Oxfam GB Evaluation Guidelines

Introduction

These guidelines have been produced to support the effective implementation of Oxfam GB's Programme Evaluation Policy. Recognising that a lot of good quality Monitoring and Evaluation Guidelines and manuals have already been developed and are widely available, these guidelines are intended to provide a top line overview, introduction to programme monitoring and evaluation that will support staff in delivering quality programme evaluations. They draw on many of these existing manuals and guidelines, and owe a particular debt to the course materials provided by the International Program for Development Evaluation Training. Importantly, the guidelines are intranet based to allow us to continue to develop and refine them as we go, and ensure that they remain useful, relevant and reflective of current thinking. Links to tools and further resources providing more detailed guidance on specific areas of M&E practice can be accessed via the intranet pages.

1 Overview: The "E" in M&E plans

All programmes should have an appropriate **monitoring & evaluation plan** from the outset.

Monitoring and evaluation (M&E) are part of everyday programme management and are critical to the success of our programmes. They help programme teams to learn what does and doesn't work in their efforts to overcome poverty and suffering and to adapt their programmes in light of what they find. This, in turn, helps Oxfam GB to maximise the effectiveness of its interventions. The processes and products of monitoring and evaluation also provide a documented record of the programme, and help strengthen accountability – supporting Oxfam GB to give an account to its wide range of stakeholders for its decisions and actions, and providing opportunities to take account of their views and opinions.

Monitoring is the routine, ongoing collection and review of information on a programme's activities, outputs, and outcomes that provides programme managers and other stakeholders with indications of progress against programmes plans and towards programme objectives. It is a collaborative process between Oxfam GB staff, partners and communities to review what has happened, identify intended and unintended changes and consider whether activities have contributed to those changes. This regular collection of information shows whether or not the programme is performing as expected, or if adjustments are necessary. Well planned, timely monitoring allows problems to be quickly identified and programme activities to be adapted in order to optimise their impact. Monitoring is more effective when it is a continuous process, included in the design of a programme and part of our day-to-day work.

Evaluations complement ongoing monitoring activities by providing more in-depth, objective assessments of a programme's design, implementation and results at particular points in time. Where monitoring shows general trends, evaluations generally help explain 'why' things are happening the way that they are. Programme evaluations may be undertaken at any point in the programme cycle where there is a need to learn more about how the programme is working, or to be accountable for the resources with which we've been entrusted. At a minimum, programme teams should be evaluated midway through implementation (a formative or mid-term evaluation) and once the programme has completed (a summative or end-term evaluation).

1.1 Key components of an M&E plan

There are four key components to an M&E plan:

1) Logic models
2) Indicators
3) Data collection (including Baseline Studies)
4) Moments for review

1.1.1 Logic Models

Good quality monitoring and evaluation depends on coherent programme design. Coherent programmes have strong internal logic which makes clear: what the programme intends to achieve, how it expects to achieve it, and what assumptions are being made – both in terms of how the programme will be implemented, as well as the programme's theory of change (i.e. how the programme expects change to happen within a given context). A programme's theory should be aligned with the country's national change strategy.

While it is understood that change is not linear, and that programmes are multi-faceted and are implemented in complex environments, it should be possible for any given programme team to clearly identify and communicate what change(s) it is trying to achieve (outcomes and impact) and how it thinks it will achieve these (strategies/ activities and outputs) in a simple programme logic model. Logic models can take many different forms (for example, a columned impact chain, or a more free flow diagram), but as a rule, they should be no longer than one page, and should communicate how the programme team expects to bring about change. This includes:

- the strategies or types of activities the programme team will undertake;
- the expected outputs of these activities, and
- the outcomes, and ultimately the impact(s), that the programme team expects these outputs to contribute to bringing about

Programme logic models help ensure that the programme objectives are realistic and that the strategies for bringing about these changes will be effective. The process of creating a programme logic model together with partners, and other key stakeholders, helps to build a joint vision of how change is expected to happen and the role each party will play in bringing about this change. Logic models also help to make explicit the assumptions that are being made about the relationship between programme activities and changes the programme is intending to bring about. Continually checking the validity of these assumptions is a critical part of monitoring a programme, and is key to assessing the programme's contribution towards that change.

Logic models and Logframes

In some situations a Logical Framework, or log frame, may double as a programme logic model, but generally speaking the purpose of these two tools differs. A logic model is provides an overview of the programme and describes its overarching logic. A log frame is a more detailed document that is generally used as a programme management tool.

Many donors require log frames as part of a programme proposal and reporting. In some situations a Logical Framework, or log frame, may double as a programme logic model, but more often programme teams will need to develop them as two separate documents. It is worth noting that different donors often have their own terminology for the different programme components. For example DFID talks about purpose, rather than outcomes, and the EC talks about results rather than outputs.

1.1.2 Indicators

Oxfam GB uses the OECD/ DAC terminology, and defines an indicator as a "quantitative or qualitative factor or variable that provides a simple and reliable means to measure achievement." They are specific measures that, when tracked systematically over time, indicates progress (or not) towards a specific target. Indicators measure change directly or indirectly (known as "proxy" indicators) and can be quantitative and/or qualitative (it is most useful to employ a balance of quantitative and qualitative indicators). Indicators are not designed to measure the whole story of the logic of the programme, they are just designed to give an indication. Good indicators are essential for a good monitoring system as they determine what data needs to be collected and analysed.

There are two different types of indicators:

- <u>Process indicators</u>: show how the project is progressing, what is happening and if the original plan is being followed. They can measure amount of work being done as well as the quality and timeliness. These indicators are developed at output level in your logic model.

- <u>Outcome/ Impact indicators:</u> are quantitative measures or qualitative judgements (or both) by which the achievements of outcomes (or the positive/negative "impact" on the target group) can be judged. These indicators are developed at outcome/objective or impact/goal level in your logic model.

It is be important to have a small number of indicators at each level in the programme logic. Most programme teams identify good output indicators, but struggle or fail to identify indicators that show change at outcome and impact levels. Without outcome and impact indicators, programme teams can only know what activities are being completed, but not what effect these activities are having. Outcome indicators allow us to test the assumptions we have made in our logic model and know if our theory of change makes sense.

All indicators should be:
- Linked to the programme logic model (or the essential logic of the log frame)
- Appropriate - different kinds of indicators are needed for different stages of the project cycle.
- Flexible – indicators should not be thought of as "set in stone" and should be changed if found to be impractical or immeasurable, or if the situation changes. However money can be wasted collecting data against inappropriate indicators – so its even better to get it right first time!
- SMART[1]: Specific, Measurable, Attainable, Relevant and Time-bound

Remember:
- Indicators measure the outcomes on vulnerable and socially excluded people, especially women, older people, men and women with disabilities or who are affected by HIV/AIDS. Wherever possible, outcome indicators should be developed in partnership with communities and partners.
- It is not helpful to develop too many indicators, or similar indicators that require different data.
- The programme team should know how it would gather the data (and fund data gathering) to measure the indicator when the indicator is agreed.
- Vague and ambiguous terms for indicators should be avoided - such as: to increase awareness…, appropriate use …, beneficiary involvement…, increased participation…, good understanding…

1.1.3 Data Collection (& baselines)

See Section 3 below

1.1.4 Moments for review

Data isn't meaningful by itself, it needs to be reviewed and analysed for meaning. It is important to plan moments to review the data that will be collecting from the outset. These will be points in time where the information being collecting is reflected on an analysed in order to consider what it is saying about the programme and, where appropriate, to make decisions to adapt, expand or stop parts of the programme. These moments for review include Monitoring Reviews, and project or programme evaluations.

---

[1] More information about the definitions of SMART and SPICED can be found in Chris Roche's book "Impact Assessment for Development Agencies: Learning to Value Change".

Monitoring Reviews are points in time where programme teams look at the monitoring data it is collecting for information about whether or not the programme is performing as expected, or if adjustments are necessary. Oxfam GB programme teams are required to hold Monitoring Reviews every 6 months, but some may choose to hold them more frequently, particularly if they are working in a rapid onset emergency, or on an advocacy and campaign initiative, where there is a need to respond quickly to rapidly changing contexts.

Project or programme evaluations also build from ongoing and one-off monitoring activities, but they provide more in-depth, objective assessments of our programme's design, implementation and results at particular points in time. In particular, they ask: 'Are the right things being done?' (theory of change), 'Are things being done right?' (implementation), 'Are there better ways?' (learning). Some examples of why you might evaluate a programme include:

- To help analyse why intended results were or were not achieved
- To explore why there may have been unintended results or consequences
- To assess how and why results were affected by specific activities
- To shed light on implementation processes, failures, or successes that may occur at any level
- To provide lessons, highlight areas of accomplishment and potential, and offer specific recommendations for improvement and reform.
- To help make resource allocation decisions
- To help rethink the causes of a problem

At a minimum programme teams should plan to evaluate their programme midway through implementation (a formative or mid-term evaluation) to check that it is on the right track, and once the programme has completed (a summative or end-term evaluation) to learn more about the results of the programmes, and make an assessment about how effective the programme has been.

2 Programme Evaluations

Oxfam GB minimum standards:

- A mid-term and a final, independent evaluations to be undertaken at least once every 5 years of (1) all major campaigns; (2) all programmes with an overall PIP value of £1M or more; and (3) all high profile, risky or complex programmes with significant learning potential.
- Plans and resources required for evaluations be included in PIP Monitoring and Evaluation plans in OPAL.
- Evaluations cover: assessment of the programme's outcomes or impact, cost-effectiveness and Oxfam's contribution.
- All evaluation recommendations are discussed, action points are agreed and learning is fed into future programme development. Managers are responsible for writing a management response and ensuring that action points are followed up.
- All evaluations and management responses are uploaded onto OPAL and sent to the Country Director, the Regional MEL lead and the global MEL team for inclusion on our website.
- Findings from evaluations are shared with partners, affected populations and donors as appropriate.

2.1 Different Types of Evaluations

There are many types of evaluations and they can be categorised in many different ways – by purpose, by timing, by strategy, by philosophy etc. Being able to identify or name the type of evaluation is less important that being clear about what is needed from an evaluation and how information can be accurately collect and interpreted.

3 main types of programme evaluations, categorised by purpose, are:

a. **Process evaluations**

Process evaluations focus on gaining a good understanding of how a programme works. They look at what is being done by a programme, for whom, and in what way in an effort to gain an in-depth understanding of the functioning of a programme. Process evaluations are used to identify where improvements might be made to programme design and delivery, and can also support decisions about whether a programme should be replicated or expanded. While they don't necessarily address the question of the effects of the programme, process evaluations will include an assessment of whether a programme is meeting its objectives. Most early or mid-term evaluations will be process evaluations.

b. **Outcome and impact evaluations**

These are sometimes referred to as final evaluations or summative evaluations. These types of evaluations deliver an assessment of the contributions the programme has made towards immediate changes and/or broader, long-term effects, and help explain why a particular intervention has or has not been successful. Outcome and impact evaluations range in rigour from those that attempt to build a strong case for attributing change(s) to a programme at one end of the spectrum to those who simply measure changes ex-anti and speculate on links between the changes and the programme at the other. Outcome and impact evaluations tend to use experimental or quasi-experimental designs.

c. **Meta-evaluations**

These evaluations pull together findings from two or more evaluations on similar type programmes. They are designed to aggregate the findings from individual evaluations, and often also draw on additional corroborative evidence, in order to draw lessons for the future with a view to improving performance. Meta-evaluations are often desk-based.

There are many different ways that a programme team can approach these different types of evaluations. The evaluation approach describes the overall methodology for the evaluation. It encompasses the values or philosophy of the programme and will inform the strategy that is used to answer the evaluation questions. For example, a programme team may decide to take a participatory approach to their evaluation - all of the evaluation types mentioned above could be approached in a participatory way. Evaluations that take a **participatory approach** emphasize the central importance of programme stakeholders, especially beneficiaries, and often involve an element of capacity building as part of the objective of the evaluation. They focus on understanding behaviours, processes and conditions as they are perceived by participants, and involve some of the stakeholders in the whole or in parts of the activities planned for the evaluation.

2.2 Managing Evaluations (and expectations)

Evaluations can be complicated, and it will be important for someone to take overall responsibility for planning and managing the evaluation, checking that quality standards are met, and ensuring the evaluation conclusions and recommendations are communicated effectively - even where they are led by an external consultant. Usually this will be the responsibility of the commissioning manager.

In addition to managing the evaluation process itself, this will involve managing expectations around the evaluation, ensuring buy-in, and managing the fears of programme stakeholders – including concerns over possible consequences of negative findings, such as issues of reputation or, in some instances, funding. Managing these fears will involve working in an open and transparent way - involving programme staff and other stakeholders in planning the evaluation, and ensuring that there are opportunities for them to review evaluation work plans, findings, and recommendations. It will also involve encouraging a culture of learning that rewards teams who are really trying to understand their programmes, learn from their experience and the experience of others, and respond to issues that arise.

Programme teams may want to consider involving donors in the evaluation in some way, even when the evaluation it not being undertaken in response to donor requirements. This can be as simple as making them aware that the evaluation is taking place. Donors tend to receive such invitations favourably, and involving them from the start of the evaluation can help ensure that they understand the findings and support actions going forward.

2.3 The Programme Evaluation Team

Oxfam GB minimum standards

- All final evaluations to be led by independent evaluators.
- Programme staff and partners play a strong role in all evaluations.

While a variety of stakeholders will have a vested interest in a programme evaluation, a small evaluation team should hold the responsibility for designing and conducting the evaluation. In some situations, the evaluation team may consist of only one member of Oxfam GB staff managing an external consultant. However Oxfam GB recommends that the evaluation team should strive to achieve a gender balance and to be as representative of the different stakeholders groups (such as partners, beneficiaries) as possible. Each member of the evaluation team should have different capacities and will fill different roles and responsibilities. Inclusive evaluation teams will support meaningful participation of different stakeholders in the evaluation process and encourage ownership of the findings and recommendations,

The key is to ensure that each person involved with the evaluation is clear about their roles and responsibilities, and has agreed to them.

2.4 Deciding the purpose and scope of your Evaluation

Oxfam GB minimum standards

All final programme evaluations should include an assessment of:

- the programme's outcomes or impact;
- Our contribution (or 'added value') to the programme
- The programme's cost-effectiveness

Programme evaluations help to meet learning and accountability needs. Within those broad overarching objectives, programme teams will need to make decisions about the specific purpose that their programme evaluation will serve. The purpose of an evaluation will be affected by: stakeholder needs, timing, type of programme, timeline for the evaluation, available resources, and what information is already available.

- **Stakeholder needs -** To help ensure the relevance and utility of the evaluation, it is important to be clear about who the key stakeholders of the programme are, and what purpose(s) the evaluation will serve. Different stakeholders may have different needs with respect to the evaluation, and an agreement needs to be reached with them about the purpose of the evaluation at this early stage in the planning of the evaluation. In

situations where there are competing stakeholders needs, commissioning managers may have to prioritise. The clearer we can be about the multiple purposes of the evaluation, the more likely it will be that the evaluation will ultimately meets the needs of all stakeholders.

- **Timing -** Consider the stage of development of the programme, and what is reasonable to evaluate at a particular point in time. Early or mid-term evaluations tend to be formative, and usually focus on learning in order to improve the programme going forward. Final evaluations, on the other hand, usually require a more objective assessment of the results of programme. In terms of timing then, an outcome/ impact evaluation would not usually be possible (or desirable) until the programme is near completion, or has closed.

- **Type of Programme** - The purpose of an evaluation will be affected by how well the programme is understood. Where programmes are highly innovative, programme teams may want to do a rigorous and objective evaluation to help test the programme's theory of change, and to determine if the approach has been effective. In situations where the programme being implemented is well known and has been proven to be effective, the main purpose of the evaluation may be to learn about implementation.

- **Available Resources** - It is unlikely that programme teams will have the time or budget to evaluate all aspects of the programme at the same level of rigor or at the same time, so you will need to prioritize. To ensure evaluation requirements are met, an appropriate budget for monitoring and evaluating should be included from the outset. This will help ensure that the programme teams are able to undertake monitoring and evaluation that will provide them with the information they need when they need it.

- **Information Available** - Not only does data collection take time and resource, but there will be some information that will not be possible to collect retrospectively which may limit the type of evaluation you can carry out. For example, it would not be possible to conduct a randomized control evaluation unless the programme team has planned for this from the outset and implemented and monitored the programme accordingly.

2.5 Defining the evaluation questions

Evaluation questions build from the purpose and scope, to give direction to the evaluation. They are the specific questions that the evaluation asks, and ultimately seeks to answer, about the programme. Identifying clear questions that the evaluation will answer helps to keep the evaluation focused, and ensure that it delivers information that is relevant and useful. The evaluation questions should be identified with the involvement of all of the key stakeholders. Bear in mind that the way questions are worded can have a big impact on what information is collected, and therefore requires careful thought.

A good place to start is the programme's theory of change – how the programme team thought change would be brought about. A clear logic model of the programme's theory of change will help to illustrate the purpose and content of the programme and will make it easier to develop meaningful evaluation questions from a variety of programme vantage points: context, implementation, outputs, outcomes and impact. If the programme does not already have a programme logic model, those people involved with the programme will need to work together to develop and agree one retrospectively. Where the programme team already has a developed programme logic model, it will be important to review it carefully, check whether the programme was implemented as originally intended, and where necessary refine or rework the existing theory of change.

Another useful framework for developing evaluation questions is the DAC criteria: Relevance, Effectiveness, Efficiency, Impact, and Sustainability - provide a good framework. Oxfam GB's seven key questions about performance and impact of our programmes also provide a useful starting point:
1. What significant **changes** have occurred in the lives of poor women, men and children?

2. How far has greater **equality** been achieved between women and men and between other groups?
3. What changes in **policies, practices, ideas, and beliefs** have happened?
4. Have those we hope will benefit and those who support us been appropriately **involved** at all stages and empowered through the process?
5. Are the changes which have been achieved likely to be **sustained?**
6. How **cost-effective** has the intervention been?
7. To what degree have we **learned** from this experience and shared the learning?

Remember:
- In addition to questions about the programme's implementation or contribution to outcomes, it will also be important to **consider the overarching situation or context in which the programme is operating** which can influence the inputs, activities and the results of a programme. Looking beyond the programme will help to flag issues that the programme team had not considered, and help to pick up on unintended results (positive and negative).

- The **number of questions that can be answered in a single evaluation must be realistic**. Programme teams won't have the time or budget to evaluate all aspects of the programme at the same level of rigor or at the same time, and so you will have to prioritize. Evaluations should focus on the issues of greatest concern – the important questions, the ones that must be answered, as opposed to those that would be nice to know.

- As with the overall purpose and scope of the evaluation, it will be important to **consider the timing of the evaluation relative to the programme cycle**. For example, questions about outcomes and impact are best answered after the programme has been operating for a few years.

Measures or indicators are needed in order to answer the evaluation questions. If the programme has had a robust and responsive M&E plan from the outset, which includes outcome level indicators, then it is unlikely that additional indicators for the evaluation will need to be identified for the evaluation. However, in some situations there may need to identify additional measures or indicators in order to answer the evaluation questions.

2.6 Terms of Reference

The Terms of Reference (TOR) is a written document that describes the overall evaluation, including a summary of the major purpose(s) and expectations of the evaluation, as well as basic information about the programme being evaluated (e.g., context, expected outcomes, and strategies). A Terms of Reference may be developed before or after the key evaluation questions are defined. The TOR will articulate and communicate initial requirements and expectations for the evaluation, and will guide the process of developing a more detailed evaluation plan. It should include:
- Information about the programme that is being evaluated (logic model)
- The purpose and scope of the evaluation
- The questions that the evaluation will answer, or the process that will be used by the evaluation team to identify the key evaluation questions
- Commissioning manager and the members of the evaluation team
- Documentation Review
- The standard against which performance is to be assessed, or analyses are to be conducted
- Reporting requirements
- Budget
- Timeframe
- Evaluator selection process (where appropriate)

In addition to being an important communication tool, the TOR serves as a management tool for an evaluation manager. It establishes the basic guidelines so everyone involved understands the expectations for the evaluation and the context in which the evaluation will take place. The process for developing the TOR can be very useful for ensuring that all stakeholders are included in the discussion and in decision-making about what evaluation issues will be addressed.

2.7 Contracting Consultants

Consultants can be brought in for the whole evaluation or only parts of the study, and may be asked to participate in or lead an evaluation for different reasons. Most often, external evaluators are brought in to an evaluation for their objectivity and professional evaluation skills to an evaluation. They may also be brought in because they have in-depth knowledge on a particular type of programme or subject matter; local knowledge or fluency in a local language; or because they are familiar with Oxfam GB or with our partners. Consultants can be expensive though, and might result in decreased ownership of the evaluation and its findings, and a loss of organisational knowledge. It is important to consider these issues when taking a decision about whether to contract a consultant to carry out all or partof an evaluation, and to try to maximise the benefits of their involvement while minimising some of the drawbacks. The decision about whether or not to contract a consultant to carry out all or part of an evaluation will sit with the commissioning manager.

If a programme team decides to bring in a consultant, an evaluation brief should be developed for the consultant's work. An evaluation brief advertises the post and asks consultants to submit a tender for the position. It should provide information on why the evaluation is needed, what the consultant is expected to deliver (including how, and by when), what approach is required, the questions that the evaluation is expected to answer, how the tender will be managed, the relevant skills sets and knowledge the consultant should have, budget information and any other practical details (such as how the tender should be submitted).

2.8 Data Collection

See Section 3

2.9 Reaching Conclusions and identifying action points

Oxfam GB minimum standards

- Programme staff and partners play a strong role in all evaluations.
- All evaluation recommendations are discussed, action points are agreed and learning is fed into future programme development.
- Findings from evaluations are shared with partners, affected populations and donors as appropriate.

Once data collection and analysis are largely completed, we can start to draw conclusions that answer the evaluation questions. This should be an inclusive process, bringing together stakeholders and the evaluation team to discuss the preliminary findings and insights of the evaluation, to reach consensus on the results and to talk through next steps.

Wherever possible, evaluations should create a space for those involved with the programme to hear and reflect on the findings of the evaluation, and to have a constructive discussion around the key issues identified by the evaluation. This provides an opportunity to involve programme team members in developing action points, recommendations and lessons, and will help ensure that programme evaluations are useful, and lead to improvements in current and future practice. This should be held as a debriefing workshop, with a cross-section of programme stakeholders (as with Real Time Evaluations), and it may

involve feeding back to different stakeholder groups in different ways. However it is done, programme stakeholders should be given an opportunity to hear what the evaluation has found and to be involved in thinking about recommendations going forward.

Once the evaluation is complete and, where appropriate, recommendations have been identified, programme stakeholders should come together to identify clear actions that will be taken to respond to the evaluation findings. These should specify who will do what by when, and should be monitored by the appropriate line manager to ensure that they actually happen.

At a minimum:
- Findings and recommendations from mid-term evaluations should be fed into the next Monitoring Review.
- Lessons from final evaluations should inform future programme development.

Involving senior managers strengthens the likelihood that action points and recommendations will be followed up.

Oxfam GB's Open Information policy means that all programme evaluations will be available to the wider aid sector, and members of the public, on demand, and it is important that findings and recommendations are shared with relevant stakeholders first. In line with this, the results of programme evaluations should be feed back to relevant stakeholder groups in an appropriate manner.

2.10 Reporting Findings

Reporting it is an integral part of evaluation. It helps programme teams to: communicate what they do, monitor and track their progress; demonstrate impact; document lessons learnt; and be accountable and transparent to donors and beneficiaries. In many cases, evaluation reports will be the only lasting record of a programme, including the results it achieved and the lessons that were learnt from its implementation.

An evaluation report will usually contain an executive summary and a full report. The executive summary should provide a quick overview of the evaluation, including the purpose and scope, the key evaluations questions, the methods used, as well as a brief summary of findings and recommendations. It should serves as a short stand-alone document that provides readers with all of the basic facts about the evaluation. The main evaluation report should be a detailed report, with complete statistical and case study analysis. It should cover the purpose of the evaluation; information about the programme, the context in which it operates/operated and it's theory of change; the specific evaluation questions it set out to answer; data collection and analysis methodology (including the limitations of this methodology); who was involved in the evaluation; findings (including supporting data); and conclusions and recommendations.

A **management response** to the evaluation is also required, which reflects careful consideration of the evaluation recommendations as a basis for management decisions going forward. The management response should be brief, normally 1-2 pages, and should be used both to comment the utility of the report and to address the recommendations (how they will be taken forward, or why they will not be taken forward).

2.11 Assessing the Quality of an Evaluation

There are many different ways of approaching and conducting programme evaluations, which produce evaluations of varying degrees of rigour, credibility, reliability and validity. Questioning the quality of the evidence and analysis, and the subsequent findings and recommendations, and making an assessment of the quality of the evaluation overall will help programme teams, and Oxfam GB, know how much weight to give to them. To avoid the suggestion of bias, this should be undertaken with support from an overview manager, or

regional MEL lead. Concerns about the quality of an evaluation should be detailed in the management response.

Rist and Kuseks[2] identify 6 characteristics of quality evaluations, which are helpful in assessing the quality of programme evaluations:

- **Impartiality** – the evaluation information should be free of political or other bias and deliberate distortions. The information should be presented with a description of its strengths and weaknesses. All relevant information should be presented, not just that which reinforces one view or another.
- **Usefulness**: Evaluation information needs to be relevant, timely, and written in an understandable form. It also needs to address the questions asked, and be presented in a form desired and best understood by the key stakeholders
- **Technical adequacy**: The information needs to meet relevant technical standards – appropriate design, correct sampling procedures, accurate working of questionnaires and interview guides, appropriate statistical or content analysis, and adequate support for conclusions and recommendations, to name but a few.
- **Stakeholder involvement**: There should be adequate assurances that the relevant stakeholders have been consulted and involved in the evaluation effort. If the stakeholders are to trust the information, take ownership of the findings, and agree to incorporate what has been learned into ongoing and new policies, programmes, and project, they have to be included in the political process as active partners. Creating a façade of involvement, or denying involvement to stakeholders, are sure ways of generating hostility and resentment towards the evaluation.
- **Feedback and dissemination**: Sharing information in an appropriate, targeted, and timely fashion is essential.
- **Value for money**: Spend what is needed to gain the information desired, but no more. Gathering expensive data that will not be used is not appropriate – nor is using expensive strategies for data collection when less expensive means are available. The cost of the evaluation needs to be proportional to the overall cost of the initiative.

2.12 Sharing Knowledge and Lessons

Oxfam GB minimum standards

- All evaluations and management responses to be uploaded onto OPAL and sent to the Country Director, the Regional MEL lead and the global MEL team.
- Findings from evaluations should be shared with partners, affected populations and donors as appropriate.
- Findings and recommendations from mid-term evaluations should be fed into the next Monitoring Review. Lessons from the programme evaluation should inform future programme development

Evaluations are key sources of information about Oxfam GB's programmes and it is important to maximise their potential to improve understanding and practice across the organisation and, in some cases, the sector as a whole, by ensuring that they are shared beyond the programme with as wide an audience as possible.

Often completed evaluation reports are just shelved and forgotten about. This can be for all sorts of reasons – some are to do with the evaluation process (such as: the right people were not involved in the evaluation and so don't feel ownership of the findings; the evaluation didn't address the right questions; the evaluation was carried out at the wrong time and so wasn't available to feed into decisions), but others are to do with how it is communicated. The formal report will likely be appropriate for donors, but this may not be

---

[2] Kusek & Rist (2004), *Ten steps to a reults-based montoring and evaluation system*. Washington, DC: The World Bank

the best way to communicate findings to other stakeholders such as field staff, partners or beneficiaries. Each evaluation should have a clear strategy for communicating its results to the full range of stakeholders.

There are many different means that can be used to disseminate the evaluation. Some to consider include:
- Short written documents such as brochures or articles
- Videos
- Oral presentations
- Stories
- Short one or two page policy briefs
- Group dissemination meetings or other opportunities to present the findings to communities, or the general public
- Good practice case studies or guidelines

At a minimum, all evaluations should be uploaded onto OPAL and sent to the Country Director, the Regional MEL lead and the global MEL team. They will be made available to all Oxfam GB staff through a searchable evaluations database, and on our website and the website of key allies and networks[3].

3 Data Collection

Data can be collected in many ways, and no single way is the best way - each will have advantages and disadvantages. Choosing the right method involves being clear about what is to be measured, and choosing a data collection method that will measure it in an appropriate and meaningful way. To do this, it is important to be aware of the different types of information that different data collection methods provide, and understanding what is entailed in the collection of that information. For an overview of different data collection methods see the Data Collection Methods Chart.

There are a variety of data collection instruments or tools that can be developed and used with each collection methods. For example, common survey tools include Household Surveys; or Knowledge, Attitudes, Practices and Beliefs (KAPB) Surveys. Depending on the information that a programme team wants to collect, they may be able to adapt an existing instrument or they may have to develop one from scratch.

Participatory Learning and Action (PLA) exercises can also be used to support information-gathering discussions with beneficiaries. Most staff will be familiar with these exercises from programme identification, design and planning stages, but they can be equally helpful in monitoring and evaluation. For examples of PLA exercises, see IFRC PLA exercises and their possible uses.

A few general tips:
- **Collect the right information.** Most programme teams collect too much information that they can not use. Information should only be collected where it is essential for influencing decision-making, promoting accountability and/ or for demonstrating impact. Developing strong programme logic and selecting a small number of indicators against which to collect data is the best way to ensure that only the most relevant data is collected.
- **Collect the information at the right time** The starting point for data collection is a baseline that shows the situation pre-intervention. Data gathering intervals will then depend on the needs of the programme and its stakeholders.
- **Plan and budget for data collection in advance.** Collecting baseline data and carrying out evaluations require an investment of staff time and funds. Make sure this is in the

---

[3] Exceptions will be made where evaluations deal with issues of a sensitive nature and/or pose a risk to our staff or partners. Where this is the case, Country Directors should indicate in writing that the evaluation is not appropriate for external audiences and why. This should accompany the evaluation and the management response.

programme plans from the beginning - establish a firm schedule for data collection, make sure that annual workplans allocate time for monitoring reviews or ongoing data collection activities, and ensure the data collectors have both the training and the tools they need for collecting data

- **Allow time to test data collection methods.** No matter what type of data collection tools are used, they should be tested with a small sample of volunteers before being implemented. This allows time to ensure that the tools are easy to understand, whether the questions are clear, relevant, and unbiased, if the data collection could be completed in a reasonable amount of time, and whether there are any improvements that could be made to the data collection tool.
- **Manage the information that you collect.** Make sure data is carefully stored; accessible to those that need it; and easily understood. Also, it will be important to ensure that confidentiality is maintained and that information cannot be associated with a particular person unless they have agreed to being quoted. Programme teams should be aware of any national legislation on this issue (for example, the Data Protection Act details regulations for how personal data can be stored and used in the United Kingdom).

Collecting baseline data

An important thing to consider in the early stages of programme design is whether you will carry out a baseline study. A baseline provides a reference point - a description and measurement of initial conditions that are addressed by the programme – that allows the programme team to measure progress (or lack of it) towards intended programme outcomes. Collecting baseline data essentially means taking the first measurements against the programme's indicators to find out the situation at the beginning, or just prior to the beginning, of a development intervention.

3.1.1 Mixed methods approach to data collection

Data collection methods tend to be broadly characterised by whether they are used to collect quantitative (generally numerical data) or qualitative (generally more in-depth, descriptive data). Taking a mixed-methods approach, collecting both quantitative and qualitative data, usually results in a better understanding of the programme, and in a stronger evaluation.

Using mixed methods provides richer detail and gives a more in-depth understanding of the issues. It allows analysis on different levels. For example, a programme team may use a survey gives a good estimate of individual, household or community welfare, and then supplement this information by using focus groups to explore and analyse processes, structures, attitudes or behaviours that may contribute to the situation. Mixed methods are also valuable when there is a limited budget and significant time constraints because different sources of information can be used to cross-check or 'triangulate' information obtained.

Triangulation is the process of using two or more methods of data collection or sources of data in order to check on the validity of the data that is being gathered, and the resulting analysis.

3.1.2 Sampling

It is rarely possible, or practical, to collect data from the entire programme population. Instead, information is often collected from a sample of the group in question, which allows an educated estimate of characteristics of the population as a whole to be made. Sampling is used with both quantitative and qualitative data collection.

Samples can be generated in a random or non-random way.

- **Random Sampling** - involves choosing a sample group randomly from the entire population that you are looking at, with each unit in the population having an equal chance of being selected. The advantage of random sampling is that it removes

selection bias. Therefore, an appropriately sized random sample should be representative of the population as a whole. This enables generalisations to be made from the sample about the population from which the sample was drawn.

Random and non-random sampling methods can be combined. For example, communities may be selected in a non-random way, but then individuals within those communities may be selected at random. Or equally, a sample of households may be selected randomly, and a decision taken to purposefully speak to only the women of those households.

When collecting information from a sample of people, rather than the entire population, the sample is used to make inferences about a larger population, and it is therefore important to understand how representative the sample is. This is where statistics come in. Statistical analysis can help us to know how valid the data is, and how probable it is that another randomly selected sample of the same population would give the same results.

Identifying appropriate samples can be quite complicated, and statisticians have developed theories and formulas for making these estimates and selecting appropriate sample sizes. If programme teams are unsure how to proceed, they should contact the country or regional MEL lead, or be in touch with the global MEL team.   In some situations programme teams may want to hire the services of a professional statistician.

3.1.3 Data Analysis

Once the data are collected, the evaluation team will need to go through it and find meaning in the words and numbers. Data collection and data analysis tend to be very linked activities, with analysis happening in real time and informing the next phase of data collection. This is particularly true of qualitative data, where interviewers or focus group leaders will be analysing the responses they are receiving and tailoring the direction of the interview or discussion accordingly.

Analysing data involves organising, interpreting and questioning the data, finding causal links, making inferences, attaching meanings, and dealing with cases that contradict the analysis. This might happen through an existing framework, or by allowing patterns, themes, and categories in the data to emerge. The processes used to analyse data should be documented, so that others are able to understand how conclusions were reached, and can verify or reproduce the process if need be.

There are many different techniques to assist with the sorting and analysis of data, some more appropriate for qualitative data, others more appropriate for quantitative data.

Qualitative Data Analysis
Qualitative data analysis is used for non-numerical data. Qualitative data often involves a combination of collection and analysis– watching for patterns or themes to emerge as data is collected and responding to them by delving into those issues in more depth. Qualitative data can give a powerful sense of the key issues for different stakeholder groups. It helps to identify key themes and trends, and the relative strength of feeling around particular issues.

Because it deals with more descriptive data - such as perceptions, opinions or feelings - analysing qualitative data can be very challenging, and care needs to be taken to accurately capture and interpret qualitative data. In particular, it is important to be careful not allow first impressions or personal biases to limit the different possible interpretations of what the data means. In most situations, there will several possible interpretations of what qualitative data means, and it will be important to be open to them all and to strive to find the most probably meaning, rather than the one that sits most comfortably with pre-conceived ideas.

Analysing Quantitative Data

Quantitative data are numerical and are analyzed using statistics, which involves grouping the data and looking for general patterns or trends. Statistics can be divided into two large categories: Descriptive statistics and Inferential Statistics.

Descriptive statistics describe the distribution of data by frequency or proportion – describing how many, or what percent (for example, 40% of participants were male, 60% were female). They may show how similar the data is (for example, most participants were female), or how different the data is (for example, the ratio of men to women was 2:3).

Inferential statistics are typically used to analyse random sample data. They allow an estimate to be made about a population based on a random sample selected from that population. As such, they should include a reliability statement that state how probable it is that the random sample represents the true population.