] The fact that each member has a chance of being selected does not necessarily mean that all the members have the same chance of selection. In statistics, the simplest case, when all the members of the population have equal probability of being part of the sample, is a simple random sample. In practice it is often not possible, desirable or efficient, to select samples using simple random sampling. Other probability-based sampling methods exist and some of them will be discussed later in this paper.

[9] Sample size will be discussed in Section 4.

2.3.1 What is the magic of probability-based selection?

The main argument for the use of probability-based selection is that it removes subjectivity from the process of selection and through this it protects the researcher from the accusation of having chosen the elements 'out of convenience' or 'influenced by the his/her own views', i.e. that the sample is not representative. Most people would be unhappy to see that a winning lottery ticket is selected by one person after examining the 49 possible numbers and choosing the winning combination with open eyes. Probability-based selection on its own does not guarantee that the sample is representative, but it does contribute to our claim that the sample is not a biased selection of elements in our population.

The question of bias is an important one. Random selection of villages for our Malawi studies ensured that we were likely to visit not only villages where access was easy and accommodation for the research team was of good quality, but also villages in remote areas where access was a major challenge and where the research team did not find comfortable accommodation facilities. Thus, we could not be accused of having biased our sample by neglecting remote areas.

Another argument for the use of probability-based sampling is that there is a large body of statistical theory that can be used to analyse data from this type of sample.

2.4 Managing uncertainty

You can be sure that if you take a sample and generalise your results to the total population you will not get it completely right. For the researcher using statistical methods the key question is not "Did I get it right?" (they should know that they probably did not), but "How far away am I likely to be from getting it right?" Statistical inferences will normally be accompanied by an indication of the 'precision' of the inference, which in practice often translates into a numerical range within which the researcher would expect to find the 'true' population value.

This 'precision' is influenced by three main factors:

1. how variable is the study population with respect to the characteristic of interest;
2. how the sample is collected; and
3. how big a sample is taken.

2.5 Conclusion

In summary, results from research based on statistical methods are accepted as applicable to the overall population because:

1. They apply directly to the population defined prior to carrying out the study and out of which the sample was taken.
2. They carry insurance against the accusation of not being representative because they use an objective, probability-based method to select the sample.
3. A measurement of the precision of the findings is given and therefore the generalisations made become 'qualified generalisations' that the user can treat as reliable or not depending on the level of precision.

For those of us interested in building bridges between participatory methodologies and statistical principles some initial questions are:

- Are the conditions that make it possible to generalise from research based on statistical methods applicable to research that uses participatory methodologies?
- Are there any consequences of adopting statistical principles that affect the fundamental principles of participation?
- Are there any ethical implications that we need to consider?

We will attempt to answer these questions in the following sections.

3 Sampling

3.1 Sampling in research using participatory methods

The need for sampling applies to research using participatory methods in the same way that it applies to survey work: if generalisations are to be made from the results of the work, adequate processes of sampling are required. A good deal of what has been written about sampling has to do with estimation of numerical characteristics. It is often written in language inaccessible to non-mathematicians and seems to have little relevance for the problems faced by those interested in research using participatory methods. However, we have found that sampling principles are useful when planning research using participatory methods. In this section, we discuss some of the key sampling issues.

In Malawi, over the last three years we have consistently used statistical sampling methods to select our study sites. Within the villages selected for studies using participatory methods, we have carried out discussions with farmers that involved varying levels of participation. We are confident that our conclusions can be generalised because:

- The results come from a considerable number of sites allowing us to look at patterns and variability (see Section 4).
- We used a method of selection of sites that avoids researcher bias and allows us to claim that the sites are representative of the study population.

3.2 Population definition and the basic unit of study

Before a sample is drawn, we need to define the population. Defining a population implies understanding what the elements are that form it and about whom the researcher is looking for information. This might be relatively straightforward in a small country where you want to run a household survey to find out if the head of the household is in employment or not. Here the basic 'unit of study' is the household and all households are probably reachable. However, in research using participatory methods, defining basic units of study is not always easy and the complexity of the issues means that more than one type of study unit may be needed.

3.3 Structure

Study units do not sit in a homogeneous background isolated from each other. It is possible to identify 'structures' of study units, which are useful for designing sensible sampling schemes. Two types of structure can be easily identified:

- The first is hierarchical. These are vertical relationships where one unit at a higher level is made up of several units at lower levels in the hierarchy. For example, individuals are found within households and households within villages.
- The second is at the same level of the hierarchy, such as classifications of units according to specific characteristics that are of importance for the research. For example, villages can be classified according to the agro-ecological zone in which they are located. Households can be classified into male- and female- headed, or by poverty levels, food security status, or whether they have been affected by an intervention like Starter Pack.

In surveys, the use of hierarchical (or multi-stage[10]) sampling is common practice. For example, a sample of districts might be drawn from all districts in the country; from within each of the selected districts a sample of villages might be taken; and from within each of the selected villages, information might be collected from a sample of households. At each of these levels, the elements are called 'sampling units'. However, often the only sampling unit that is a 'unit of study' is the one at the bottom level of the hierarchy. The unit of study in an income and expenditure survey would be the household, even if – in the selection of the households – the researcher went through a process of selecting regions, then districts and then areas within districts.

There is no reason not to use a similar approach when selecting sites for research using participatory methods. However, with this type of research, the objectives often require information collection at several levels of the hierarchy. For instance

[10] For a discussion about multi-stage sampling see De Vaus (1996), Barnett (1991), Cochran (1977). The mathematical content of these discussions increases from De Vaus to Cochran.

information might be required at district, village and household level. Elements at each of these levels would then be regarded as units of study.

3.4 Stratification

In statistics stratification means forming groups (strata) of study units in such a way that the units within one group are as similar as possible while the units in one group are as different as possible to the units in another group[11]. It makes sense to use stratification when the criterion used to form strata is believed to have an effect on the issues of research interest (one could think of stratification as further refining the definition of the study population). In theory, stratification brings two advantages:

1. results and generalisations can be made specifically for the stratum if the sample size within the stratum is sufficiently large; and
2. the variability in the overall results will be decreased by having removed between-strata variability at the design stage.

3.5 Sampling strategy – a Malawi example

Recognising the structure of the units of study in our Malawi research helped us in deciding on the integration of statistical and participatory approaches. One strategy that worked well was to use statistical sampling in the hierarchy down to the level of village and participatory methodologies for collecting information at village and household levels. If necessary, participatory methods could also be used for sampling households or individuals within the village (see Section 3.6).

Figure 3.1 illustrates the structure for our SP2 Module 4 study about the impact of the free inputs on the sustainability of the smallholder agricultural enterprise (Cromwell, 2000). In this study, we took a statistical sampling approach to the selection of villages. We used a stratified, two-stage sampling scheme where the strata were areas of the country classified into 'Spheres of Influence' by the Famine Early Warning System (FEWS) (see Figure 3.1 and Appendix 1). At the first stage, Extension Planning Areas (EPAs) were selected from within each stratum. At the second stage, villages were selected from within each selected EPA. From the village level downwards, PRA tools were used to stratify farmers into sustainability groups and collect information.

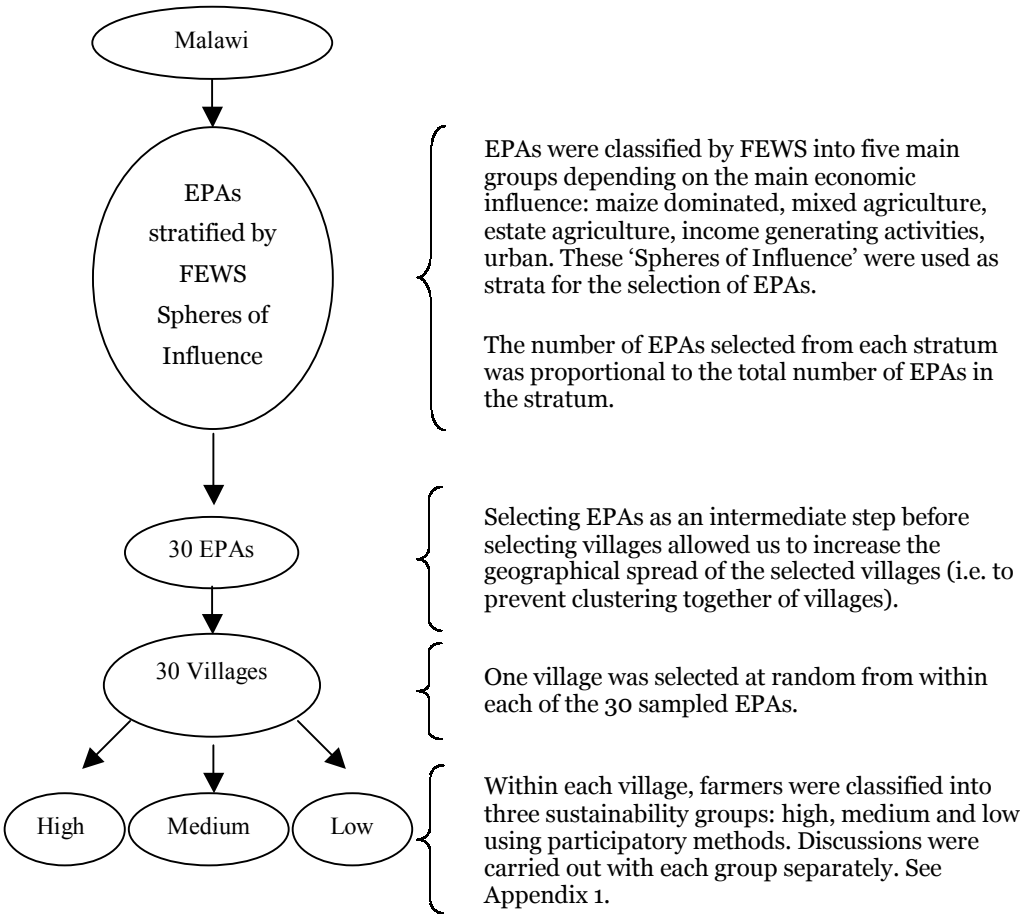
[11] Sometimes 'stratification' is based not in the principle of having units that are as similar as possible within each stratum but is used instead because grouping units seems necessary from a practical point of view. In these cases, stratification does not have the advantages described in this section.

It should be observed that the structure that the researcher identifies in the units of study is a convenient conceptualisation of the situation on the ground. Generally, the structure is built up by taking into account:

a) geographical or administrative structures; and
b) information about the study units that the researcher knows in advance and that might affect the issue under investigation.

Geographical or administrative structures often influence the decision to adopt hierarchical sampling schemes, while information about the study units within one level of the hierarchy leads to stratification.

Figure 3.1: Hierarchy and sampling approach (SP2 Module 4)



3.6 When participatory methods help in sampling

In order to draw a sample, the researcher needs a sampling frame. A sampling frame is a complete list of all the members of the population. If such a list is available and up-to-date, the researcher is able to claim that in choosing the sample from the list, all members of the population are given a chance of selection. Moreover, the researcher should be able to work out the probability of selection for each member of

the population. However, often an appropriate sampling frame does not exist, or if it exists, it is incomplete or out of date.

In the case of smallholder farm households in Malawi, it was not possible to construct a sampling frame for the study population. We solved this problem by using a multi-stage sampling scheme where the households of the farmers were selected at the last stage of the sampling process, after having chosen districts, areas and villages.

This approach assumes that at village level it will be possible to construct a sampling frame of smallholder farmers. The traditional survey mechanism for constructing the sampling frame involves visiting all households in the area, writing a number over the door of the main dwelling of the household and then listing all the households[12].

The formal approach to sampling frame construction is possible if the limits of the village are clearly defined and a definition of household is available and unambiguous[13]. However, in many development settings these two conditions are not easy to fulfil. For example, in Malawi, villages are in constant change: they grow and shrink depending on seasonality and whether there is a food crisis[14]; they merge or split up according to the dynamics of the local power structures; and in some cases villages do not 'officially exist' although they are to be found in the field for all intents and purposes.

Under these circumstances the construction of a sampling frame of households by someone who does not live in the village is a very difficult task. We have used community mapping with cards to solve the problem[15]. By involving the community in the definition of the limits of the village and the identification of households we have been able to construct useful and reliable sampling frames.

Community mapping with cards is not merely an efficient means of generating a list of households. Important numerical data can be obtained from sessions based on mapping with cards. However, from a statistical point of view, community mapping has particular significance because it can be used to identify characteristics of each household that allow the researcher to:

---

[12] Over the last decade, community mapping has been used to construct sampling frames for surveys by some researchers. For a discussion, see Marsland (2001).

[13] National statistical offices often use census enumeration areas to define the areas and official definitions for the identification of households.

[14] As the situation becomes more difficult, poorer households tend to merge in an effort to pool resources. Another strategy is to send children to live with relatives.

[15] For a full description of community mapping with cards, see Section 5.4.

a) stratify households into different groups of interest to the study, e.g. wealth groups, food security status categories or farming practice groups;
b) find out the size of each stratum within the community and some basic quantitative information about it using participatory techniques;
c) conduct a survey (selecting households at random from among the cards, either for the village as a whole or for strata within it).

In the case of SP2 Module 5, in addition to the participatory sessions, the research required a full enumeration of each village (see Box 3.1).

Box 3.1: Measuring the size of the rural population in Malawi

The SP2 Module 5 study (Wingfield Digby, 2000) defined the study unit as the household because registration for Starter Pack was done on the basis of one pack per farming household. One of the main objectives of the study was to find out whether the Starter Pack register was being inflated by households registering more than once and by the registration of non-existent households. Some stakeholders argued that Starter Pack was providing incentives for the proliferation of households (because if you wanted to receive a pack you needed to be a household).

The selection of villages for the study was carried out using a probability-based sampling scheme[16]. At village level we relied on a community mapping exercise to locate every household in the village. We believe that no other tool would have been able to do this, as participation was necessary to establish the boundaries of the village, which households belonged to it and which of them were real households. On the basis of the map, a full questionnaire-based enumeration of all households was carried out (see Appendix 1).

3.7 Within the community

We have pointed out that at village level, community mapping with cards can be useful not only for collecting information about households but for stratifying and selecting a random sample. In this section, we address the issue of whether respondents *should* always be selected at random from within the community (given that we have the means to do so). We also look at alternative approaches that can be used when community mapping and random sampling are not possible.

[16] For a full description of the sampling scheme see Annex A in Wingfield Digby (2000).

3.7.1 Voluntary versus random approaches

Participatory tools usually rely on voluntary participation of individuals. Should the participants in the discussions be selected at random? From a participation point of view, random selection of participants makes little or no sense. From a statistical point of view, random selection of participants may not always be necessary for the generation of useful numbers. In our work, we have opted for participation on a voluntary basis within the community. However, we believe that it is important to understand the implications of adopting either a voluntary or a random approach.

In a household survey, the selection of households – the study unit – is based on probability sampling. It is a common approach in surveys that whatever the restrictions imposed in the selection of units at the top of the hierarchy, the selection of the study units at the bottom of the hierarchy is at random. However, once the enumerator reaches the household, the survey protocol indicates who in the household should be interviewed. The selection of the respondent within the household is not at random. It would be silly to choose a 3 year-old child to ask him/her questions about health issues and provision of services. The choice of respondent is made on the basis of who in the household can provide complete and reliable information about the study unit. It is not uncommon to find that a questionnaire has modules that apply to different members of the household and these members are clearly defined in the protocol.

A parallel can be drawn for research using participatory methodology. Sampling within the community is unnecessary for generating information (numbers or otherwise) when the study unit is defined as the community itself. If the information belongs to the community and can be provided in an unbiased way by key informants or groups of key informants, there is no reason to select those informants at random. It would be silly to do so. PRA tools are often used to discuss issues that deal with characteristics or properties of the communities rather than characteristics of individuals. A resource map, an institutional diagram or a transect walk collects and discusses information that 'belongs to' or 'is a property of' the community. In these cases, as long as the participants are believed to know about the issue under discussion, the fact that they are volunteers does not affect the validity of the information.

The main implications of having communities as study units are that the study will require probability-based selection of *communities*, and that a reasonable number of communities should be visited (see Section 4) if the intention is to draw general conclusions.

However, there are cases when unrestricted participation in the within-community discussions is not desirable and cases when random selection of households or individuals would be appropriate. We believe that in these cases, it is still possible to reconcile statistical sampling with the principles of participation.

We have already discussed stratification as a result of community mapping (see Section 3.6). An easy extension of the identification of strata is that of working with a specific stratum in the analysis of specific topics, in other words, engaging in discussions through participatory methods with specific groups within the communities. Managing participation is common practice in PRA. PRA practitioners often separate women from men, young people from older people, or divide participants either in such a way that one group is not affected by the participation of another group, or with the purpose of empowering particular groups (Chambers, 1994b). The same practice can be applied to research work that uses participatory tools, with the groups being formed to suit the research objectives[17].

Once the community has been split into strata, in-depth studies may be desirable among the members of the strata[18]. In such cases, the selection of the study units (individuals, household enterprises, etc) can be done on a random basis within the stratum using the cards generated by the community mapping (see Section 5.4).

If any of the selected individuals or households refuses to participate there is no loss of information because a replacement can be obtained through the same selection process. An advantage is that in such cases of 'non-response' there is the potential to gain further information: the reasons for refusal may offer interesting insights. If units are selected through voluntary inclusion, these opportunities may be lost.

3.7.2 When community mapping is not applicable

There are some situations in which community mapping may not be possible and alternative solutions need to be found. For example:

- Settings where key informants do not know every member of the community in enough detail to be able to provide reliable information. This might happen in

[17] There may be ethical issues here. Is it acceptable to manage participation to suit research objectives? How different is it from the accepted practice of enabling specific groups (usually those without power) to communicate their points of view? Could this practice it be seen as ethical under the assumption that their views will be potentially useful in influencing policy? (See Section 9).

[18] It is sensible to ask key informants, such as participants in a community mapping session, to provide basic information about all households – e.g. sex of household head, or whether the household received free inputs – but it would not be sensible to rely on them for more detailed information about at household level.

very large rural communities and urban environments or if the information is relatively private in nature.

- Populations that are difficult to identify or to access: e.g. drug users, people basing their livelihood strategies on activities that are regarded as illegal, or people engaged in activities that carry a social stigma.

A solution for the first case would be to work with part of the community. The challenge lies in selecting a sub-set of the community that is representative of the whole. Wilson (2002) suggests that Ranked Set Sampling may offer a solution.

If populations are not easy to identify or are scattered and intertwined with other populations, the development of a sampling frame for selecting a representative sub-set of the community may not be possible at all. For these situations, Wilson (2002) suggests a sampling method based on a protocol.

3.7.3 When sampling is not possible

There are some situations in which the researcher may not feel that it is desirable or possible to use probability-based sample selection at community or sub-group level. In the remainder of this section, we discuss one example from Malawi where the group members were voluntary participants.

The TIP1 Module 3 study (Dzimadzi, 2000) set out to discover how different sex and age groups perceived the printed messages that were included in the packs. The study focused on the perceptions of boys and girls, men and women. The methodology established that the leaflets were to be discussed in focus groups. The discussions were facilitated by one member of the study team and recorded by another member of the team, who assumed the role of a note-taker.

The focus group discussions were carried out with voluntary participants. The study unit was not the community or households within it but segments of communities that could be described as target audiences. The problem lies in the fact that there was no guarantee that the focus group comprised a representative sample of the members of the target audience. We could have been 'unlucky' if we had disproportionately included participants that were well-informed and literate. We could not claim that the focus groups were formed through an objective procedure that offered insurance against the accusation that the participants were not representative of the target audience.

Our answer to the problem was to collect key information about the focus group participants. We recorded the name, sex, age, educational level, religion and reading ability of each participant. After the field work was finished, we looked at the data collected on all participants and made a judgement about whether or not their key characteristics showed patterns that made them different from the expected target

audience[19]. We concluded that the composition of the focus groups did not show any peculiarities that made us suspicious about the external validity[20] of the results.

However, the results of these focus group discussions do not carry the same insurance against the accusation of not being representative as they would if an objective method of selection of focus group participants had been used. The conclusions of the part of the study based on these focus group discussions remain open to criticism on the grounds that there may be characteristics of the target audiences that are different from those of our focus groups.

4 Sample size

What sample size is required to ensure that the study results can be relied upon to make generalisations? This is a difficult question to answer because the reply depends on a number of complex factors. The elements to be considered when deciding on sample size are the focus of this section.

Before starting the discussion, we feel that there is a question that requires a straight answer: "Can statistics provide a formula that will give the correct sample size for a study that uses participatory methods to collect information?" The short answer is that statistics does not have such a formula. The reason for this is that statistical theory deals mainly with making inferences about the numerical characteristics of a population. If a sample size calculation is to be made, it is done for one characteristic at a time. This is not helpful for our problem of deciding how big a sample is needed when we are interested in researching the type of complex issues usually addressed through participatory methods.

4.1 Precision and variability

Although statistics does not provide a formula, understanding the principles underlying sample size calculation can be useful for our objectives. At a basic level, there are two statistical concepts that affect the decision on sample size for the estimation of a numerical characteristic of a population:

1) how *precise* do we need to be in estimating that characteristic; and
2) how *variable* is the population with respect to the characteristic of interest[21].

[19] This is possible by comparing the data with information from other sources, e.g. the 1998 census and our surveys within the TIP1 evaluation.

[20] i.e. that the results could be applied to the target audience.

The following hypothetical example explores these issues using an intuitive approach.

In Figure 4.1 a dot has been drawn for every one of 100 villages in District A. The dots represent the percentage of households that run out of maize six months after the harvest in each village[22]. For example, there is one village where 30% of households run out of maize, in three villages 41% run out of maize and one village shows 70% without maize. On average 49.7% of the households run out of maize.

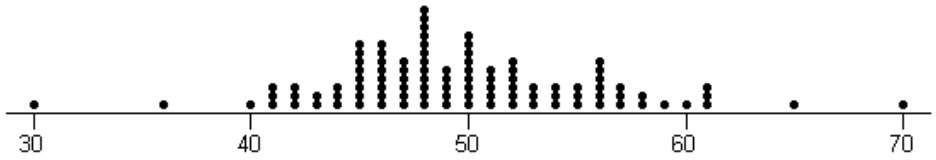


Figure 4.1: District A

If a team of researchers were to visit a random sample of three villages (a sample of size 3) to find out through participatory methods how many households had run out of maize, they might find that the percentages were 30%, 52% and 48%. The percentage for the district would be calculated to be 43.3%. However, we are in the (unrealistic) privileged position of knowing what the true percentage is 49.7%. We can tell that they got it wrong by 6.4 percentage points! Was this an unlucky sample of villages? On paper it is easy to check what would happen with another random sample of three villages: 52%, 49% and 56%, with an estimated district percentage of 52.3%. They got it wrong again! But this time only by 2.6 percentage points.

To take this exercise a bit further, we simulated what would happen if the study were to be repeated ten times[23]. The resulting estimates for the district were in ascending order: 43.3%, 45.3%, 45.7%, 46.0%, 46.7%, 48.3%, 48.3%, 52.3%, 53.0% and 53.0%. Not once did we get it exactly right, and the averages varied within a range of nearly 10 percentage points.

To demonstrate the effect of increasing the sample size, we repeated the same exercise but this time with a sample of size 10. In other words, our team went to ten villages each time, instead of only three. The resulting estimates for the district, in

---

[21] The overall variability can be split into variability that can be explained by known factors (e.g. geography, proximity to a trading centre, etc) and random variability. It is the random, unexplained variability that is important in the calculation of sample size.

[22] This data does not reflect the real situation in any real district in any country.

[23] This is not an approach that would be sensible in practice and by no means a methodology that is to be recommended.

ascending order, were: 46.6%, 47.1%, 49.1%, 49.4%, 49.9%, 50%, 50.3%, 51.4%, 52.3% and 53.6%. The range of the estimates is 7 percentage points.

We tried the same exercise with 20 villages (sample size 20) and found that the range of the estimates was reduced to 5 percentage points. Again, we did not hit the true district percentage with any of the samples taken, but we got closer to the true percentage for the district.

The effect of increasing the sample size was to reduce the width of range in which the percentage estimated by the sample was likely to fall. In other words, we increased the precision of the estimate[24]. It is important now to come back to reality and stress that in practice only one sample would be taken. We only have one opportunity to get it right, so we would want to increase the chances that the estimate will be close to the real value for the district. The choice of sample size will depend on the level of precision required.

Suppose now that the percentage of households without maize in villages in District B looks like this:

Figure 4.2: District B

In Figure 4.2, the percentages are not concentrated around the 50% mark but spread between 30% and 70%. The district percentage is still 49.7%.

When we took ten samples of different sizes the results were:

Sample size	% households as estimated by each sample	Range of estimates in percentage points
3	36.7, 54.7, 48.7, 48.7, 57.3, 47.0, 43.7, 49.7, 40.7, 42.0	21
10	46.6, 46.6, 50.4, 50.8, 48.3, 56.5, 54.9, 51.5, 51.2, 53.3	10
20	50.4, 53.4, 49.1, 46.5, 47.4, 49.4, 47.5, 51.4, 52.5, 50.6	7

District A and District B are similar insofar as they both have 100 villages and the percentage of households running out of maize is 49.7%. The difference is that in District A the village percentages are less variable than in District B. This example

[24] Strictly speaking, the precision of an estimate is measured by its standard error. We use the range of the estimates in a number of samples as an illustration only.

illustrates how the variability in the population affects in the precision of the estimate. For example, the range that we found in District A with ten samples of size 3 was 10 percentage points while in District B it was 21 percentage points.

In summary, the more variable the characteristic of interest, the larger the sample that is needed to achieve the required level of precision.

What if we want to be confident that 95% of the possible samples yield an estimate that is not more 3.5 percentage points away from 49.7 (the district true percentage) – i.e. between 46.2 and 53.2. How large a sample would be needed? We simulated the study 10,000 times for District A assuming that each time ten villages were visited. The result was that 95% of the estimates lay between 46.1 and 53.5. For District B, we had to increase the number of villages to 40 in order to get 95% of the estimates between 46.1 and 53.4.

For policy makers in District B, the average village percentage alone is of little use. They would be interested in knowing that there is high variability, and where that variability comes from. For this purpose we need a reasonably large sample size.

4.2 Population size is irrelevant

A common misunderstanding when choosing a sample size is that one needs to take into account the size of the population, i.e. the study of a larger population requires a larger sample size. This is not correct. In the examples above, the number of villages in the district could have been 100, 1000 or 10,000 and, as long as the percentage of households running out of maize showed the patterns of variability described, the estimates from samples of size 3, 5 or 10 would have shown the same behaviour. From an estimation point of view, the choice of sample size is not made as a fixed proportion of the population. Rather, it is guided by two main criteria: the required level of precision for the estimate and how variable is the population with respect to the characteristic of interest.

4.3 Other factors

Enough of statistical gymnastics! How do we choose a sample size in practice? Apart from the criteria of precision and population variability, there are other factors that affect the choice of sample size. Among these are the structure of the units of study, the level of aggregation at which the results are to be presented and the resources available for collecting the information. The final decision should be based on a trade-off amongst all these elements.

The decision about the sample size is never easy. One of a series of best practice guidelines commissioned by DFID and published by the Statistical Services Centre at The University of Reading observes that:

> "Making your own sampling decisions based on careful thought and detailed planning is a positive challenge: dialogue with some people experienced in sampling issues is often a great help to ensure the proposal covers all the necessary issues in a sensible way, and we would advocate that at least one of them should be a sympathetic statistician. Even if experience shows you that your solution is less than perfect, you can take comfort from knowing the alternative is worse: adopting pre-packaged solutions and standard plans without thought would probably be a recipe only for disaster!" (SSC, 2001).

4.3.1 Structure of the units of study

In relation to the structure of the units of study, we have already pointed out that in research using participatory methods it is common to have multiple units of study. For example, we might want information from the village level as well as information from household and individual level. This means that there is a need to answer three questions: how many villages? how many households? and how many individuals? Wilson (2001) presents a useful discussion about study design compromises. He argues that although visiting a very small number of villages and a large number of households in each village (in order to obtain a large sample of households) might seem an attractive way to reduce costs, the potential for generalisations would be "extremely weak […] insofar as we have almost no primary information about between-community variability".

A potential problem with having few units at the higher levels of a hierarchy is that what happens in units in lower levels might be affected by what happens in units at higher levels. For example, in the evaluation of the free inputs programmes in Malawi it was found that whether a household received a pack or not was highly influenced by what happened at the village level. In some cases, whole villages were left out or refused to receive packs. If a study decided to visit only a small number of villages and to collect information from a large number of households within each village, there would be the risk that an 'unlucky' sample might be drawn – one that selected a village where no one received a pack. The effect on the estimation of the proportion of households that received the inputs, or on the level of fairness of the programme as perceived by the smallholder farmers, would be disproportionate. On the other hand, with a large sample there would be two advantages: a) if the problem was not detected, the effect would not be as strong; and b) there is the opportunity to detect the problem and get an idea of how widespread it is[25]. The problem is that unless we

[25] In a survey of 135 villages for TIP2, we found that four villages did not receive any inputs and that in four villages every household received a pack. From these results, one could infer that such anomalies happened in around 6% of villages.

have a large enough sample it is difficult to know whether our sample has been 'unlucky'. If the researchers suspect that this type of problem is likely to occur, the design of the study needs to consider visiting a large number of villages.

4.3.2 Level of aggregation

The level of aggregation of the final results is another important consideration. For instance, in Malawi we ask the question "Do we want to present the results at national, regional or district level?". Let us take as an example a case where the first phase of a project is working in only one district within a region. The implementing body sets up an evaluation study with sample size X that yields information that can be regarded as a valid and reliable assessment of the impact of the project work (i.e. the results of the evaluation can be generalised to the work of the project). In the second phase, the project will expand to the whole region and this means adding four new districts to the project area. An evaluation is to be carried out to assess the impact of the second phase of the project. If the results are to be presented at regional level only (no breakdown by district is required), the sample size needed is likely to be similar to the one used in the first phase of the project provided that there are no major differences between districts and the information requirements have not changed. However, if the information needs to be broken down by districts, and the implementing body wants as high a level of precision at district level as the one obtained in the evaluation of the first phase, the sample size required is likely to be several times bigger.

When thinking about sample sizes it is useful to have a clear idea of how the results are to be presented and who is going to use them. For example one study might want to give the percentage of households that are food insecure in Malawi: this will probably be suitable for stakeholders working at an Africa level. The Government of Malawi might be interested in the results at regional level (North, Centre and South). The district authorities might need the information for their own district broken down by sex of head of household. The required breakdown of the results is important for the decisions about sample size.

4.3.3 Availability of resources

In the case of the Malawi studies using participatory methods, we worked with no less than 20 villages throughout the country and up to 48 in one case. Twenty villages was the minimum number that we were prepared to work with. We would have felt happier working with 30-40 villages in all cases, but in practice resource constraints

did not always make this possible (see Section 6.5 on cost considerations). A sample of 40 villages would have given us a sample of some 2,500 households[26].

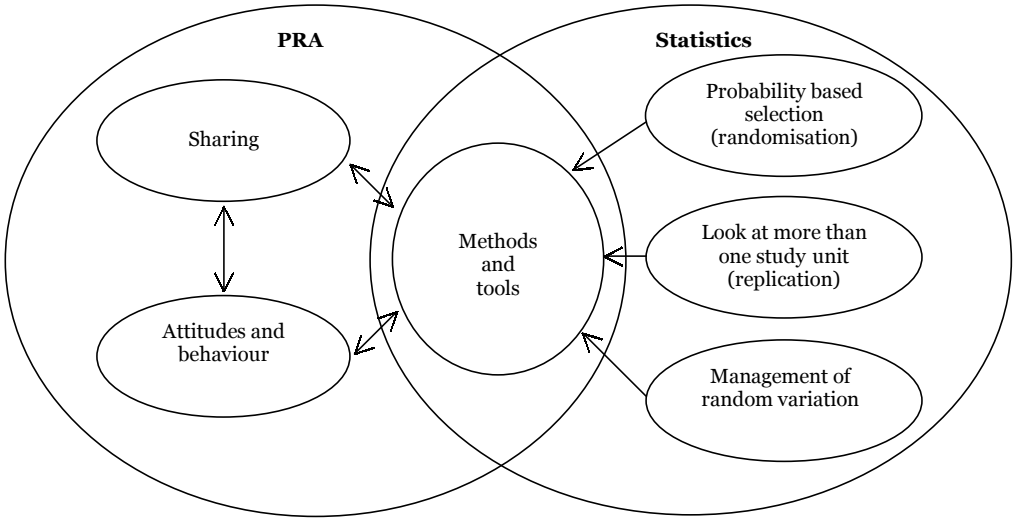
5 Adapting PRA tools

In previous sections, we have discussed some of the basic statistical principles that can be used when designing research studies using participatory methods. This section deals with adapting PRA tools for the generation of numbers. In particular, we look at the issues of standardisation and comparability.

5.1 Building bridges

Readers familiar with PRA will easily list its three basic components: 1) sharing; 2) attitudes and behaviour; and 3) methods[27]. Readers familiar with statistical methods, particularly those who have taken a course on experimental design, will probably remember that there are three basic principles of design of experiments: 1) randomisation; 2) replication; and 3) management of random variation.

Figure 5.1: Integration of PRA and statistical design of experiments



When looking to integrate the two methods, we would argue that there is no need to adapt the first and second components of PRA or the three basic statistical principles. However, we do need to adapt the PRA tools to satisfy the requirements of collecting information from multiple sites – which is necessary for statistical processing and

[26] This estimate is based on TIP2 Module 2, where 21 villages were visited and information was collected from a sample of 1,343 households.

[27] Mascarenhas (1991) quoted by Chambers (1994b).

generalisation of results. At the practical level, the integration of the two approaches comes when designing the tools (see Figure 5.1).

The full potential for integration is realised when the design of the research takes into account basic statistical principles while the praxis of the research is based on a philosophy of participation. Both design and practice meet through the methods and the tools.

5.2 Standardisation

The main differences between traditional PRA approaches and research using participatory methods (as defined in Section 1.1) are that in the latter case a) we need to work in a fairly large number of sites, b) the researcher/facilitator has a clear purpose of acquiring data for informing policy and c) the impact of participation will go beyond the local setting. As a result, standardisation of PRA tools is needed so that consistent results are produced. This means that across all sites:

1. The tools should be used in the same manner.
2. The same type of information with the same level of detail should be collected.

The first condition implies the development of a Field Manual that specifies in detail the tools to be used and the steps to be followed for each activity throughout the period that the research team interacts with the community in each site. The Field Manual is developed by the researchers in a process that involves discussion and appropriation of research objectives, preliminary testing and improvement of tools. In addition, we standardise the 'human element' as much as possible by careful training of field teams (see Section 6.3). It should be emphasised that in order to standardise the tools, a preliminary design phase is required. It is not possible to develop standardised PRA tools sitting in an office. Consultation with stakeholders at community level as well as testing of different approaches are essential for the development of the PRA tools (see Section 6.2).

The second condition is fulfilled by the use of a unique instrument for recording the information acquired in each site. We call it a Debriefing Document. This complements the Field Manual by providing the framework for recording the information with the same level of detail across all sites.

We should stress that although our emphasis is on the information that can be integrated across sites for the purpose of making generalisations, there is a wealth of context-bound information that comes from the use of PRA tools, of benefit for local stakeholders as well as the researchers. This plays an important part in the processes of learning and sharing. We believe that the ideas presented here 'add to' rather than

'take away from' the demonstrated potential of participatory methods. We assume that researchers interested in participatory methods practice principles of participation such as reversal of learning, offsetting biases, optimising trade-offs, triangulating, seeking diversity, "handing over the stick", self-critical awareness, personal responsibility and sharing (Chambers, 1994a).

In terms of recording information, particularly non-standardised information, the researcher's skill in using a notebook remains an irreplaceable asset, whether in the context of research using participatory methods or other types of research.

Although we are making a case for standardisation, it is important to note that we are not proposing to remove the flexibility of PRA tools. Standardisation should apply to the process of collecting and recording key pieces of information that are needed to fulfil the research objectives. When adapting the PRA tools, the challenge lies in identifying those aspects that require researcher control while retaining as much of the flexibility of the methods as possible.

5.3 Comparability

In this section, we examine how two PRA tools have been adapted to ensure comparability of information collected between sites. Ensuring comparability means using the same time periods, definitions, indicators, etc in all sites, and using absolute rather than locally referenced, relative measurements. This frequently means alterations to the methodology, but it is essential if we are to aggregate the results.

5.3.1 Trend analysis

In the SP2 Module 4 study about the impact of free inputs on sustainable agriculture in Malawi (Cromwell, 2000), we were interested in examining trends for a set of fifteen indicators of sustainability[28]. For this purpose we decided to use PRA trend analysis, as it would allow the research teams to identify the main changes and to get a deeper understanding of the reasons for them.

In such cases, it is important first to identify the structure of the information. The researchers should ask themselves: What do we want to know? What do we know already? How do we take advantage of what we know to make our research more efficient? The answers to these questions are important for the design of tools.

[28] See report by Cromwell (2000) for a full discussion about how the indicators were decided upon through consultations with communities in the preliminary phase of the study.

What did we want to know? The first priority was the shape of the trend for each sustainability indicator, and whether the farmers in the discussion groups felt that the indicator was of importance for them. In ordinary trend analysis, the exercise would start with reaching an agreement with the group about the period of time to be considered. However, this was an element that we could not afford to leave to the discretion of each group. We realised that for the information to be comparable across sites, the period of time discussed by the groups needed to be the same. What did we already know? We knew (or were willing to assume) that the factors affecting the sustainability of the smallholder agricultural enterprise in Malawi started having a major impact in the 1970s. In consequence, as part of the Field Manual we established that the period of time under discussion should start in 1970 and go up to 2000 with time references for the 1970s, 1980s and 1990s.

5.3.2 Scoring versus ranking

Another example of how a PRA tool has been adapted in our work is the use of scores for preferences. The TIP1 Module 4 study (Chinsinga, 2001) facilitated 'voting' for different types of direct welfare transfer by scoring of the alternatives: cash and in-kind transfers. The objective was to gather evidence that would allow the researchers to answer the question whether it would be best to give cash or in-kind transfers to the work-constrained poor. The process is described in Box 1.1 (Section 1.4).

In order to make the results of the scoring comparable across sites we decided to pre-determine the scoring scale. A scale from zero to five was selected, where a low score represented low preference[29]. The scoring was done individually and in secret after a full discussion of the alternatives. When the 'voting' took place, the score given by one person should not have influenced the score given by another. Moreover, each alternative was scored in turn, independently of the other.

The total score was the number of grains of maize given by all individuals in the group to each alternative. In addition to the total score, the total number of people that scored the option was recorded. At the analysis stage, we used the fraction (total score)/(total number of participants × 5) as the indicator of preference.

It is arguable that preferences for different alternatives could be addressed using a ranking process. However, throughout the studies carried out in Malawi we have preferred scoring to ranking. The main reason is that ranks give information that is relative to the set of items under consideration. Also, ranking gives no indication

[29] Local materials were used for scoring, generally grains of maize. The six-point scale (0, 1, 2, 3, 4 and 5) was adopted because it does not encourage people to go for the centre of the scale, does not require highly levels of numeracy and the value of individual scores is intuitive.

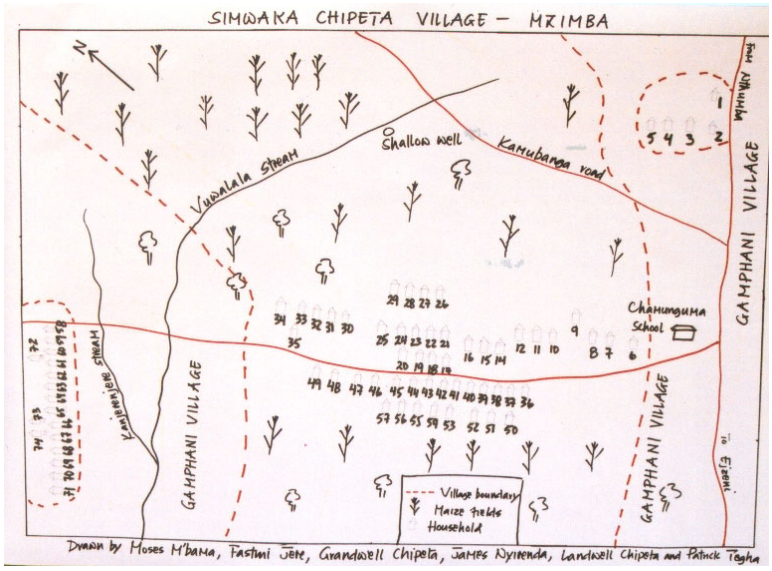
about the size of the gap between different elements in the list. For example, the in-kind transfers might be preferred to cash, so in-kind transfers would have rank 1 and cash would have rank 2. However, this does not tell us much about the strength of the preferences. If a third alternative were available in some areas, the comparison of ranks from sets of two options with ranks from sets of three options would not be possible.

The difference between scores and ranks is that scores are *absolute* while ranks are *relative*[30], and therefore not useful for integration across sites. If scores are obtained by a process of independently scoring each alternative, they should be good raw material for integration across sites. They also have the advantage of allowing the integration of results even if the sets of alternatives are not exactly the same in all sites. For a fuller discussion of scores and ranks see Abeyasekera (2002).

5.4 Community mapping with cards

A highly efficient tool for generating numerical data as well as facilitating discussion of complex issues is the community map combined with household cards[31]. The following section looks in detail at this technique, as we believe that it has great flexibility and potential to generate numbers from participation.

Picture 1: A community map (TIP2 Module 2)



[30] We refer to ranks as being 'relative' although in a mathematical sense they are better described as 'conditional'.

[31] In our Malawi studies, this is usually referred to as 'social mapping' and 'card games'. In this paper, we refer to 'community mapping with cards' because the main product is not an analysis of social features, although it is possible to combine the approach with a social emphasis.

Community mapping with cards evolved from social mapping. As with social mapping, it begins with asking participants to map the community, including key geographical, social and institutional features. However, an additional activity is required that is not part of social mapping processes in general. This is the marking on the map of every household in the community, identified by a number (see Picture 1). As each household is marked, a numbered card is produced for the household and certain characteristics of the household are recorded on the card[32].

Community mapping with cards can generate information about percentages of households belonging to different categories – well-being or poverty groups, food security groups, farming practice groups etc. It can also be used to produce data that allows us to analyse the correlation between such categories and other indicators, or on the level of agreement between different stakeholders[33]. In Malawi, three of our six research studies using participatory methods employed this tool:

- Module 4 of the SP2 evaluation mapped every household in the village using cards, and then divided the households into 'high', 'medium' and 'low' sustainability farming practice groups for focus group discussions.
- Module 4 of the TIP1 evaluation used community mapping with cards to identify all the households in the village and place them in well-being categories (including the poor and vulnerable). Focus groups then used the cards to simulate the selection of households for different benefits – direct welfare transfers and TIP – and to identify the criteria they used in the selection process.
- Module 2 of the TIP2 evaluation used community mapping with cards to categorise households as 'food secure', 'food insecure' or 'extremely food insecure', and to collect basic information about the household including gender of household head, receipt of TIP and whether – in the views of different stakeholders (the Village Task Force, the TIP beneficiaries and the non-beneficiaries) – the household met any criteria for selection as a TIP beneficiary.

In the remainder of this section, we take the example of Module 2 of the TIP2 evaluation to show how the process works and what sort of numerical information can be generated from it. Our emphasis is on generating data, but it should be pointed out that this is not a mechanical process of producing numbers. The data comes from discussions with key informants or members of groups of interest to the researcher within the community. This participation element gives credibility to the information collected, as it should have been checked by various stakeholders in the process of collection. It also provides an opportunity for the facilitators to record the views and behaviour of the participants. This information is vital for interpreting the numbers.

---

[32] Similar tools have been developed in relation to censuses in India and the Philippines (Chambers, 1994a).

[33] It is possible to explore such relationships because each household card contains several pieces of information and therefore households can be classified by more than one factor at a time.

5.4.1 An example of community mapping with cards

5.4.1.1 Background and objectives

In the 2001-02 TIP, there were enough packs available for roughly one-third of households in rural areas, and these were intended to go to the poorest households through a process of community targeting in which Village Task Forces (VTFs) were to select the most deserving beneficiaries to receive a pack. The Module 2 evaluation team's main challenges were to find out whether the community targeting process had been fair, and whether one-third was an appropriate proportion of rural households to target or whether a larger proportion qualified as eligible beneficiaries. The team's specific objectives were to establish:

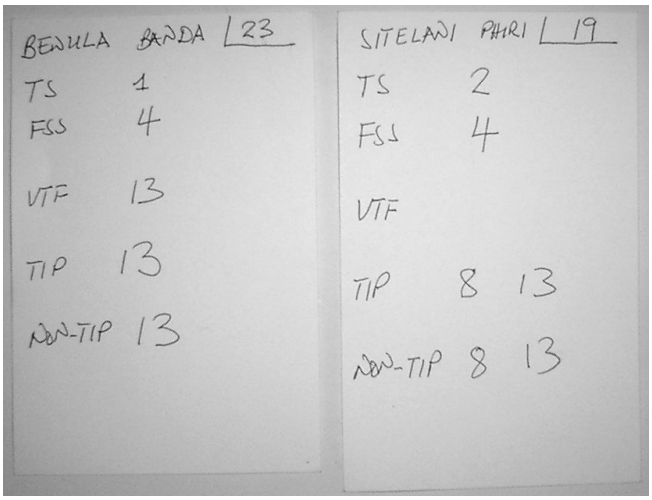
1. what proportion of the population was food insecure, moderately food insecure and extremely food insecure (food insecurity was used as a proxy for poverty);
2. whether food insecure households were given priority in the allocation of packs;
3. the relationship between receipt of free inputs and gender of household head;
4. the relationship between receipt of free inputs and household characteristics that met the official criteria for receiving the inputs (e.g. whether the household head was elderly or a widow/widower, and whether the household kept orphans);
5. what proportion of those who benefited from TIP in 2001-02 should not have received a pack, in the view of local stakeholders; and
6. how much agreement existed between different stakeholders – the VTF, the TIP beneficiaries and the non-beneficiaries – about who should receive a TIP pack and who could safely be excluded from the programme.

5.4.1.2 How the information was collected

a) The mapping exercise

The mapping exercise was carried out by 5-8 people from the village with enough knowledge of their community to produce a map. The participants were asked to show on the map the location of every household in the village and to prepare a card for each household, with the name and sex of the household head and a household number (Chinsinga, 2002). It was vital that every household in the village appeared on the map and that a card was prepared for each household with the same number that appeared on the map (see Pictures 1 and 2). The household number was written on the top right-hand corner of the card.

Picture 2: Household cards (TIP2 Module 2)



During the mapping exercise, the participants were asked to "categorise households into three predefined food security categories: Food Secure (FS), Food Insecure (FI) and Extremely Food Insecure (EFI). The decision to use predefined categories was in order to ensure comparability between sites", Chinsinga (2002)[34]. Each household card was marked with a food security status (FSS) code. Participants were then asked to say which households received a TIP pack, and each household card was marked with the TIP status (TS) code for receipt or non-receipt of TIP. The codes in Table 5.1 were used, with the aim of recording the information in a manner that would not be obvious to participants in the next stage of the process (the focus group discussions). Thus, the participants in the next stage of the process were 'blind' to the information collected in the previous stage (see Section 9.2).

Table 5.1: TIP and food security status codes

TS Code	Description	FSS Code	Description
1	Household received a TIP pack	3	Food secure
2	Household did not receive a TIP pack	4	Food insecure
		5	Extremely food insecure

Source: Chinsinga (2002).

The information produced during the mapping addressed points 1-3 of the study's specific objectives. Section 5.4.1.3 shows how the data was presented.

b) Focus group discussions on beneficiary selection and community targeting

[34] For further information about definitions and other aspects of the exercise, see Chinsinga (2002) Appendix 3: Field Manual.

Points 4, 5 and 6 of the specific objectives required a discussion of each household card to determine whether or not it should qualify as a TIP beneficiary. In order to achieve this, the study team organised three focus group discussions: one with the VTF or the Village Head (VH) if there was no VTF; one with a group of TIP beneficiaries; and one with a group of non-beneficiaries (see Box 5.1). The objective of the game was that all three groups of stakeholders should decide independently whether each household in the village met any criteria for inclusion in TIP; the researchers would then write the code for the relevant criteria (one or more might be mentioned) on the household card against the relevant focus group – VTF, TIP and non-TIP. A list of 'core codes' was provided for criteria that were likely to apply (see Table 5.2); if the focus group mentioned a different criterion, this was given a new code and recorded in the Debriefing Document.

Box 5.1: Card game procedure – from Field Manual for TIP2 Module 2

For VTF focus groups or VHs:

"The game will use the cards generated from the Social Mapping exercise. For this exercise select the household cards in numerical order. Go through each card one by one. Explain to the VTF/VH that you would like to know whether the households on the cards fulfil any of the criteria for selection. Don't mention whether they received TIP or not. The other team member should record the **code** representing the reason for **inclusion** on the household cards; write code 10 if the household does not meet any of the criteria. If the focus group participants refuse to play the game with any of the households, put "-" (a dash) in the appropriate place on the household card(s)".

For beneficiary and non-beneficiary focus groups (separately):

"Explain criteria for inclusion ***actually used by the VTF/VH***. Present all the household cards from the Social Mapping one by one. Ask the participants:

- Does the household meets any of the criteria established by the VTF? If so, which criteria?
- Should the household receive TIP and if so, by what criteria? *(If it does NOT qualify for inclusion by any criteria, write 10).*

Write down the **code** representing the reason for **inclusion** on the household cards; write code 10 if the household does not meet any of the criteria. If the focus group participants refuse to play the game with any of the households, put "-" (a dash) in the appropriate place on the household card(s)".

Chinsinga (2002), Appendix 3.

Table 5.2: Core codes – criteria to qualify for TIP

Code	Description
6	Elderly
7	Household keeping orphans
8	Widow/widower
9	Household keeping disabled
10	Does not satisfy any criteria to receive TIP
11	No support or income
12	Divorced
13	Without seeds
14	Without fertiliser
15	Sickly household head
16	House of a junior wife in polygamous household
17	Energetic/able-bodied OR has help from able-bodied person

Source: Chinsinga (2002).

The study found that the stakeholders involved in the three different focus groups were able to decide whether households fulfilled any criteria to receive TIP; they were also able to identify households that did not qualify by any criteria to receive TIP. We found that there was substantial agreement between stakeholders on who deserved to be included and who might be excluded from the programme.

The cards allowed us to keep information at household level which in turn allowed us to match this information with information about receipt of TIP. For example, we might have a household that was agreed to qualify by all three focus groups because the household head was elderly, and our information might indicate that the household received a TIP pack. We might have another household that did not receive a pack; this household might have been recorded as qualifying (code 6) by two focus groups, but as not fulfilling any criteria (code 10) by the third group.

5.4.1.3 How the information was presented

a) The mapping exercise

The information produced during the mapping was used to produce the following tables. These could be produced at village, regional or national level.

Table 5.3: Relationship between receipt of TIP and food security status

Food security status	TIP recipients (%)	Non-recipients (%)	Total households (%)
FS	21.2	33.5	28.9
FI	38.5	39.7	39.3
EFI	40.3	26.8	31.8
Total	100.0	100.0	100.0

Source: Adapted from Chinsinga (2002).

Table 5.3 shows the relationship between receipt of TIP and food security status at national level. If targeting had been fair, all EFI households should have received a pack; FS households should not have received one. This was not the case, although the proportion of EFI TIP recipients was slightly higher than for non-recipients. A formal statistical test could be used to check if the distribution of inputs really matched food security status. However, in our view formal tests are unnecessary because this exploratory analysis presented enough evidence to show that TIP distribution did not effectively favour the food insecure.

Table 5.4: Relationship between receipt of TIP and sex of household head

Sex	TIP recipients (%)	Non-recipients (%)	Total households (%)
Female	41.6	32.7	36.0
Male	58.4	67.3	64.0
Total	100.0	100.0	100.0

Source: Adapted from Chinsinga (2002).

Table 5.4 shows the relationship between receipt of TIP and sex of household head at national level. It shows that there was a slight preference for female-headed households in the selection of TIP beneficiaries.

b) Focus group discussions on beneficiary selection and community targeting

Our data analysis focused on several aspects of beneficiary selection and community targeting:

a) Did the VTFs/VHs who selected the beneficiaries in 2001-02 select only the deserving and exclude only the non-deserving?
b) Is there a group that can be agreed by all local stakeholders to be deserving, and what proportion of rural households does this represent?
c) If communities were to undertake beneficiary selection in future and enough packs were provided for all deserving households, would we expect them to target the deserving households only?

The information from each focus group was analysed as follows. First, the cards of TIP beneficiaries were classified according to all qualifying criteria and Code 10 (did not satisfy any criterion to receive TIP); the same was done for the non-beneficiary cards. This produced six tables of data similar to the example in Table 5.5[35], which shows the national-level results for the TIP beneficiary cards classified by the non-beneficiary focus group. We can see that the TIP beneficiaries qualified to receive TIP for many reasons, but according to the non-recipient focus groups, 18% of cases[36] did

[35] The full set of tables and analysis is presented in Chinsinga (2002), Section 4.5.

[36] We use percentage of cases rather than percentage of responses because this is a multiple response (i.e. one household may qualify for one, two or more reasons).

not meet any criteria and therefore should not have received TIP. A similar analysis of non-beneficiary cards classified by the three focus groups showed that over half of non-beneficiaries did meet criteria to qualify for TIP although they did not get a pack.

Table 5.5: TIP recipients – criteria used by non-recipient focus groups

Category label	Code	Count	Pct of Responses	Pct of Cases
Food insecure	4	1	.2	.2
Extremely food insecure	5	34	6.7	8.5
Elderly	6	114	22.4	28.4
Household keeping orphans	7	43	8.4	10.7
Widow/widower	8	37	7.3	9.2
Household keeping disabled	9	15	2.9	3.7
Does not satisfy any criteria to receive	10	73	14.3	18.2
No support or income	11	37	7.3	9.2
Divorced	12	28	5.5	7.0
Without seeds	13	29	5.7	7.2
Without fertiliser	14	17	3.3	4.2
Sickly household head	15	33	6.5	8.2
House of a junior wife in polygamous HH	16	3	.6	.7
Energetic/able-bodied	17	9	1.8	2.2
People involved in development work	19	1	.2	.2
Those without large gardens	20	1	.2	.2
Households with a large number of depend	21	6	1.2	1.5
Households keeping the elderly	24	1	.2	.2
Those who volunteered to receive TIP	27	18	3.5	4.5
Newly married without resources	28	3	.6	.7
Members of VTF	29	7	1.4	1.7
Total responses		510	100.0	127.2

43 missing cases; 401 valid cases

Source: Chinsinga (2002).

The second part of the analysis of data from the focus group discussions focused on Code 10. First, we looked at Code 10 households as a percentage of total households in the communities visited (as well as percentages of TIP beneficiaries or non-beneficiaries). We found that at national level between 32% of households (according to VTFs) and 36% (according to TIP beneficiaries) did not deserve to receive TIP, i.e. 64-68% did deserve to receive a pack (see Table 5.6). This indicated that the number of packs distributed in 2001-02 was, in the opinion of those doing the targeting and the recipient communities, well below the number of deserving households.

Table 5.6: Code 10 analysis at national level

Focus Group Discussion	Code 10 households		% of total households
	Recipients %	Non-recipients %	
VTF	14.1	41.9	31.6
TIP recipients	16.5	47.3	35.8
Non-recipients	18.2	41.2	33.1
Total	100.0	100.0	100.0

Source: Adapted from Chinsinga (2002).

We then looked at whether it would be possible to target 64-68% of households in Malawi. The study concluded that proportion of households to be targeted would have

to be higher – 80% at national level – because of 'inclusion errors' – the percentage of households included as recipients despite not meeting any criteria probably because they are part of local power structures – and disagreements about specific cases between stakeholders (VTFs, beneficiaries and non-beneficiaries).

A further problem was variations between regions and between sites, which "would give rise to allegations of unfairness at local level even if the overall level of coverage were adjusted to take account of regional variations" (Chinsinga, 2002). Table 5.7 shows the results on a village-by-village basis, indicating the extent of variability.

Table 5.7: Variability in the proportions of deserving TIP recipients across sites

		According to VTF		According to Recipients		According to Non-recipients	
Village	Region	% should receive	% should not receive	% should receive	% should not receive	% should receive	% should not receive
Chilarika II	North	32.4	67.6	36.8	63.2	33.8	66.2
Kafwala	North	51.2	48.8	78.0	22.0	82.9	17.1
M. Chivunga	North	63.3	36.7	45.6	54.4	55.0	45.0
Mdambazuka	North	43.2	56.8	40.5	59.5	40.5	59.5
S. Chipeta	North	51.4	45.5	59.5	40.5	51.4	48.6
Chatowa	Centre	100.0	0.0	90.0	10.0	85.0	15.0
Daudi	Centre	100.0	0.0	100.0	0.0	93.3	6.7
Matapa	Centre	53.5	46.5	55.6	44.4	69.7	30.3
Mdala	Centre	81.0	19.0	77.6	22.4	67.2	32.8
Mkanile	Centre	77.9	22.1	83.8	16.2	82.4	17.6
Mkhomo	Centre	96.0	4.0	100.0	0.0	-	-
Nkhafi	Centre	87.1	12.9	93.5	6.5	95.2	4.8
Chimwaza	South	-	-	51.8	48.2	46.4	53.6
Chintengo	South	84.5	15.5	50.0	50.0	43.1	56.9
Chisalanga	South	-	-	89.8	10.2	95.9	4.1
Makuta	South	78.4	21.6	40.5	59.5	100.0	0.0
Mbepula	South	78.8	21.2	88.2	11.8	82.4	17.6
Mwambeni	South	74.5	25.5	60.1	39.9	57.5	42.5
Njuzi	South	62.7	37.3	41.8	58.2	83.6	16.4
Sitima	South	54.5	45.5	72.7	27.3	-	-

Source: Chinsinga (2002).

5.4.1.4 Conclusion

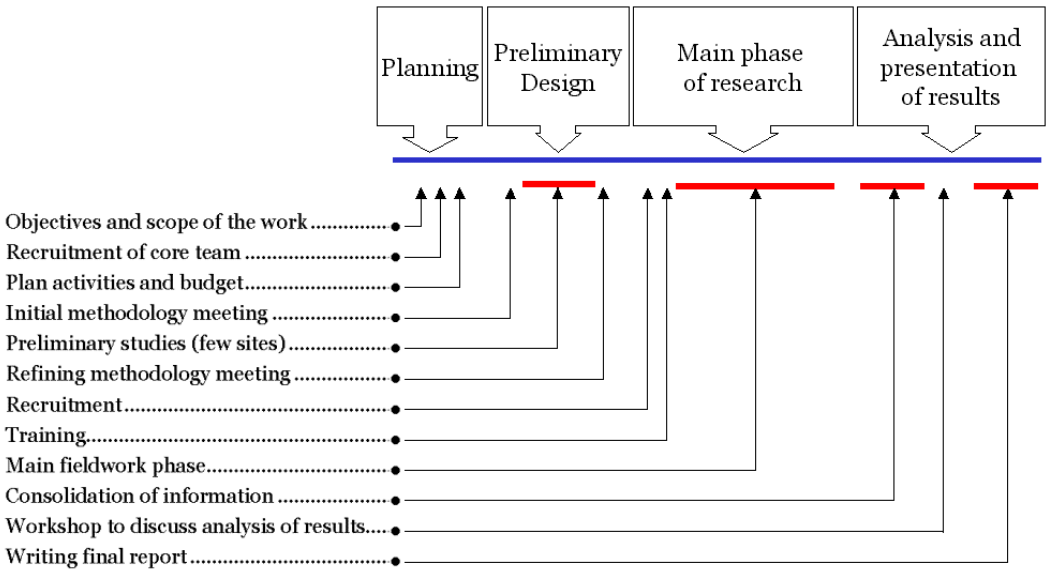
This example of community mapping with cards shows how data can be generated to provide important insights into complex policy issues. The information obtained requires a participatory process of data collection; it would be impossible to collect using a questionnaire.

It should be noted that the main conditions that are necessary for the data to be suitable for statistical analysis are 'standardisation' of tools (see Section 5.2) and adjustments of tools to ensure comparability of data between sites (see Section 5.3). This implies a careful process of planning, preliminary work, training and main phase fieldwork organisation (see Section 6).

6 Planning, organisation and cost

We have argued that in order to produce generalisable results, including numerical data, the researcher needs to work in a relatively large number of sites and adapt the PRA tools to ensure standardisation and comparability across sites. This section looks at the planning, organisation and cost implications of these requirements. We draw on the lessons of several of the Malawi research exercises. Our discussion reflects the consecutive stages of the process, as shown in Figure 6.1. In order to generate reliable results which will be taken seriously by policymakers, it is important that each stage be properly thought out. In particular, we give considerable weight to the preliminary design stage. This is intended to make sure that the approach is tried and tested before it is applied in a standardised manner across a relatively large number of sites.

Figure 6.1: Stages and activities



6.1 Planning

The main activities of the planning stage are:

1. Deciding the objectives and scope of the work;
2. Recruitment of the core research team;
3. Producing an overall plan of activities and budget.

First, it is essential to clearly define the objectives and scope of the research exercise. In particular, those initiating or commissioning the study should:

- agree on what is the population of interest;
- determine the nature of the information to be collected and who should provide this information (see Section 1.4);
- examine the structure of the information separating what the study needs to find out from the information that is already available (see Section 5.3.1);
- consider what research methods will be most appropriate (e.g. a survey, participatory methods, semi-structured interviews with key informants) and, if participatory methods are involved, what sort of tools will be used; and
- establish what are the resource requirements: staff, time and money.

The remainder of Section 6 assumes that the research methods used will be similar to those of our Malawi evaluation studies which used participatory methods.

Once the objectives and scope of the work have been agreed, a core research team should be identified. This team will be responsible for producing an overall plan of activities and budget and for the subsequent stages of the work, including heading the field teams, analysing the information collected in the field and writing the final report. The team members must be good participatory facilitators; they should have good critical analysis skills; they should have field experience; and they must be willing to work in difficult conditions. At least one member of the team should have project management skills, including the ability to co-ordinate a large team and to implement a budget. One or more of the team members must have good writing skills and be capable of pulling together the findings into a final report.

The number of core research team members will depend on the number of sites to be visited and the time available. For instance, if the study plans to visit 40 sites and estimates that the work in each site including travel will take four days (total of $40 \times 4 = 160$ days), then a team of four would be reasonable: each core team member would lead a team that would cover ten sites. The fieldwork would take 40 days: $160/4 = 40$. However, if only one month were available for the main fieldwork, then the core team would need to be made up of five people: $160 / 5 = 32$.

Once the core team has been recruited and has appropriated the objectives and scope of the work, its first task is to draw up an overall plan of activities and budget. These should focus clearly on meeting the study objectives and ensuring that the outputs are delivered within the time and resources available. A Gantt chart is helpful for planning activities[37].

6.2 Preliminary design

The preliminary design stage should comprise an initial methodology meeting, preliminary studies in a small number of sites (three to five[38]) and another meeting – after the preliminary studies – to refine the methodology. This stage should take two to three weeks, and by the time it is finished the methodology for the study – including the PRA tools – will have been completely defined and a Field Manual and Debriefing Document will have been prepared. The Field Manual will contain a programme of activities to be carried out in each site and detailed instructions for each activity, so that the work can be carried out in the same way in each site, while the Debriefing Document will enable researchers to record key findings, including numerical data, in the same way in every site.

The product of the initial methodology meeting – a workshop involving all core team members (and possibly other stakeholders) – should be agreement on what tools will be used to collect each piece of information required for the study objectives. Each proposed tool should be discussed in detail so that the meeting agrees when, how, and with whom they will be used. These guidelines should be written down as a draft Field Manual for use in the preliminary studies in a small number of sites. If there is doubt or disagreement about the best method of collecting the information, alternative methods should be included. The draft Field Manual should include a 'programme of activities' for each site, which recommends the order in which the activities should be tested; however, at this stage, sequencing remains flexible and it may be desirable to test different sequences of activities.

The preliminary studies in a few sites should be carried out soon after the initial methodology meeting by the core team members, working in groups of two or three. The main aim of the preliminary studies is not so much to collect information as to

---

[37] A Gantt chart is a project management tool which shows a breakdown of the activities to be carried out during the project and when each activity will start and finish. Examples and free software can be found by searching the world wide web for "Gantt chart".

[38] The number depends on the diversity of situations that the researchers expect to find in the main fieldwork phase. The sites may reflect different strata or conditions likely to be encountered. They may be selected purposively or at random within strata.

test the proposed tools, so that the researchers can be sure that they will work well – if necessary after modifications – in the main fieldwork phase. Additional objectives are to consult communities about definitions, indicators etc.; to think about how the information from each participatory session will be recorded; and to observe whether the researchers have the right skills for conducting the study and are working well together[39].

The purpose of the meeting for refining the methodology (immediately after the core researchers' return from the field) is to make the necessary modifications. The core team members should discuss their experience in the field – in particular which part of the methodology and sequence of activities worked and which did not – in order to improve the tools and sequencing for the main fieldwork phase. The meeting should also sketch an initial plan of analysis of the information that will be produced in order to be sure that all the information to be collected in the main fieldwork phase will be relevant and nothing will be missing. Once all refinements have been agreed, the team will produce the final Field Manual (including the final programme of activities for each site) and a Debriefing Document. By this point, the approach has been standardised for the main phase of fieldwork.

6.3 Main fieldwork phase

This stage comprises recruitment and training of research assistants and carrying out the main phase of field research in a relatively large number of sites.

Recruitment of research assistants is desirable in order to reduce the amount of time and resources spent on fieldwork. This is particularly important if the study is working in a large number of sites. For our purposes, we found that field teams made up of one core researcher and three research assistants were ideal, as these could be split into two teams of 'facilitator + note taker' if we wanted to run two activities at the same time – e.g. two focus group discussions in parallel.

The key considerations when recruiting research assistants are that they have good facilitation and note-taking skills and enthusiasm for the work. Languages, age and gender may also be important factors – depending on the context that the study will be working in. Once they have been recruited, the research assistants must be fully trained so that they understand the objectives of the study, the use of the Field Manual, facilitation and note taking and how to complete the Debriefing Document.

---

[39] If it becomes apparent during the preliminary design stage that any or all of the core research team are not making a positive contribution to the study, it will be necessary to replace them because in the main phase, if any team produces sub-standard results, the information from that team's sites may have to be abandoned.

Thorough training is essential when teams are to work in a large number of sites because the 'human element' needs to be standardised as much as possible. In other words, the teams should use the tools in the same way in different sites and different teams should use them in the same way. If, for example, a tool depends on a definition such as food insecurity, the definition must be conveyed in the same way to communities in the sites visited by Teams 1, 2, 3 and 4. If Team 3 were to misunderstand the Field Manual and define food insecurity differently from Teams 1, 2 and 4, the information collected by Team 3 would not be comparable with that of the other teams. For this reason, all the core researchers who will lead the teams in the field must be present throughout the training.

Immediately after training, the teams should depart for the field. If the preliminary design and training have been carefully carried out, the main fieldwork phase should be straightforward and enjoyable. In our experience, the main problems that may be encountered are difficulty in accessing a site and the occasional refusal by a community or group to take part in the research. In both cases it is important to document the reasons for these 'non-responses'[40].

6.4 Analysis and presentation of results

The analysis of information from the main fieldwork phase is based on the Debriefing Documents, which carry the key numerical data, drawings, diagrams and summaries of text findings for each site, supplemented by insights from the researchers' notebooks. We find it useful to begin by asking each team leader (the core team member in each field team) to consolidate the qualitative information from his/her sites and present the main findings to other core team members at a workshop. Following the team leaders' presentations, the structure of the final report should be agreed and a plan of analysis for the numerical data discussed. After the workshop, each team leader will continue to consolidate the information from his/her sites, and to refer any queries to the research assistants. The numerical data from the Debriefing Documents will also need to be processed at this point (see Section 7.1).

The core team then needs to consolidate the information about each topic from all sites. This may be done by asking each core team member to take a topic and integrate the findings from his/her sites with the findings from the other teams' sites.

[40] In cases of extreme difficulty of access to a site, we have followed a policy of replacing the site using the same selection method as for the original site. However, replacements should be kept to a minimum, as they introduce bias. Remoteness is not acceptable a reason for replacing a site: if all remote sites were replaced by sites close to roads this would introduce severe bias in our samples.

Once the 'qualitative' findings have been summarised by topic and integrated with the analysis of the numerical data for that topic, the final report can be written. The report should have appendices containing the list of villages visited by the study and their location, the Field Manual and the Debriefing Document. The core researchers may share the writing up of the final report or it may be the responsibility of only one or two of the core team members. In either case, it is vital that all of the core team members read the draft of the report critically to ensure that their findings are reflected in the report, and that it does not give undue weight to the sites visited by the person(s) who have done the writing up.

6.5 Cost considerations

It is often argued that the cost of carrying out participatory studies in a large number of sites is prohibitive. Costs vary enormously depending on the cost structure of the organisation carrying out the study and the conditions in the country where the study is carried out. However, in our experience good planning, organisation and financial management make it possible to carry out research using participatory methods in a relatively large number of sites at a reasonable cost.

6.5.1 Comparison of costs

The following is a comparison of costs between survey work and research using participatory methods[41]. The amounts are given in US dollars. We use hypothetical examples based on the cost of the TIP2 studies carried out in Malawi in 2001-02:

- For the hypothetical study based on participatory methods, the whole study takes four months. It covers 40 villages scattered throughout the country (in a period of 40 days), plus four villages in the preliminary phase. We have four core researchers and twelve research assistants.
- For the hypothetical survey, we cover 3,000 households in 135 villages throughout Malawi. The period of the study is four months, and the main fieldwork phase takes 50 days. The study employs four researchers, 10 supervisors, 19 enumerators and 12 data entry clerks.

Tables 6.1 and 6.2 show that the hypothetical research study in Malawi using participatory methods in 40 sites (around 2,500 households, see Section 4.3.2), costs some $34,000. The hypothetical survey covering 3,000 households in 135 sites costs some $39,000. The cost of the two studies is of a similar order of magnitude.

[41] This example does not include salaries/consultants' fees for the four core researchers in the study based on participatory methods or the four researchers conducting the survey. Nor does it include any administrative overheads. We can assume that both studies cost the same in term of salaries and administrative overheads, as they employ the same number of researchers for the same period of time (four researchers for four months).

Table 6.1: Hypothetical research study using participatory methods

	Cost
Methodology and results analysis workshops	
Room hire & refreshments (3 * 2-day workshops)	$130
Training of research assistants (3 days)	
Attendance allowances (4 researchers, 12 research assistants)	$520
Room hire & refreshments	$80
Fieldwork	
Preliminary design phase	
Researchers' subsistence allowances** (40 person days)	$530
Hire of 2 vehicles for 10 days (including drivers)	$1,920
Fuel	$450
Main fieldwork phase	
Researchers' subsistence allowances* (160 person days)	$2,130
Research assistants' fees/allowances (480 person days)	$11,520
Hire of 4 vehicles for 40 days (including drivers)	$13,060
Fuel	$1,960
Stationery and communications	$1,400
Total cost (US dollars)	**$33,700**

** Field allowances only; salary or consultants' fee not included.

Table 6.2: Hypothetical survey of 3,000 households in 135 sites

	Cost
Methodology and results analysis workshops	
Room hire & refreshments (3 * 2-day workshops)	$130
Training (5 days, including 1 day field exercise)	
Researcher meals allowances (4 researchers)	$80
Supervisor attendance allowances (10 supervisors)	$630
Enumerator attendance allowances (19 enumerators)	$1,050
Room hire & refreshments	$120
Transport, including field exercise	$200
Fieldwork	
Pre-testing questionnaire	
Researcher meals allowances (4 person days)	$16
Hire of 1 vehicle for 1 day (including driver)	$80
Fuel	$10
Main fieldwork phase	
Researcher subsistence allowances** (62 person days)	$830
Supervisor fees/allowances (500 person days)	$6,600
Enumerator fees/allowances (950 person days)	$10,500
Hire of researchers' vehicles for 62 days (including drivers)	$7,660
Fuel for researcher vehicles	$1,650
Supervisors' motorbike running costs, including fuel	$1,360
Public transport refunds for enumerators	$654
Data entry	
Payments to data entry clerks	$730
Transport for data entry clerks	$90
Hire of computers	$2,340
Supervision and management fees	$290
Stationery and communications, incl. photocopying questionnaires	$3,680
Total cost (US dollars)	**$38,700**

** Field allowances only; salary or consultants' fee not included.

6.5.2 Cost control

Financial planning, organisation and management are key factors in any effort to keep down the costs of research using participatory methods:

- Financial planning (budgeting) allows the research team to compare different options for use of resources and to ensure that inputs of time and money are minimised without jeopardising standards. For example, see Box 6.1.
- Good organisation helps to ensure efficient use of the resources available. For example, if transport costs are high, sites should be divided up between the field teams in a way which minimises the amount of travel between sites.
- Financial managers should check that actual expenditures do not exceed the amounts contained in the budget. With research exercises in a large number of sites, spending can get out of hand if it is not strictly controlled[42].

Box 6.1: Person days in the field – a trade-off

The core researchers might feel that they would like to do all the fieldwork themselves, as the quality would be better than if the team were expanded by bringing in research assistants. However, if the core team were to carry out all the fieldwork in 40 sites in which the programme of activities lasted three days with one day allowed for moving to the next site, the fieldwork would last 160 days over five months. If the core team was made up of four highly qualified and highly paid consultants, this would imply paying high rates for a total of 160 × 4 = 640 'person days' in the field. If, on the other hand, each core team member formed a team with three research assistants, the same work could be carried out in 40 days and 480 of the 640 'person days' would be paid at research assistant rates rather than highly qualified consultant rates. The only additional cost would be that of training the research assistants to a high standard.

7 Data analysis

7.1 Data management

An information collection exercise of the size described in this paper places data management demands on the research team. Systems need to be put in place to

[42] Financial incentives and penalties can also be used to enhance the quality of the research work and ensure that outputs are delivered on time.

ensure that the data[43] are managed in such a way that all the information is kept safely, completely and accessibly. This includes hard-copy and electronic versions.

The first and primary deposits of information for our studies are the Debriefing Documents and notebooks. The Debriefing Documents contain concentrated data about the key issues of interest, collected using standardised tools, recorded in a uniform format and completed *in situ*. The researcher's notebooks keep more detailed information in a free format.

In our studies, the Debriefing Documents are gathered together after the field work and kept under the responsibility of the core team manager. Part of the information contained in them may be computerised to help the task of statistical analysis. When data are transferred from paper to computer it is important that the information is complete and that checks are made to ensure that the electronic copy is a faithful transcription of the originals. The choice of software for data management depends on the resources available to the research team. Availability of software and experience in the use of specific packages varies from team to team. We have used simple spreadsheets and complex database systems with custom-made facilities for data entry, filtering and retrieval of data.

Whatever software is chosen a special effort must be made to avoid errors of transcription: much effort and resources will have been invested in collecting the data, and it would be a pity to enter wrong information onto the computer[44].

The importance of periodical backups should also be stressed. We recommend keeping two backups while entering the information: one that is kept in the office and one that is kept away from the office.

7.2 Statistical analysis

The first condition for any dataset to be suitable for statistical analysis is that it can be trusted. Trust comes from the process of data collection and data management. In our experience there is no perfect dataset. Therefore, a data quality assessment helps to establish trust in the information. Such a quality assessment should point out areas were problems have been found and detail the solutions given, identifying areas

---

[43] The Oxford English Dictionary defines datum (singular of data) as "…something known or assumed as fact; and made the basis of reasoning or calculation". We take the definition in its wider sense so as to include not only numerical information but all the information that is collected through the process of research.

[44] Guidelines on how to use spreadsheets for data entry, and general recommendations for data management can be found on www.reading.ac.uk/ssc under the DFID section of the website.

where the user should be cautious while interpreting the information. It should describe any actions taken by the research team to deal with problems such as replacement of study units, refusals or special circumstances found in the field.

We believe that if the design of research using participatory methods takes into account the elements discussed in this paper, the numeric data generated will be suitable for standard statistical analysis. The same considerations apply here as for the analysis of studies that do not use participatory methods, and the same potential for making generalisations exists. There may be an issue of sample size, as for some study units the sample may not be large enough. This is no different from studies using other methods, and it should be known about and dealt with at the design stage.

The type of statistical analysis that is carried out on the data depends on the type of data and the objectives of the study, not on the tools used to collect the information (participatory or otherwise). However, in studies using participatory methods, the researcher has the advantage of having access to narratives from notebooks and Debriefing Documents that help to explain the results from any statistical analysis.

As with any statistical analysis, the first stage in processing the information should be exploratory. The use of simple descriptive statistics: average, minimum, maximum, median, standard deviation for continuous data, or frequencies and tabulations for categorical data, are recommended. More complex statistical methods of analysis may be required, but no researcher should jump straight into the deep end of the statistical methods pool. Like any other area of speciality, statistics offers tools that can yield great benefits if used appropriately, but can lead to wrong results if used without appropriate knowledge. As a rule of thumb, if a statistical analysis is getting you confused and detached from the information, stop and re-think. In these cases you have two options: abandon the use of such analysis, or seek help from a 'friendly' statistician who can advise on the use and interpretation of the analysis.

Statistical data processing almost inevitably leads to loss of detail as information from different sites is aggregated, but gives the advantage of allowing generalisations. An advantage of using participatory methods in research is that the non-numerical, in-depth information can be incorporated into the analysis[45]. Researchers should be able to establish clear, credible links between the qualitative and the quantitative information in their final analysis.

[45] Here we use analysis in the more general sense as the process of reasoning that enables learning. This includes the results from the 'statistical analysis' or statistical processing.

7.3 Scaling up

How do you scale up study results? If you ask a statistician this question, he/she will think about weighting the results so that they apply to the full population. We will explain the statistician's point of view using a simple hypothetical example. For an example of how one process of scaling up was done, see Wingfield Digby (2000).

Suppose that we are working on an island where a programme of distribution of free inputs to farmers has been in place for one year. There are three districts on the island: Chambers Hills, Potterland and Wilsonia. For the sake of simplicity, the villages on the island all have the same number of households. The free inputs programme has been managed by each district independently. In order to find out the percentage of households that receive the benefit and other information relevant for policymaking, the research team decides to carry out research using participatory methods. In each district, a total of 15 villages are visited. The results for the percentage of households receiving free inputs in each district are as follows:

District	**% of households that received inputs**
Chambers Hills	40
Potterland	25
Wilsonia	10

However, from the island census of 1998 it is known that Chambers Hills has 20,000 households, Potterland has 10,000 and Wilsonia has 15,000. The whole island has 45,000 households. If the research team is interested in calculating the percentage of households receiving the free inputs programme in the whole island, weighting of the district results would be necessary before scaling up the results for the whole island. The weighting would be done as follows:

Contribution from Chambers Hills:	40% × 20,000 /(45,000)	= 17.8%
Contribution from Potterland:	25% × 10,000 / (45,000)	= 5.6%
Contribution from Wilsonia:	10% × 15,000 / (45,000)	= 3.3%
All districts together:	17.8% + 5.6% + 3.3%	= 26.7%

The percentage of households receiving the free inputs in the whole island is 26.7%.

The derivation of weights depends on the sampling design. In our example, equal samples were taken from each district, despite the fact that districts were of different sizes. Therefore re-weighting of the district results was needed to produce island-wide figures. However, the sampling could have been done in such a way that each household in the island had the same probability of being part of the sample. In this case, the sample would be self-weighting and there would be no need for the application of weights.

Information is not always available to construct useful weights. In such cases, the statistician might decide to use secondary sources to help in the task of deriving weights, or decide that no weighting would be preferable to using incorrect weights.

8 Empowerment

The research using participatory methods discussed in this paper is designed to produce results at area, district, regional or national levels. It can be used to empower stakeholders at these levels – from NGOs and members of civil society to local, regional or national officials. This is, in our view, a very important function, and one that is often neglected: empowerment should not only involve local communities, but those whose decisions on a higher level of aggregation may affect the lives of people in communities spread across large areas.

However, research using participatory methods is generally weak on local-level empowerment. Unlike the participatory work that aims to empower local groups or communities, research using participatory methods does not usually focus on a local development outcome. As a result, it often fails to feed information back to decision makers and other stakeholders in the communities where the research is carried out. In our Malawi studies, for example, the data generated by the research was processed by the researchers on return to base, and the analysis was not shared with the communities that produced it.

We believe that – while the focus of research using participatory methods will continue to be aggregation to the higher level – there is potential to develop techniques for sharing results with the communities that take part in the studies. The initial data handling could be carried out by the researchers while in the field. For example, community-level tables comparing FSS and TS like that in Section 5.4.1.3 could be produced from the community mapping with cards and presented for discussion in the community where the data comes from.

However, our experience suggests that if the issues are of a sensitive nature (like Starter Pack/TIP), care is likely to be needed when making feed-back to local communities. This is important mainly because if sites are selected at random, many of them may be in places where there are no social programmes or NGOs. There would be no follow-up once the research programme had finished, and the researchers would not be in a position to mediate any conflicts that were produced.

Thus, empowerment is an area which has considerable potential in relation to research using participatory methods, but a sensitive approach is required. Work is needed to develop guidelines about what sort of numerical data might be fed back

into communities, how it should be done, and what sort of local empowerment outcomes might be expected to come from it.

9 Ethical issues

In this section, we discuss some of the ethical issues that have arisen in connection with our research in Malawi. Most of these are questions to which we do not have clear answers. We restrict the discussion to those ethical issues most closely related to research using participatory methods (as distinct from other types of participation). Unlike survey research – for which there are established ethical procedures – this is an area in which the ethical rules are not clear and need to be developed.

9.1 Ignoring the evidence

Should the researcher collect data for bodies (national, regional or local) that have no commitment to using the evidence as a basis for policymaking? The process described in this paper requires an intensive use of resources, not least in terms of time spent by the members of communities who take part in the studies. While community-led participatory studies are designed to feed back into local communities or groups within them by empowering them to make improvements (see Section 8), research using participatory methods that aims to influence policy at a higher level will only impact positively on the lives of the participants if policymakers use them to reduce poverty, increase livelihood opportunities or otherwise benefit the target populations. We should not ask communities (or groups within them) to spend time on research if we do not believe that the policymakers will take the findings seriously.

By introducing statistical principles that make the data collected representative of a population of interest and by making the research process a rigorous, well-organised one, we argue that it is more likely that policymakers will believe in the findings of research using participatory methods and act on them. However, more important factors such as party politics, personal interest or budget limitations may outweigh the evidence. In such cases, we believe that research using participatory methods should be avoided, as it will raise expectations that will not be fulfilled.

In cases where governments at national, regional or local levels are themselves undertaking or commissioning research using participatory methods, we believe that there should be a commitment beforehand to take the results seriously. Otherwise, resources will be wasted and participants' expectations disappointed.

Box 9.1: What is the benefit for the communities?

In the Malawi studies, when the researchers using participatory methods introduced themselves to the communities, they explained that they wanted to find out about how people lived and what were their problems with the intention of transmitting this information to policymakers so that they could make changes for the better. Whether or not free inputs were mentioned in the introduction or later on (see Section 9.2), people tended to interpret this as meaning changes to Starter Pack/TIP, as this was such an important food security issue. They were happy to participate in the research on this basis. However, in 2000-01 when some of the same villages were visited as in 1999-2000[46], people complained that nothing had changed for the better – in fact the programme had been scaled down from nearly 3 million to 1.5 million packs.

9.2 Transparency and consent

We believe that it is important not to hide information from the communities in which we work. At the outset, the teams make the general purpose of the visit clear to the VH and potential participants at an open village meeting and seek their consent for carrying out the research (see Box 9.1). However, we do not always disclose our whole agenda at once. For instance, we have often opted to talk about general livelihoods, problems, food security, etc. before specifically discussing Starter Pack/TIP. This is partly because free inputs are such an emotive issue in Malawi that once they are mentioned it is almost impossible to channel the discussion in any other direction; it is also partly because we often want participants' responses not to be influenced by what they think we should hear or will want to hear in relation to the free inputs programme. Therefore we intentionally avoid mentioning Starter Pack/TIP until we have covered topics that may be adversely affected, disclosing information in a pre-arranged sequence.

The ethical question here is whether it is right to restrict information flow to participants at any stage in the process. Medical trials have established procedures for what is and is not acceptable in this respect, but practitioners of participation generally favour complete openness. In the example of community mapping with cards (see Section 5.4.1), the information on the cards about receipt of TIP was deliberately coded during the mapping so that the VTFs, beneficiary and non-

[46] This was done by a case-study based module – Van Donge (2001) – that purposively selected villages that had been visited the previous year.

beneficiary focus groups would not be confronted with information about whether the household on the card that they were looking at was a beneficiary or not (although of course most of them knew anyway). This was intended to get the focus group participants thinking along the lines 'Does this household deserve to receive or not?' rather than, for example, 'This household received – should it have done so or not?' We felt that by removing the direct association between receipt of TIP and criteria for receiving, we would get a more 'objective' assessment of each household's situation.

However, in our Dealing with Data from Participatory Studies workshop[47], one of the participants asked about the community mapping with cards, "Is the card game participatory if it has so much to hide?" We feel that – although our approach has merit in terms of the quality of information collected – the answer to this question is unclear. Some participatory practitioners will feel that our procedure is not open enough. If it is felt that the principle of openness should not be sacrificed, it may not be possible to collect certain types of sensitive information.

A possible compromise might be to collect the information by keeping some participants blind to information provided by others, for example by using codes, and then feed back the results so that participants can learn from them (see Section 8). This would based on a 'contract' whereby participants agree to 'play the game' first, with some elements unexplained; and facilitators agree to 'reveal all' at the end. This process could be iterative, in that participants might then want to 'correct' some of the information given in the light of the full picture. They would be allowed to do so, and both initial and final positions would be recorded.

9.3 Flexibility

Participatory work is highly flexible where the agenda is determined by the local community or groups within it and the objective is one of local empowerment or development. By contrast, research using participatory methods is highly structured: research objectives are fixed and information collection techniques, including certain assumptions, definitions and indicators, are standardised (see Section 5). Some would argue that this approach is incompatible with the principles of participation, which requires flexibility of discussion. However, we believe that our approach is compatible with participatory principles if the following conditions are met:

- standardisation in the main fieldwork phase is based on consultation with the communities in the preliminary design phase;

[47] Statistical Services Centre, The University of Reading, 2-13 September 2002.

- we fix the elements that we know about in advance, but leave as flexible as possible the unknown elements – the issues to be discussed; and
- although the activities and information collection tools are standardised and certain information needs to be recorded in every site, discussion is not limited – the facilitators should allow participants to develop the discussion in a way which deals with their concerns as well as those of the researchers.
- the main principles of participation are embraced.

9.4 Sharing information

In the short term, research using participatory methods takes more from the communities that participate in the research than it gives back to them. We are transparent about the objectives of the research (e.g. to inform policy at national level), and it is on the basis of this 'informed consent' that people are asked to participate at the start of work in each site (see Section 9.2). Nevertheless, it is a matter of concern that there is little immediate benefit for participants' communities.

One possible improvement would be to include a commitment to information-sharing as part of the terms of reference of research studies using participatory methods. This would have the aim of facilitating learning as part of the research process. Exactly what information would be shared would depend on the nature of the study. For instance, a study relating to HIV/AIDS, such as those carried out by TIP1 Module 3 or TIP2 Module 2, might explicitly stipulate that certain information (e.g. about measures to prevent the spread of the disease or how to treat patients) should be shared with participants. This would imply bringing in a health worker during the training at the start of the main fieldwork phase (see Section 6.3) to clarify what messages should be transmitted. Similarly, a study looking at use of the free inputs might incorporate as part of its objectives a discussion of how to plant the seed and apply the fertiliser in order to get the most from the inputs. This would imply bringing in agronomic expertise during the training for the main fieldwork phase.

9.5 Confidentiality

A final area of concern with research based on participatory methods is that of confidentiality[48]. In survey work, clear rules have been established to ensure that confidentiality of responses is honoured, including requirements to aggregate responses to the level where no individual respondent can be identified. However, participatory approaches are open about who provided the information and

[48] According to De Vaus (1996), "Confidentiality simply means that the researcher can match names with responses but ensures that no one else will have access to them".

encourage acknowledgement of sources. This is because the information is normally collected for the use of the community that provided it and the discussions are open to all participants. Thus, confidentiality is not an issue. However, for research using participatory methods, where the end-user is outside the local community, confidentiality is an important consideration. It is conceivable that information might be given by Village A or by Participant B that might make them unpopular with the authorities or other stakeholders. If the source of information were to be made public, Village A or Participant B might be victimised. This is an area which, in our view, requires urgent debate, and for which guidelines need to be established.

10 Conclusion

This paper sets out to show that it is possible to generate statistics which will be taken seriously by policymakers using participatory methods in research. We have argued that in order to achieve this objective, it is necessary for research using participatory methods to produce results from a 'representative' sample, which can be generalised in order to reach conclusions for the population of interest. This is usually seen as characteristic of research using surveys, but there is no reason why it should not be possible for research using participatory methods. Indeed, the Malawi experience presented in this paper shows that it is possible.

The approach described here combines two elements to achieve this result:

First Statistical principles are employed in the design of the research, including the selection of the sample of communities in which the research will be carried out, and, if necessary, of samples within the communities.

Second The PRA tools are adapted through a process of standardisation, based on careful planning and organisation of fieldwork, taking into account the requirements of comparability of data between sites.

We have argued that if the design of research studies using participatory methods follows this approach, the numeric component of the data generated will be suitable for standard statistical analysis, including scaling up of results by weighting if necessary. The statistics produced by such studies should be capable of informing policy at national level.

We have pointed out that the choice between surveys and participatory methods in research depends on the type of information (including the numerical data) which the researcher needs to collect, in particular the level of complexity. Research using participatory methods should by no means replace survey work. Rather, we show that the two approaches are complementary. They can be combined within a 'modular' research programme – as was the case with our Malawi studies. Surveys can also be useful add-ons to studies using participatory methods.

A key characteristic of our approach is that research teams need to work in a relatively large number of sites compared with traditional participatory studies. Therefore, the paper provides some pointers about cost considerations. A comparison of costs between two hypothetical studies, a research study using participatory methods and a survey, based on information from Malawi in 2001-02, shows that the cost of the two types of study is of a similar order of magnitude.

Finally, we have raised some concerns about empowerment and ethical issues, which present challenges for the future. Are the fundamental principles of participation undermined by the modifications that we need to make if we are to generate data for use beyond the local level? Does research for end-users outside the community imply a need to introduce new safeguards and seek commitments from such end-users about the use of the information? In our view, there is a need to develop these matters further and to produce guidelines for future studies.

List of references

Abeyasekera, S., Ritchie, J.M., and Lawson-McDowall, J. (2002): 'Combining ranks and scores to determine farmers' preferences for bean varieties in southern Malawi'. Experimental Agriculture (2002), volume 38.

Barnett, V. (1991): *Sample Survey: Principles and Methods*. London: Edward Arnold.

*Binauli, L., Kunkwenzu, E., Polepole, M., and Chipeta, L. (2000): 1999-2000 Starter Pack Evaluation Module 3: Sustainable livelihoods, intra-household distribution and gender aspects of Starter Pack.

Chambers, R. (1994a): 'Participatory Rural Appraisal (PRA): Analysis of Experience', World Development Vol. 22, No. 9.

Chambers (1994b): 'Participatory Rural Appraisal (PRA): Challenges, Potentials and Paradigm', World Development Vol. 22, No. 10.

*Chinsinga, B., Dzimadzi, C., Chaweza, R., Kambewa, P., Kapondamgaga P. and Mgemezulu, O. (2001): TIP 2000-01 Evaluation Module 4: Consulting the Poor on Safety Nets.

*Chinsinga, B., Dzimadzi, C., Magalasi, M. and Mpekansambo, L. (2002): 2001-02 TIP Module 2: TIP Messages.

Cochran, W.G. (1977): Sampling Techniques. New York: Wiley.

*Cromwell, E., Kambewa, P., Mwanza R. and Chirwa, R. with KWERA Development Centre (2000): 1999-2000 Starter Pack Evaluation Module 4: The Impact of Starter Pack on Sustainable Agriculture in Malawi.

*Dzimadzi C., Chinsinga, B., Chaweza R. and Kambewa, P. (2001): TIP 2000-01 Evaluation Module 3: Agricultural Communications.

De Vaus, D.A. (1996): *Surveys in Social Research (4th edition)*. London: UCL Press.

*Levy, S. and Barahona, C. (2002): 2001-02 TIP, Main Report of the Evaluation Programme.

Marsland, N., Wilson, I., Abeyasekera, S., and Kleih, U. (2001): 'A Methodological Framework for Combining Quantitative and Qualitative Survey Methods' in *Combining Quantitative and Qualitative Survey Work: Methodological Framework, Practical Issues and Case Studies* – Report for DFID NRSP (SEM) Project R7033 (www.reading.ac.uk/ssc)

Statistical Services Centre, The University of Reading (2001): *Some Basic Ideas of Sampling, Statistical Good Practice Guidelines* (www.reading.ac.uk/ssc).

*Van Donge, J.K., Chivwaile, M., Kapondamgaga, P., Kasapila, W., Mgemezulu, O., Sangore, N. and Thawani, E. (2001): 2000-01 TIP Module 2 (2): A qualitative study of markets and livelihood security in rural Malawi.

Wilson, I.M., (2001): *Sampling and Qualitative Research*. Report for DFID NRSP (SEM) Project R7033 (www.reading.ac.uk/ssc).

Wilson, I.M. (2002): 'Some practical sampling procedures for development research', presented at the conference Combining Qualitative and Quantitative Methods in Development Research, Centre for Development Studies, University of Wales.

*Wingfield Digby, P.K. (2000): 1999-2000 SP Module 5: Measuring the size of the rural population in Malawi.

* References marked with an asterisk belong to the Malawi Starter Pack/TIP monitoring and evaluation archive, and can be obtained on CD-ROM from the DFID offices in Lilongwe or from c.e.barahona@reading.ac.uk

Appendix 1: Malawi study outlines

1999-2000 Starter Pack evaluation (SP2)

Module 3: Gender and intra-household distribution

Reference: Binauli (2000).

This study assessed the sustainable livelihoods, intra-household distribution and gender-related aspects of the impact of Starter Pack. It examined the livelihoods of the communities visited; the intra-household distribution aspects of food and child nutrition; and gender differences in use of Starter Pack inputs and outputs.

Selection of sample:

For the main phase of the study, 48 villages were selected from 24 Extension Planning Areas (EPAs). Stratification of EPAs was on the basis of region and 1996 Famine Early Warning System (FEWS) Vulnerability Assessment Mapping (VAM) Food Deficiency Index. The sampling was done in two stages: in the first stage, a sample of EPAs was selected at random from within the strata, and in the second stage villages were selected at random within the EPAs.

Within villages:

Villagers were called together by the chief, and the main objectives of the study were discussed. People were asked to take part in three focus groups – male household heads, married women and female household heads. The focus groups discussed livelihoods and problems faced by communities. They also used a tool known as the 'Bingo Game' to collect data about poverty based on ownership of assets, and answers to some key questions about farming, use of seed and fertilizer and the relationship between Starter Pack and food self-sufficiency (see Binauli, 2000, Appendix A).

The team also conducted a survey covering household characteristics, socio-cultural practices, gender relations, food self-sufficiency, intra-household food distribution and child nutrition. For the survey, respondents were selected at random from the focus groups by picking a piece of paper with a number (the aim was to select 3 male heads, 3 married women and 4 female heads in each village). A total of 476 interviews were carried out with those selected.

The study did not make a link between the information collected by the Bingo Game and that collected by the survey. This would have been easy to do, simply by

numbering the Bingo Game responses and then recording the participant's Bingo Game number on the survey questionnaire when he/she was selected.

1999-2000 Starter Pack evaluation (SP2)

Module 4: Sustainable agriculture

Reference: Cromwell (2000).

This study focused on the contribution of the Starter Pack scheme to sustainable agriculture. It used participatory M&E approaches to elicit evidence from smallholder farmers. It aimed to capture regional variations and differences between farming practice groups (high, medium and low sustainability) as well as to report on national trends and patterns.

Selection of sample:

EPAs were stratified according to the 1996 FEWS VAM Sphere of Influence clusters. Thirty EPAs were selected at random; the number selected in each stratum was proportional to the number of EPAs in the stratum at national level. In the selected EPAs, one village per EPA was selected at random, to give a sample of 30 villages. Villages with very small numbers of Starter Pack beneficiaries were excluded, as it would have been difficult to conduct the research using participatory methods in such villages.

Within villages:

A preliminary study in three villages explored the suitability of a range of PRA tools for the main phase of the study and identified 15 sustainability indicators that were mentioned consistently by farmers. In the main phase, the field work in each village comprised:

- introductions;
- resource, social and institutional mapping and a transect walk;
- ranking of relative importance of Sustainability Indicators;
- classification of households into 'high', 'medium' and 'low' sustainability farming practice groups, using the list of Sustainability Indicators;
- focus group discussions with each farming practice group about:
    - the relative importance of different Sustainability Indicators;
    - trend analysis of factors influencing the sustainability of farming over time;
    - impact of Starter Pack on farming;
    - ideal contents of a 'Dream Pack'; and
    - changes to Starter Pack agricultural extension and logistics.

1999-2000 Starter Pack evaluation (SP2)

Module 5: Measuring the size of the rural population in Malawi

Reference: Wingfield Digby (2000).

This study critically examined the number of rural households registered by the Starter Pack Logistics Unit (SPLU). The SPLU figures were far out of line with the information from the 1998 census, and it was felt that registration to receive Starter Pack might have encouraged the recording of more individuals as heads of household than would be found using any conventional household definition. However, the Ground Truth Investigation Study (GTIS) conducted by Module 5 supported the SPLU's estimates, producing an estimate of 2.78 million rural households – compared with 2.89 million estimated by the SPLU and 1.95 million estimated by the census.

Selection of sample:

For the GTIS, a sample of 30 EPAs was selected at random from within Agricultural Development Divisions (ADDs), following allocation of EPAs to the country's eight ADDs in the following manner: Karonga 3, Mzuzu 3, Kasungu 4, Lilongwe 5, Salima 3, Machinga 4, Blantyre 5 and Shire Valley 3. Two villages were selected at random from within each of the 30 selected EPAs. The sample comprised 12 villages in the northern region, 24 in the central region and 24 in the southern region. However, the team did not carry out the study in 6 villages in 3 EPAs in the southern region. Wingfield Digby (2000) suggests that the sampling process could have been improved by ensuring a greater spread of EPAs within ADDs and a mix of large and small villages.

Within villages:

The GTIS team carried out a detailed mapping exercise in each village which involved firstly establishing the boundaries of the village, and secondly discussing the definition of a household in order to identify every household on the map. The result of this was a participatory household listing, showing 6,326 households in the 54 villages studied. The mapping and household listing process was accompanied by discussions about the Starter Pack registration process, the de-registration of some beneficiaries in 1999-2000 and the distribution of the free inputs.

In each village, the mapping and household listing provided the basis for an enumeration of households, in which every household was interviewed using a simple questionnaire to obtain demographic data (household size and composition) and information about which members of the household cultivated a garden, registration

for SP2, deletion from the SP2 register and receipt of packs. Using the mappings and the survey, the researchers were able to provide regional and national population estimates by grossing up the data (using weights and adjustments for non-response, including failure to visit 6 of the sampled villages).

2000-01 TIP evaluation (TIP1)

Module 3: Agricultural communications

Reference: Dzimadzi (2001).

This study was designed to learn from the communications experience of the 2000-01 TIP with a view to increasing the impact on farmers of future communications messages, including the leaflet on use of TIP inputs and the HIV/AIDS leaflet.

Selection of sample:

Strata were defined on the basis of the predominant languages and religion for administrative districts: language was used for stratification as it was thought likely to affect comprehension of communications messages. Religion was used for stratification on the hypothesis that it might affect attitudes to HIV/AIDS. The information about language and religion was obtained from the National Statistical Office (NSO) tabulations of the 1998 census data.

For the main phase of the study, 27 villages were selected using systematic circular sampling from within the list of villages in each stratum. The number of villages selected from each stratum was proportional to size of the stratum. Very small villages were removed from the sampling frame because it was likely to be difficult to find enough TIP beneficiaries to work with, while very large villages were removed because participatory workshops might have been difficult to handle.

Within villages:

At the village level, a combination of approaches was used: a survey, semi-structured interviews with key informants and participatory focus group discussions. For the focus groups that looked at agricultural communications, the study convened one focus group for men and another for women in each village. For the HIV/AIDS focus group, the participants were split by age and sex – men, women, boys and girls – because the study was interested in potential differences in attitudes between generations as well as sexes.

The research using participatory methods element of this study was not based on an objective method of selection of focus group members. The team invited people to

volunteer to join the discussions, aiming to make up groups of eight to ten participants. An attempt was made to get a mixture of literate and illiterate participants. However, participation was voluntary. Basic information about each member of the focus group was recorded in the Debriefing Document so that it would be possible to tell (using other data sources such as the 1998 census) whether or not the characteristics of the groups were similar to those of the general population in the same age and sex categories.

2000-01 TIP evaluation (TIP1)

Module 4: Consultations with the poor on safety nets

Reference: Chinsinga (2001).

This study was intended to contribute to the development of the Government of Malawi's National Safety Net Strategy. It looked at existing interventions such as Starter Pack and TIP and planned interventions such as direct welfare transfers. The study provides evidence about what the rural poor themselves think about different forms of safety net provision, methods of beneficiary selection and ways of managing the benefits. It discusses the feasibility of community poverty targeting for different types of intervention.

Selection of sample:

Owing to time limitations, this study was only able to visit 20 villages in the main fieldwork phase. In each district, one village was selected using simple random sampling. In order to reduce the number of districts from which villages would be selected from 27 to 20, the study excluded from the sampling frame the districts that were visited during the preliminary phase. Likoma Island was also excluded because it would require too much time to reach. Another three districts, one in each region, were randomly selected for exclusion from the sampling frame. Very small villages were removed from the sampling frame because it was likely to be difficult to find enough TIP beneficiaries to work with, while very large villages were removed because participatory workshops might have been difficult to handle.

Within villages:

Focus group discussions were held on traditional support systems, management of direct welfare transfers, TIP registration, distribution and management. The study made use of specially adapted PRA tools: community mapping with cards, seasonal calendars and scoring. The study also sought the views of the village heads on how the transfers could be managed to ensure fairness, transparency and accountability through individual interviews with each village head.

2001-02 TIP evaluation (TIP2)

Module 2: TIP messages - beneficiary selection and community targeting, agricultural extension and health (TB and HIV/AIDS)

Reference: Chinsinga (2002).

The study looked at whether the TIP2 communications experience represented an improvement compared with TIP1. It elicited information from stakeholders in 20 communities throughout Malawi about beneficiary selection and poverty targeting; it assessed the effectiveness of the TIP2 agricultural extension campaign; and it looked at the impact of the TB and HIV/AIDS leaflets that were distributed with the packs.

Selection of sample:

Owing to resource limitations, this study was only able to visit 20 villages in the main fieldwork phase. Villages for the main study were selected using a two-stage sampling process. In the first stage, 20 districts were selected randomly from within each region. The north, centre and south were allocated 5, 6 and 9 districts respectively. The districts visited during the preliminary phase were excluded at this stage. At the second stage, a village was selected at random from within each district. Very small villages were removed from the sampling frame because it was likely to be difficult to find enough TIP beneficiaries to work with, while very large villages were removed because participatory workshops might have been difficult to handle.

Within villages:

The study used group discussions and semi-structured interviews with key informants. For the beneficiary selection and poverty targeting part of the study, the discussions focused on two PRA tools: community mapping with cards and scoring. For the part of the study looking at the agricultural extension campaign, the study used semi-structured interviews with extension workers and farmers (during field visits), as well as focus group discussions exploring the leaflet on use of TIP inputs. For the study of TB and HIV/AIDS leaflets, focus groups were organised with men, women, boys and girls separately, as with TIP1 Module 3. As with that study, basic information about each member of the focus group was recorded in the Debriefing Document so that it would be possible to tell (using other data sources such as the 1998 census) whether or not the characteristics of the groups were similar to those of the general population in the same age and sex categories.

For all the Malawi studies, the sampling frame for villages came from the SPLU register of beneficiaries, later referred to as the TIP Logistics Unit (TIPLU) register.