



Save the Children International (previously called The International Save the Children Alliance Charity) is a company limited by guarantee, registered in England and Wales with Company No. 3732267 and Registered Charity No. 1076822. Its registered office is St Vincent House, 30 Orange Street, London, WC2H 7HH. A list of directors is available on request.

Contents

1	Introduction	4
1.1	What is evaluation?.....	4
1.2	Why do we need evaluations?	4
1.3	Who is this handbook for?	5
1.4	The handbook’s structure.....	6
2	Standards and principles for Save the Children International evaluations	8
2.1	When should you carry out an evaluation?	8
2.2	Enabling children and young people to participate	9
2.3	Quality standards and evaluation criteria.....	10
3	Designing your evaluation.....	12
3.1	Understanding the context.....	12
3.2	What do we mean by evaluation design?	12
3.3	What questions do you need to ask?.....	13
3.4	When should you select the evaluation design?	16
3.5	Choosing your evaluation design.....	16
3.6	Choosing your data collection methods and tools.....	20
3.7	Choosing your sample	22
3.8	Ensuring the quality, validity and reliability of data.....	24
4	Planning your evaluation.....	27
4.1	Drawing up the terms of reference.....	27
4.2	Involving children and other stakeholders	28
4.3	Setting your budget.....	29
4.4	Agreeing a timeline	29
4.5	Getting the right evaluation team	29
4.6	Finding the right consultant to lead your evaluation team.....	32
4.7	Managing roles and responsibilities.....	32
4.8	Managing external consultants	34
5	Implementing your evaluation.....	36
5.1	Fine-tuning your evaluation plan.....	36
5.2	Facilitating children and young people’s participation	37
5.3	Designing and testing data collection tools.....	38
5.4	Training.....	39
5.5	Managing data.....	39
5.6	Writing the evaluation report.....	42
5.7	Reviewing the draft report	42
5.8	Approval procedures	44
5.9	Planning for action and communication.....	45

6	Using the results.....	47
6.1	Preparing a management response to the evaluation report.....	47
6.2	Developing an action plan.....	48
6.3	Sharing the findings with external audiences and stakeholders	49
6.4	Sharing the findings with internal audiences.....	50
6.5	Using your findings to improve programme strategy and implementation.....	51
	ANNEX 1: DAC evaluation criteria.....	53
	ANNEX 2: What you need to know about sampling.....	55
	ANNEX 3: A template for drawing up your terms of reference.....	67
	ANNEX 4: Quality review guide.....	69
	ANNEX 5: A template for your management response plan.....	71
	ANNEX 6: How to communicate statistics	72
	ANNEX 7: How to involve children and young people in evaluations.....	76
	ANNEX 8: Gender equality audit.....	91
	ANNEX 9: The contract manager’s responsibilities	95

Acknowledgements

This handbook has been developed by the Core Process Harmonisation Monitoring and Evaluation Design Group, co-chaired by Helen Cheney, Save the Children Australia and Jo Feather, Save the Children UK.

We are grateful to the following people who peer reviewed the handbook: Joanna Arulraj from Save the Children in Sri Lanka, Monica Zikusooka from Save the Children in Somalia Somaliland, Suzanne Wessels from Save the Children in South Africa, Tarun Adhikari from Save the Children in Nepal, Larry Dershem from Save the Children USA (Middle East/Eurasia Region), Yosef Gebrehiwot from Save the Children USA (Africa Region) and Simon Early from Plan. We would like to thank Katy Sklan, Save the Children UK who helped compile Annex 7. We would also like to thank many other colleagues who took the time to review and comment on the handbook, including those from Save the Children Australia, Save the Children Denmark, Save the Children Finland, Save the Children International, Save the Children Japan, Save the Children Norway, Save the Children Sweden, Save the Children Spain, Save the Children UK and Save the Children USA.

The handbook was edited by Kathryn O'Neill, proofread by Kathleen Richardson and designed by Officentral Desktop and Online Publishing.

Illustrations are by Jess Stockham and were originally published in Save the Children's *Advocacy Matters: Helping Children Change their World*.

Cover image: Muskan (age 10), taking a photo with a plastic camera, India. Photo: Rachel Palmer/Save the Children.

Design group members:

Ayano Suzumura, Save the Children Japan
Barbara Rieckhoff, HIV/AIDS Global Initiative
Brynjar Sagatun Nilsen, Child Protection Initiative/Child Rights Governance
Hana Crowe, Save the Children International
Helen Cheney, Save the Children Australia
Ingunn Tysse Nakkim, Save the Children Norway
Jo Feather, Save the Children UK
Luc Vanhoorickx, Save the Children International (Asia Regional Office)
Marianne Bo Paludan, Save the Children Denmark/Child Rights Governance
Marta Arranz, Save the Children Spain
Mulu Chekol, Save the Children USA
Nicola Chevis, EVERY ONE Campaign
Nitika Tolani Brown, Education Global Initiative
Sarah Lilley, Child Protection Initiative
Tove Stromberg, Save the Children Sweden

I Introduction

I.1 What is evaluation?

An evaluation is the systematic and objective assessment of an ongoing or completed project, programme or policy and its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, development efficiency, effectiveness, impact and sustainability.

Evaluation is different from monitoring, which is a continuing assessment based on systematic collection of data on specific indicators as well as other relevant information. Monitoring enables managers and other stakeholders to measure progress in terms of achieving the stated objectives and the use of allocated funds.

Monitoring and evaluation (M&E) is a key part of realising Save the Children's theory of change and common values and strategies, inherent in the child rights programming (CRP) framework. The principles, rights and obligations set out in the United Nations Convention on the Rights of the Child 1989 (UNCRC)¹ provide a fundamental framework for the work we carry out with children and young people around the world. All of Save the Children's programme and advocacy work should aim to address violations of children's rights and gaps in service provision, as well as supporting children as rights-holders and helping states, as duty-bearers, to meet their obligations. Our vision, mission, values and theory of change² reinforce this. It is vital that we clearly **articulate**, **demonstrate** and **document** the outcomes of our work for girls and boys and their carers.

The work of Save the Children and our partners often has immediate benefits for children and young people. It is important that we assess the longer-term outcomes and impact at local, national, regional and global levels. Assessing the quality of our programmes helps us to ensure that we are constantly learning from and improving what we do. This includes rigorous monitoring and evaluation, so that we can feed back lessons learned to improve the design and delivery of the project or programme. Evidence will be routinely collected during project/programme monitoring, including against the agreed set of Global Indicators and, where appropriate, should inform the evaluation process. We use evaluations to learn about our interventions, whether direct interventions providing services for children, or policy and advocacy work to advocate for better practices and policies to fulfil children's rights.

This handbook focuses on the key role evaluations play in gathering evidence to inform this learning cycle. It explains Save the Children's evaluation requirements and sets out good practice to help you improve the evaluation process for your project or programme.

I.2 Why do we need evaluations?

Evaluations can help us to:

- Identify the outcomes and impact³ of our work (at project, programme, country, regional or global level)

¹ See UNICEF, 'Convention on the Rights of the Child', www.unicef.org/crc/

² Save the Children's theory of change shows three major elements of our common identity: striving to create impact for children by being innovative, by acting as the voice for and of children and by achieving results at scale. For more information, see: www.savethechildren.net/xtranet/alliance/one_save_the_children/One_Save_the_Children-English.pdf

³ Save the Children uses the following definition of impact: the totality of effects produced by an intervention, whether they be positive or negative, intended or unintended, direct or indirect, primary or secondary. Note: in emergency interventions, effects may be assessed in terms of outcomes rather than impact.

- Determine what helps and/or hinders our efforts to improve the quality and effectiveness of our programmes and projects
- Identify value for money and ensure that funds are used efficiently and effectively
- Strengthen a learning culture through reinforcing the links between evaluation findings and programme design and decision-making
- Validate our contribution to the key components of the theory of change
- Ensure accountability and transparency to our stakeholders, including children and their carers
- Provide robust evidence for policy and decision-making and for our advocacy, fundraising and external communications about the impact and value of our work
- Improve future project design and management, not just by proving impact but also by improving practice and by developing and documenting replicable, innovative solutions to the problems facing children.

How evaluations can support our Theory of Change?

Evaluations give us the evidence to prove that our innovative approaches to addressing children's rights bring results. In many cases, evaluations provide evidence of best practice to support leverage and scale-up of our interventions. They can help us to learn about the effectiveness of our partnerships too.

1.3 Who is this handbook for?

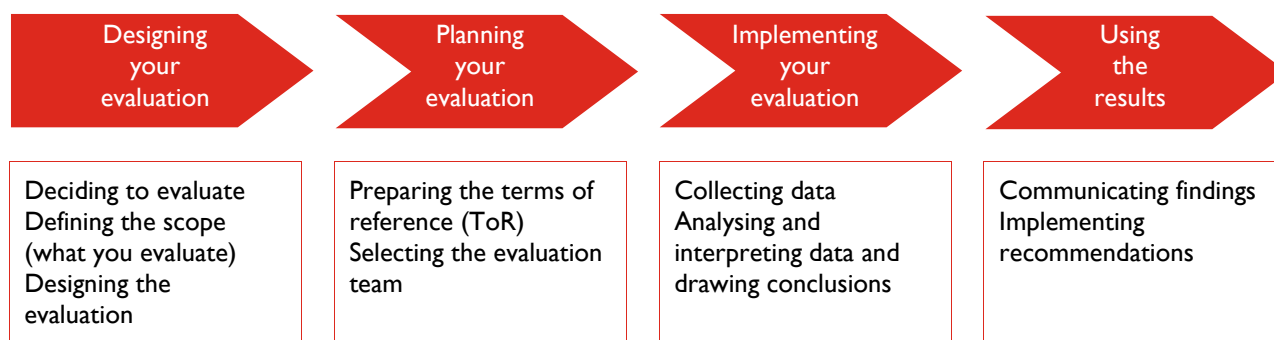
This handbook has been developed for country offices to provide a step-by-step, practical guide to the evaluation process. Its purpose is to help managers, teams and technical specialists to ensure that evaluations are systematically designed and implemented to a high standard. This includes where Save the Children is conducting internal evaluations or commissioning and managing external evaluations. It also aims to ensure that learning is fed back into our programmes in a systematic way. It sets out the Save the Children International (SCI) Management Operating Standards (MOS) for evaluations and follows the Standard Operating Procedures (SOP) on evaluations. It will give teams an oversight of SCI standards and principles for designing, planning and implementing evaluations, as well as guidance on how to communicate the findings and follow up on recommendations made.

We have kept the level of technical detail to a minimum, but there is valuable information in the Annexes. We want country offices to use the handbook as a reference document, so we have included guidance on best practice as well as useful tools, methods and resources for carrying out evaluations. It is important that country offices source technical assistance as required at any point during the evaluation process. In the first instance, contact your Regional M&E focal point, who will be able to source appropriate technical support for your country office.

1.4 The handbook's structure

Though evaluation is not a linear process, the chapters in this handbook follow the phases outlined in Figure 1 below.

Figure 1: The evaluation process



Chapter 2 sets out SCI's requirements for evaluations as defined in the Management Operating Standards (MOS) and Standard Operating Procedures (SOP). It also outlines the ethical standards and participation practice standards, particularly regarding children's participation – one of our fundamental ways of working.

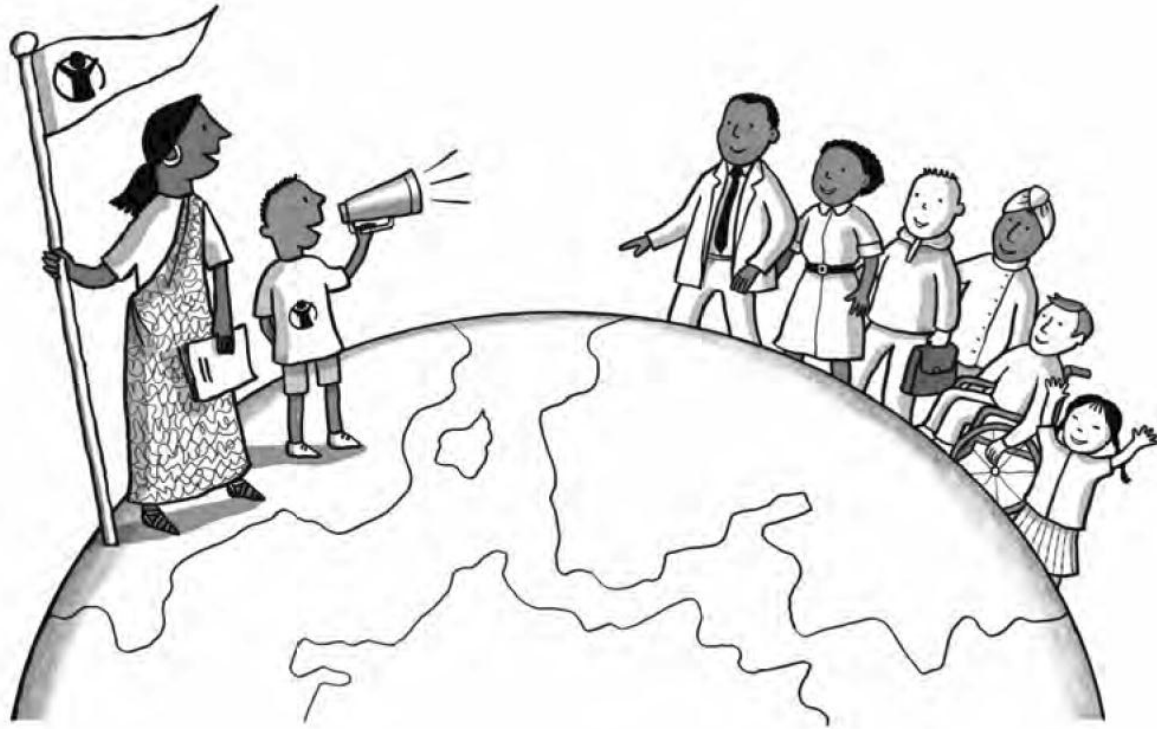
Chapters 3 to 6 comprise the step-by-step guide. Chapter 3 takes you through how to design the evaluation, from clarifying its purpose and defining the scope, objectives and questions, to deciding the design and methodology. It will help you consider issues like the level of rigour needed, data sources, sampling and data collection and data quality.

Chapter 4 explains how to plan your evaluation. It covers how to involve stakeholders, preparing a budget and timeline and forming the evaluation team. It also takes you through the management process, which includes defining roles and responsibilities for the evaluation team and preparing the terms of reference (ToR) and contracts.

Chapter 5 takes you through the implementation phase of the evaluation. It starts with the inception phase, when you will need to carry out detailed planning. It gives practical examples of how to involve children, as well as how to design and test the data collection tools. It covers providing training for those involved in the evaluation and how to manage data, including collection, input, cleaning, analysis and storage. There is a section on writing the evaluation report and how to review and get feedback. It also outlines the relevant approval procedures.

Chapter 6 focuses on how to communicate and disseminate your evaluation findings and how to follow up any recommendations. It will help you to prepare the organisational response to the evaluation. It also shows you how to develop an action plan to ensure that learning informs future project or programme planning, design and implementation – in your own team and country office and across the organisation.

Finally, we have included a number of tools and templates as annexes. There are templates to help you draw up the terms of reference and the management response plan, some guidance on data sampling, a more detailed outline of evaluation criteria, information on how to conduct a gender equality audit, advice on how to communicate statistics, guidance for contract managers and resources to help you make sure that children are meaningfully involved in the evaluation process.



2 Standards and principles for Save the Children International evaluations

This chapter sets out the standards that all SCI evaluations should adhere to. They are included in the SCI Management Operating Standards (MOS) and Standard Operating Procedures (SOP).

2.1 When should you carry out an evaluation?

The decision to evaluate your project, programme or thematic area should be in line with the thresholds set out in Save the Children’s Management Operating Standards (see box below). You should check donor requirements too, as these may mean bringing the thresholds forward. See the Evaluation Standard Operating Procedures for examples of who might initiate and/or commission an evaluation. The entity that commissions the evaluation is responsible for oversight of decisions and approvals at all stages (see Chapter 4 for more details).

SCI evaluation thresholds*

Programme/project length and budget	Mid-term evaluation	Final evaluation
Less than 1 year Less than \$1 million	Not required**	Not required**
Less than 1 year More than \$1 million	Not required	Required Internal/external
Between 1 and 3 years All budgets	Not required	Required Internal/external
4 years and over All budgets	Required	Required External recommended

* Evaluation thresholds do not apply to emergency response programmes

** Final project report is still required for smaller projects

In line with the Management Operating Standards, all innovative projects/programmes and those of strategic importance should be evaluated at the end of the intervention period, regardless of the budget. In declared emergencies, a real-time evaluation and Evaluation of Humanitarian Action must be conducted in accordance with the Emergency Rules and Principles.⁴

Aside from the thresholds set out above, there may be other reasons why you want to evaluate a project or programme. It may be that you are considering starting work with new partners, implementing, replicating or scaling up a pilot project, or conducting high-risk activities or complex interventions with multiple partners and/or donors. You may also want to evaluate across projects and programmes to assess synergies or, alternatively, conflicting objectives.

⁴ See Monitoring and Evaluation document library on the SCI Xtranet

2.2 Enabling children and young people to participate

As an organisation, our work is underpinned by the United Nations Convention on the Rights of the Child (UNCRC). This guiding framework sets out a number of key reasons why children and young people must be meaningfully involved in the evaluations we carry out. Children and young people have an inherent right to participate in evaluation activities because the project or programme involves “matters affecting the child” (Article 12).

It is crucial that you enable stakeholders of all ages to participate in the evaluation. (A stakeholder is any person or organisation that may have an interest in your project or programme, whether they are external or internal.) You also need to think about how children and young people – our key stakeholders – can participate meaningfully at all stages of your evaluation process. This handbook gives lots of examples of how you can do this and you will find many examples of creative, age-appropriate ways to involve children and young people in Annex 7. Getting a wide range of stakeholders to participate in the evaluation builds ownership of the process and promotes accountability. It can also help to build local capacity and ensures that learning is comprehensive and can really be used to improve activities, systems and processes.

A number of tools are available to help you identify your stakeholders (you can find examples in the *SCI M&E Handouts Package Volume 1*). You need to think systematically about who should be involved in your evaluation so that you avoid any unintentional bias or discrimination. Look back at your child rights situation analysis (CRSA) to help you. Try to include a fair representation of children and young people (girls and boys) and adults involved in the programme. Make sure that children from different ethnic, social and religious backgrounds have the chance to participate, as well as children with disabilities and children who may be excluded or discriminated against in their community.

SCI country offices must ensure that children and young people are meaningfully involved in monitoring and evaluation activities, to comply with the Management Operating Standards, Standard Operating Procedures, Ethical Standards, Child Safeguarding Principles and Practice Standards in Children’s Participation (see box below).

The Practice Standards in Children’s Participation

Standard 1 – An ethical approach: transparency, honesty and accountability

Standard 2 – Children’s participation is relevant and voluntary

Standard 3 – A child-friendly, enabling environment

Standard 4 – Equality of opportunity

Standard 5 – Staff are effective and confident

Standard 6 – Participation promotes the safety and protection of children

Standard 7 – Ensuring follow-up and evaluation.

Each standard sets out criteria you can use to assess whether the standard is being adhered to. Standards 1, 2 and 7 particularly emphasise the fact that children should be involved in monitoring and evaluation as a holistic process and not only as informants. The Child Participation Global Indicator is also a useful tool that will help you assess the scope and quality of children's participation in your project or programme. See *Practice Standards in Children's Participation* (International Save the Children Alliance 2005); and Global Indicator technical guidance (*SCI M&E handouts Package, Volume 2*).⁵

Safeguarding children

Children's safety and well-being should always be paramount. Children and young people should not be exposed to any risks, whether psychological (such as shame or fear), social (threatening their family/community relations), political (discrimination, threats or violence due to positions or acts such as being affected by armed forces or groups) or physical (abuse or violence). Evaluation activities should provide a safe, creative space where children feel that their thoughts and ideas are important. All staff, partners and consultants should adhere to Save the Children's Child Safeguarding Policy and Code of Conduct.⁶ Managers should ensure that training and all other necessary steps are taken to identify risks to children's safety and well-being and to minimise and manage those risks.

Staff will need to be prepared to support a young person if he or she chooses to disclose abuse or report a concern as part of an evaluation activity or discussion. Staff should consider the potential for this kind of disclosure – especially if the content is sensitive. There are a number of ways you can safeguard children taking part in evaluation activities, including:

- Drawing up a contract with the children and young people involved regarding confidentiality and respect at the beginning of the meeting and make sure they know who to go to if they wish to disclose something in private
- Carrying out a risk assessment that includes any ethical risks related to children or young people's participation.

2.3 Quality standards and evaluation criteria

There is a high degree of international consensus with respect to the criteria to be applied in evaluations. Bilateral aid programmes are largely evaluated using the five criteria set out by the Development Assistance Committee (DAC) of the Organisation for Economic Co-operation and Development (OECD) (see Annex I).

These criteria are:

- Impact
- Effectiveness
- Sustainability
- Relevance
- Efficiency.

⁵ See Monitoring and Evaluation document library on the SCI Xtranet

⁶ See Monitoring and Evaluation document library on the SCI Xtranet

Taken together, these criteria will cover critical areas of most interventions, but not all of them are equally applicable to all evaluations. Which ones you need to apply will depend very much on the purpose of your evaluation.

Other criteria might also be appropriate for specific evaluations. Save the Children's theory of change provides another set of criteria you can use – for example, if you are considering replicating or scaling up certain activities, or assessing some partnerships in more depth. Or you could use the principles of the UNCRC as the basis on which to assess impact or outcomes, using criteria such as the best interests of the child, participation and non-discrimination.⁷ The five dimensions of change⁸ could also serve as useful criteria in some evaluations.

Evaluations of humanitarian interventions should also assess coverage, coordination, coherence and protection, as well as performance against accountability standards (for example, the Sphere Project's *Humanitarian Charter and Minimum Standards in Humanitarian Response* and the Humanitarian Accountability Partnership (HAP)'s *2007 Standard in Humanitarian Accountability and Quality Management*).⁹

In Chapter 3, we show you how to design your evaluation and decide what methods to use.

Checklist – Evaluation standards and principles

Deciding to evaluate

- ✓ Does the project or programme need to be evaluated according to Save the Children International's MOS thresholds in relation to timescales and budgets?
- ✓ If so, what type of evaluation is required?
- ✓ Is the project or programme innovative or of strategic importance?
- ✓ Is it an emergency response, or are there other reasons to carry out an evaluation?
- ✓ Have you chosen the criteria you want to use for the evaluation?
- ✓ Have you thought about how you can ensure the meaningful participation of children and young people in the evaluation activities?

⁷ For more information, see *Getting it Right for Children: A practitioners' guide to child rights programming*, International Save the Children Alliance, 2007. Available at: <https://sites.google.com/a/savethechildren.org/m-e-training/training-materials/other-materials/GettingitRight.pdf?attredirects=0&d=1>

⁸ *Getting it Right for Children*, Chapter 2

⁹ For more information, see The Sphere Project's website: www.sphereproject.org/. See also HAP International's website: www.hapinternational.org/

3 Designing your evaluation

3.1 Understanding the context

This chapter looks at some of the ways you might structure your evaluation and explores how to choose the evaluation design that best meets your needs.¹⁰ In order to establish the purpose of your evaluation, you must first consider these key questions:

- What are you trying to find out? Are you looking for evidence of any changes brought about by the project or programme? Or do you want to have a better understanding of the process and what works? Do you want to draw out lessons learned? Or assess the extent to which children, young people and communities gave their views and feedback?
- How do you define and measure progress and success? What would constitute positive impact or change for the better in your context?
- How do you know what elements of the programme need to change in order to increase its effectiveness and relevance?
- If you do identify positive change, how do you know if it was caused by Save the Children's programme or something else (the attribution question)? And how is it linked to other interventions, structures or broader changes in society?
- Even if it seems the project or programme is having positive results, what can you improve on and what elements can be changed or removed without affecting the programme's success?
- Who will be interested in the findings of your evaluation? (Consider the needs of internal and external stakeholders.)

We begin by providing examples of questions your evaluation should ask. We then provide a list of factors to consider when choosing the right design for your evaluation, ranging from simple factors to more complex issues facing country offices. We describe the most commonly used evaluation designs. Next, we provide you with a simple decision tree that will help you choose the right evaluation design, based upon your evaluation questions, objectives and resources. Finally, we review quantitative and qualitative data collection methods and tools, including sampling techniques and help you to ensure the quality, validity and reliability of the data you collect.

3.2 What do we mean by evaluation design?

Every evaluation is essentially a learning activity. It may be about determining how effective and/or relevant your project or programme has been overall, which parts of it are working well or have worked well, or whether some participants respond to certain methods or conditions differently from others. If your evaluation findings are to be reliable, you have to use a structure that will tell you what you want to know. That structure is the evaluation design.

¹⁰ This chapter draws from M Bamberger, J Rugh and L Mabry, *RealWorld Evaluation: Working under budget, time, data and political constraints*, Sage, 2006; also D Schober and C Holt (eds), *The Community Tool Box*, Chapter 37, Section 4, 'Selecting an appropriate design for the evaluation', available at: http://ctb.ku.edu/en/tablecontents/chapter37_section4_main.aspx

Avoiding bias

To avoid bias, as far as possible, you need to be systematic:

- Reflect stakeholders' information needs in your evaluation questions.
- Use data collection techniques that will enable you to collect all the data you need.
- Follow analysis steps: data analysis, interpretation, conclusions and recommendations.

In the evaluation design, you need to make a number of deliberate choices, including what to evaluate (and what not to), which techniques you will use to collect data and who you are going to involve in the process. The design should also include the “level of rigor or precision” needed to answer the evaluation questions, which will be determined by time, costs, context and capacity. These choices bring some degree of subjectivity to the evaluation, so it is important to follow the highest standards on ethics, stakeholder participation and transparency.

3.3 What questions do you need to ask?



The evaluation design depends most importantly on what kinds of questions your evaluation needs to answer. Some of the most common evaluation questions are set out in Table I below.

Table 1: Common evaluation questions using criteria from the DAC, Save the Children's theory of change and child rights programming (CRP) framework

Effectiveness

- ✓ Did Save the Children and/or partners implement the project or programme as planned and if not, why not? Did a particular project or programme cause a particular change in children's lives or the lives of others, whether in terms of physical or social conditions, health, education, nutrition or protection outcomes, or other indicators of success? How significant was this change and how did it happen?
- ✓ What component(s) and element(s) of the project or programme were responsible for the change?
- ✓ If your project or programme tried a new methodology or approach, what was the result? What lessons were learned and what recommendations made?
- ✓ On what basis was the beneficiary population selected? Were there other demographic groups that might have been included to give the project more impact?
- ✓ Were the partnerships appropriate? Have they been managed effectively?

Efficiency

- ✓ How cost-effective is the project or programme?
- ✓ Could the same results have been achieved with fewer resources?
- ✓ How has knowledge been shared?

Relevance

- ✓ Was the project or programme appropriate for the context where it was implemented?
- ✓ To what extent has it taken into account people's different needs according to age, gender, ethnicity and other social identities? How has the project or programme adapted to meet those differing needs?
- ✓ How has the project or programme ensured that children's voices are heard and reflected, both in project activities and more broadly, in our interaction with governments and other stakeholders?

Sustainability

- ✓ Will the changes caused by the project or programme continue beyond the life of the programme?
- ✓ Has the project or programme resulted in any leveraging of our knowledge and interventions to ensure sustainable impact for children at scale?
- ✓ How has the project or programme worked with local partners to increase their capacity in a sustainable way?

Impact

- ✓ What are the project or programme's unintended effects and how did they influence the outcomes?

Replicability and Scale

- ✓ Will the project or programme work in a different context? What would happen if we scale up in one context rather than another?
- ✓ Have we developed and or proved new, evidence-based and replicable breakthrough solutions to problems facing children, working with others and sharing these effectively to ensure greatest impact for children?

Coverage and Non-discrimination

- ✓ Have all target groups been reached?
- ✓ Have there been any constraints in terms of access to the services or goods provided through the project or programme?
- ✓ Has our work resulted in more equitable and non-discriminatory projects, programmes and/or policies for children and young people at *national, regional and global* levels?



Encouraging children to participate

Children, young people and community members should be invited to participate in the research and to develop evaluation questions. Country office staff could convene an evaluation design workshop with a representative sample of children and young people to help design the questions. See Annex 7 for more creative ideas about how to involve girls and boys at different stages of your evaluation.

Looking only (or mainly) at results and whether you have achieved the project or programme's objectives offers a limited perspective and one that does not allow lessons to be drawn out. There are, however, other approaches and evaluation models that are more comprehensive. They ask not just *what* has been achieved (or has not), but *how*. They include analysis of elements such as project/ programme design, processes and structure and should also consider SCI's theory of change.

3.4 When should you select the evaluation design?

Ideally, you should decide the evaluation design at the very beginning, when you plan a project or programme. This means you can include the evaluation in all relevant proposals, plans and budgets. It also allows you to collect baseline data and to monitor any changes over the life of your project. All this data can be used in the evaluation.

Having a clear monitoring and evaluation (M&E) plan from the outset will help to ensure this occurs (see Design, Monitoring and Evaluation (DM&E) Standard Operating Procedures/ Management Operating Standards).

However, if this is not the case and you are planning an evaluation for a project /programme that has been running for some time – maybe months or even years – you still need to choose the evaluation design carefully, especially since you may not have been collecting all the required, relevant data during the project.

In some cases, you may decide and plan the evaluation during the implementation of the project or programme – for example, when monitoring data reveals a need to understand results in more depth or find out why something is not working. It might also be that you need to decide whether to continue, exit or scale up the project or programme, or increase accountability to stakeholders and/or donors and an evaluation is the best way to help you do so.

3.5 Choosing your evaluation design

There are a number of different types of evaluation design, but no one design is ideal. The right one for your situation depends on the purpose of your evaluation, its timing, where you are up to in terms of the implementation phase, what data you already have, the level of rigor required (eg, testing a pilot model or an innovation) and your available resources (money, time, etc).

When choosing the appropriate evaluation design for your project or programme, you should consider these points:

- *The nature of the evaluation questions you are trying to answer.* If you are trying to answer questions about whether an intervention led to changes in children and young people's outcomes and the scale of that change, you will need a more complex evaluation design. If you are trying to answer questions about whether activities were implemented as planned and on time, a simpler design may be more appropriate.
- *The theoretical or conceptual basis (theory of change or project logic model) for the outcomes you want to study.* It is helpful to conduct a literature review when you begin planning your evaluation. This will help you understand the theoretical basis for the relationships between the outcomes you are interested in and the interventions you are implementing. A literature review may also help you choose the best evaluation design.
- *The audience for the evaluation and the way you want to use the findings.* If you are trying to bring about a change in government policy through the evaluation findings, you will need information that is robust, reliable and evidence based so that policy-makers take it seriously. This means you may need a more complex evaluation design. If you are conducting an internal review to draw out lessons learned, you may be able to use a simpler design.

- *The stage in the project cycle that you are planning the evaluation.* If you are planning the evaluation at the beginning of a project or programme, you can use a more complex design, as you will be able to collect baseline and follow-up data. If you are planning it at the end of the project, you may be limited to using simpler designs, depending on the data you have access to.
- *Ethical considerations.* If you judge that it is ethically sound to do so, you may use control or comparison groups in your evaluation design, or, for example, methods of collecting data from young children. If you or other stakeholders have ethical issues with using control or comparison groups or using particular data collection methods, a simpler design may be more appropriate. Ethical considerations may also help determine the kinds of questions you can ask and who you can ask.
- *Budget and time constraints.* If you have a large budget and lots of time, you will be able to implement more complex evaluation designs. If you have a small budget and little time, you will need to use a simple design.

Programming that focuses on advocacy work, changing systems and policies and strengthening civil society is more likely to use qualitative methods and look at broader outcomes at regional and national levels. If you are evaluating advocacy or policy work, the evaluation design should include an assessment of the policy situation and government position at the start of the project or programme, so that any developments or changes can be identified and, if possible, attributed.

The most common types of evaluation design, including their advantages and limitations, are summarised below.

End-of-project analysis (with or without a comparison group):¹¹ This is the simplest evaluation design. It can tell you how the project group perceives the project and how they feel it has changed or improved their lives. It is normally used when the project or programme has no access to baseline data to measure outcomes or impact. But it is probably the least accurate or desirable evaluation design, because it cannot tell you, in a measurable way, the degree of change that has occurred or what caused the change. Using a comparison group can help you to document any differences between the project and comparison groups. But the evaluation design will not allow you to find out *what* caused those differences (that is, whether any changes can be attributed to the project or programme).

Baseline and end-of-project /programme analysis: This evaluation design is applied to evaluate the process of implementation, outcomes or impacts and is planned at the beginning of the project or programme. It assumes that differences at the beginning and the end of the project will be observed and that any positive or negative change was caused by the intervention. It can help you determine whether certain things have happened, but it will not tell you why. It does not account for the influence of other, external factors that might have affected children and young people's lives. Nor does it tell you anything about changing trends or the progress of change during the evaluation period – it simply gives you a snapshot about project beneficiaries' situation at the beginning and at the end of the project or programme.

¹¹ The use of comparison groups is usually not appropriate for this type of programming, unless you want to compare attitudes and beliefs between cohorts. However, it is still important to use an evaluation design that, at a minimum, captures baseline and end-of-project information and can provide credible evidence for Save the Children's likely contribution to any advocacy or policy change outcome. For example, in advocacy and policy change work, the design should include an assessment of the policy situation and government position at the start of the programme, so that any developments or changes that occur during the programme cycle can be easily identified.

Baseline and end-of-project analysis of project/programme and comparison groups: This evaluation design can use one of two methods for determining the comparison group.

- 1 **Quasi-experimental design:** In this type of design, project and comparison groups are selected non-randomly (eg, phasing a programme in different areas, using matched pairs, identifying groups with similar characteristics, or other quasi-experimental techniques). This is a more complex evaluation design, but it is flexible and can be used in a range of situations. Because there is both (i) baseline and end-of-project data and (ii) a comparison group, the design provides reasonably good estimates of the scale of the changes caused by the project or programme. The comparison group is usually similar to the project group, but either receives no intervention at all, or receives a different intervention.

You should be aware that there are important ethical considerations for involving comparison groups. The challenge with this type of evaluation design is finding a satisfactory comparison group. Two groups may appear to have similar characteristics, but may differ in important ways. For example, two groups of children may have similar characteristics and histories, but if one group has higher levels of motivation – which is not easy to assess from the outside – the results are not likely to be comparable. The extent to which the project group is motivated and determined, rather than the effectiveness of the programme, may influence the amount of change observed.

- 2 **Randomised controlled trial (RCT):** This is a complex evaluation design, but it allows you to statistically control for sample selection bias (eg, when results are distorted in a positive or negative direction), as participants are randomly assigned to the project and comparison groups. This design should provide reliable and precise estimates of the scale of the changes caused by the project or programme.

The challenge with this design is that it may not be ethically or practically appropriate to randomly assign participants to our programmes. Also, this design is likely to be more costly and resource intensive, given the need to collect baseline and endpoint data for both project and comparison groups.

Comprehensive longitudinal design with baseline, mid-term, end-of-project and project follow-up (eg, one year after the project has ended): This is a complex but strong evaluation design that studies the implementation process, outcomes and sustainability. It may be suitable if you are testing a new project innovation that, if you can prove change in children's outcomes, could be scaled up. However, it is likely to be much more costly and resource intensive than some of the other types of design.

Table 2 (*below*) presents a simple decision matrix to help you choose the right evaluation design.

Table 2: Choosing the right evaluation design

Type of evaluation (please refer to the M&E glossary of terms for definitions)	Where are you in the project timeline?			Use of control group		Suggested evaluation design
	Start	Mid-term	End of Project	Yes	No	
Process evaluation: Were the project activities carried out as planned? What was the implementation process?	x				x	Baseline and end-of-project analysis of project group, includes real-time evaluations (no comparison group)
		x			x	End-of-project analysis of project group, includes real-time evaluations (no baseline data, no comparison group)
			x		x	End-of-project analysis of project group, includes real-time evaluations (no baseline data, no comparison group)
Outcome evaluation: What are the likely or achieved short-term and medium-term effects of one or more interventions' outputs?	x			x		Baseline and end-of-project analysis of project and comparison groups, where both groups are selected non-randomly (eg, phasing programme in different areas, matched pairs, identifying groups of similar characteristics, or other quasi-experimental techniques)
				x		Baseline and end-of-project analysis of project group, includes real-time evaluations (no comparison group)
		x		x		End-of-project analysis of project and comparison groups (no baseline data)
					x	End-of-project analysis of project group, includes real-time evaluations (no baseline data, no comparison group)

Type of evaluation (please refer to the M&E glossary of terms for definitions)	Where are you in the project timeline?			Use of control group		Suggested evaluation design
	Start	Mid-term	End of Project	Yes	No	
Impact evaluation: ¹² What are the positive and negative,	x			x		Baseline and end-of-project analysis of project and comparison groups, where both groups are selected randomly (eg, randomised control trial (RCT))

¹² Impact evaluations are interested in effects caused by a development intervention – that is, *attribution*. This means going beyond describing what has happened to look at why it happened. Evaluating impact will, therefore, often require a counterfactual, or an assessment of the effects the development intervention has had, compared with what would have happened had the intervention not taken place (Evaluating Development Impacts – An Overview) (http://www.oecd.org/document/17/0,3746,en_21571361_34047972_46179985_1_1_1_1,00.html)

Type of evaluation (please refer to the M&E glossary of terms for definitions)	Where are you in the project timeline?			Use of control group		Suggested evaluation design
	Start	Mid- term	End of Project	Yes	No	
intended and unintended, direct and indirect, primary and secondary effects produced by the intervention?	x					Comprehensive longitudinal design with baseline, mid-term, end-of-project and project follow-up (eg. one year after project end) analysis of the project and comparison groups
		x		x		End-of-project analysis of project group, includes real-time evaluations (no baseline data, no comparison group)
			x	x		End-of-project analysis of project group, includes real-time evaluations (no baseline data, no comparison group)
Proof of concept: Are we evaluating a highly innovative, strategic pilot project and want to conduct a proof-of-concept evaluation?	x			x		Baseline and end-of-project analysis of project and comparison groups, where both groups are selected randomly (eg. randomised control trial (RCT))

3.6 Choosing your data collection methods and tools

There are a range of quantitative and qualitative data collection methods and tools you can use. Quantitative methods generally tell you ‘what’ and ‘how much’ was achieved. Qualitative methods generally tell you ‘why’ and ‘how’ certain outcomes were achieved and which segments of the population have benefited. Qualitative methods reflect a more holistic, contextual and exploratory approach that can help the evaluation team to understand the broader impacts of our programmes. The data collection methods you choose will depend on the evaluation questions they need to answer. In general, we recommend that you use a mixed method that combines quantitative and qualitative data collection techniques (see Design, Monitoring and Evaluation (DM&E) Management Operating Standard/Standard Operating Procedures). The validity and reliability of the evaluation data can be strengthened by using more than one methodology, through triangulation. The most common quantitative and qualitative data collection tools are summarised below.

Using mixed methods (qualitative and quantitative) to measure impact

An evaluation of an early childhood programme used both qualitative and quantitative methods to inform the findings and recommendations. These included surveys of parents and teachers, case notes from childcare workers, drawings done by the children and project records and management information.

Quantitative methods

Quantitative methods are used to collect data that can be analysed in a numerical form.

- *Surveys* – most often used when you need to collect information from a large number of people. Most surveys use ‘closed-ended’ questions, where respondents are asked to choose answers from a range of predetermined responses. Surveys allow you to cover a wide range of topics, but they also require a great deal of planning and preparation, particularly to train data collectors (enumerators). Surveys can also have some drawbacks; they may lack depth and miss important information on the context your project or programme operates in.
- *Counting/record-keeping* – a simple method that counts the number of beneficiaries reached or number of activities implemented, for instance.
- *Use of secondary data* – data sources can include existing reports, national household surveys, national censuses, management information systems developed by government ministries, etc.

Qualitative methods

Qualitative methods help us build an in-depth picture among a relatively small sample. They can tell you how the community is organised, what the key relationships are and how different aspects of people’s lives are connected.

- *Key informant interviews* – interviewing individual respondents to explore topics in more depth. The purpose is to collect information from a wide range of people who have first-hand knowledge of a particular project or programme. Compared with focus group discussions, individual interviews can result in more candid discussions on sensitive topics. Interviews can be structured or unstructured. The latter method, where the interviewer begins with some guiding questions but has no set questions to follow, is particularly useful for exploring a topic in its broader context. But it is usually more difficult to analyse this type of data across respondents, as the content of the interviews will vary.
- *Observations* – one of the most common but most demanding of qualitative techniques. Direct observation requires an observer to watch, as unobtrusively as possible, what is happening in a particular project or programme site. Its purpose is to enable the evaluation team to better understand people’s behaviours and interactions and the context in which the project or programme operates. This technique may allow you to identify unanticipated outcomes or issues that participants are unwilling to discuss in an individual interview or focus group discussion. Direct observation can also include ethnographic diaries and photography and video. Videos and photos may be extremely useful, both for aiding memory and as a focus for subsequent questions.
- *Focus group discussions (FGDs)* – involving between 6 and 12 participants, guided by a facilitator, talking freely and spontaneously about certain issues. Group discussions may stimulate a richer dialogue; they also give you the chance to study how groups interact in a more natural setting than individual interviews allow. However, the quality of the data gathered largely depends on group dynamics and the skills of the facilitator.
- *Case studies* – using interviews or direct observation to create a comprehensive and systematic picture of a particular child or adult’s situation. Case studies can focus on individuals, households, communities, programmes, organisations, events, or policies. They are particularly useful in the following situations:¹³
 - Where complex questions have to be addressed in complex circumstances

¹³ C Roche, *Impact Assessment for Development Agencies: Learning to value change*, Oxfam Novib, 1999

- Where individual rather than standardised outcomes are sought
 - In providing a focus for debate and further probing of sensitive issues in informal interviews with other respondents
 - For showing how typical and/or limited the situation is, or to highlight particular issues
 - For demonstrating and communicating the impact of your project when presenting and disseminating the evaluation findings.
- *Narrative inquiry or story-telling* – a relatively new method for gathering qualitative information. Narrative inquiry methods include field notes, interviews, journals, letters, autobiographies and oral stories. For example, a researcher might carry out a study on why girls drop out of primary school. They would look at things like school records and might also interview some girls and spend time observing them and their families. They would then construct a narrative or write a story about their findings. The Most Significant Change technique is one story-telling methodology that is particularly helpful to document stories of change and impact among children, young people and other beneficiaries.
 - *Formative dialogue research* – building on participatory action research and process evaluation to gather evidence through dialogue. This method encourages people to listen to each other; children, young people and adults can all be included and feel respected. It also encourages researchers to be more responsible for sharing feedback and information with children and young people who have given their views.
 - *Participatory methods* – many are specifically designed for use with children and young people. Examples include body mapping, risk mapping, construction of local timelines, network mapping, drawing and the spider tool. Annex 7 has lots more ideas about how you can involve children and young people in your evaluation, including data collection activities. If you want more guidance on using participatory tools, see the '*Kit of Tools for Participatory Research and Evaluation with Children, Young People and Adults*'.¹⁴

3.7 Choosing your sample

You have to decide whether it is feasible or desirable to collect data from the entire population you intend to study. This will be especially important when evaluating activities that are implemented directly at community level. You are most likely to take a sample (a subset) of the population, which is less costly, but can still provide representative data. Depending on how you select this sample, you may be able to make inferences about the whole population based on the results – this is called generalising the findings. You can find out more about sampling in Annex 2.

¹⁴ *A Kit of Tools for Participatory Research and Evaluation with Children, Young People and Adults*, Save the Children Norway, 2008. available at: http://tn.reddbarna.no/default.asp?V_ITEM_ID=19028

Getting help with sampling

Sampling can be a complex process and we do not expect you to have the technical skills to undertake this. You can get support from technical advisers within Save the Children. You can also approach universities or local research institutions. It is also helpful to build relationships with local, national or regional academics. They are a good source of expertise and often have researchers who can be made available to help you carry out evaluations and baseline studies.

For more detailed guidance on the technical aspects of sampling, see Annex 2

You may want to select a 'purposive' sample – deliberately choosing certain individuals to be involved to achieve a certain goal. In a focus group, for example, you may seek out certain respondents because they have more knowledge about an issue, or to ensure that different groups and views are represented. Purposive samples are often used when evaluating pilot projects. A purposive sample does make it more difficult to draw strong quantitative inferences.

When selecting your sample, ask yourself these questions:

- What are our evaluation questions? (This will influence sample size)
- What is the study population (girls, boys, women and men) we are interested in, from which we want to draw a sample?
- How many girls, boys, women and men do we need in our sample?
- How will these children, young people and/or adults be selected?
- Who else might be interested in the findings?

You can use random sampling or non-random sampling. The former is usually associated with quantitative methods, the latter with qualitative methods.

Once you have decided how to select your sample, you need to work out the desired sample size. How you do this depends on whether you are using quantitative or qualitative data collection methods. For quantitative methods, there are formulae that you can use depending on the data you want to collect. For qualitative methods, there are different approaches depending on your evaluation questions and the focus of the information you want. You need to be aware that sampling can lead to biased results. To find out how you can avoid bias, see Annex 2 (section 4, 'Bias in sampling').

Disaggregating data

As Save the Children is dedicated to changing the lives of the most vulnerable and disadvantaged children, it is vital that we are able to report on exactly which children our programmes have targeted and reached. Disaggregating data enables us to assess any changes in equality and discrimination over time and to understand which children are being included in – or excluded from – the benefits of a programme or project. This provides vital information for managers to make adjustments to programme strategies or implementation.

Data are often disaggregated into these key categories: sex, age, disability status, area of residence (rural or urban), ethnic group, migrant status, refugee status, single/double orphan, child-headed household. There are separate tools you can use to assess the gender impact of your programme (see box below). You will need to decide which categories to use when designing the data collection tools for your evaluation to make sure they capture the information you need. Be aware that as the number of categories increases, the sample size needed in order for your data to be representative increases substantially.

Carrying out a gender equality audit

A gender equality audit is a way of looking at the organisation, its systems and procedures and the understanding of staff involved in a project or programme, to get a picture of what is being done on gender equality. This is a good way to help you identify any challenges and needs around gender equality and to develop an action plan to address them. A gender equality audit should seek to answer questions around levels of commitment, understanding and the capacity and resources available for work on gender equality.

A gender equality audit should help you get answers to these questions:

- What level of commitment is there in the team/office/among partners to promoting gender equality?
- How well do staff and partners understand terms such as gender, gender equality, women's rights, mainstreaming? Who finds the terms confusing, who is opposed to them and who questions their meaning?
- What skills, resources and capacity do staff have to support work on gender equality and what do they need to do it better?
- Is gender integral to all of the team/office/organisation's work, including key documents such as plans and reports?
- How does the focus on gender equality fit with other organisational priorities?
- Who is responsible for ensuring that gender is integrated into work at all levels of the country office and across all functions? How is that monitored?

To find out more about how to conduct a gender equality audit, see Annex 8.

3.8 Ensuring the quality, validity and reliability of data

Validating your findings

Triangulation – the use of two or more theories, sources or types of information, or types of analysis to verify and substantiate an assessment – requires a mixed method approach to evaluation design. By combining multiple data sources, methods, analyses and theories, the evaluation team can seek to overcome the bias that may arise from single informants, single methods, single observer or single theory studies. Triangulation allows you to test the consistency of the data obtained during the evaluation.

Triangulation is also referred to as “cross-examination”. When a finding is confirmed or validated across multiple data sources, we can be more confident about its accuracy. There are several different categories of triangulation,¹⁵ including:

- Data triangulation (using more than one sampling strategy to gather slices of data at different times and in different social situations)
- Investigator triangulation (using more than one researcher to gather and interpret data)
- Theoretical triangulation (using more than one theoretical position when interpreting data)
- Methodological triangulation (using more than one method for gathering data)
- Secondary data triangulation (using secondary data to confirm the findings and analysis).

To sum up, triangulation can be a useful tool to enhance the credibility of your evaluation by confirming the validity of its findings (for guidance on how to involve children see Annex 7).



Involving children in validating and triangulating data

You could convene a workshop with the children and young people who were involved in collecting data and spend some time reviewing the findings. Ask them to group different themes together and to identify, for example, the five key findings from information collected using participatory tools. Encourage them to ask questions about the findings and to see if there are any issues that are repeated, or even any unexpected findings.

Checking quality, validity and reliability of data

The quality of the data you collect is of fundamental importance (see Design, Monitoring and Evaluation (DM&E) Management Operating Standards/Standard Operating Procedures). Data should, at an absolute minimum, be *traceable back to its source*. This means keeping clear records of where the information has come from – for instance, from a field report, attendance records, discussions with clients, observation in a classroom or clinic setting, or through feedback/complaints mechanisms, etc. Data quality can be assessed according to its accuracy, completeness, precision, timeliness and integrity. For a good example of a tool to assess data quality, see Spencer *et al* (2003), *Quality in Qualitative Evaluation: A framework for assessing research evidence*.¹⁶

Data validity means confirming the accuracy and meaningfulness of your data. There are four types of validity commonly examined in research and evaluation:

- 1 *Conclusion validity* asks if there is a relationship between the programme and the observed outcome.
- 2 *Internal validity* asks if there is a relationship between the programme and the outcome observed and whether it is a causal relationship.
- 3 *Construct validity* asks if our measured outcomes actually reflect the real outcomes we are trying to evaluate.

¹⁵ See N Denzin, *Sociological Methods: A sourcebook*, Aldine Transaction, 2006

¹⁶ L Spencer, J Ritchie, J Lewis and L Dillon, *Quality in Qualitative Evaluation: A framework for assessing research evidence*, National Centre for Social Research, Cabinet Office, 2003, available at: www.civilservice.gov.uk/Assets/a_quality_framework_tcm6-7314.pdf

4 *External validity* refers to our ability to generalise the findings of our study to other settings.

Data validation should be an integral component of any evaluation. Methods for data validation include testing of data collection instruments before use and peer review of the methods and approach.

Data reliability is the consistency of your measurement, or the degree to which an instrument measures the same way each time it is used under the same condition, with the same subjects. In short, it is the ability to repeat your measurement. A measure is considered reliable if a person's score on the same test given twice is similar. Methods for checking data reliability include comparing results using repeated measurements (test/re-test).

Checklist – Designing your evaluation

Defining the scope

- ✓ Have you established the purpose of evaluation?
- ✓ Have you agreed the questions your evaluation needs to answer?
- ✓ Have you considered the information needs of different stakeholders?

Designing the evaluation

- ✓ How will you involve children and young people in designing the evaluation?
- ✓ What is the level of rigor needed to answer your evaluation questions?
- ✓ Have you aligned with relevant program M&E plans and appropriate SC indicators?
- ✓ Which DAC criteria or other criteria will you use?
- ✓ Does your evaluation design tell you what you want to know?
- ✓ What is the theoretical or conceptual basis for the outcomes you want to study?
- ✓ Who is your audience for the evaluation and how will you use the findings?
- ✓ What stage of the project cycle are you planning the evaluation at?
- ✓ Are there any ethical considerations?
- ✓ What are your budget and time constraints?
- ✓ Which qualitative and quantitative methods and tools will you use?
- ✓ How will you to ensure the quality of the data collected?
- ✓ How will you involve children and young people in checking the quality, validity and reliability of data?

4 Planning your evaluation

After you have chosen which type of evaluation is right for your project or programme and defined the evaluation objectives, questions and methodology, you need to start planning. This chapter takes you through how to plan for the evaluation.

4.1 Drawing up the terms of reference

You need to draw up a document that clearly sets out the terms of reference (ToR) in order to achieve a good quality evaluation (see box below and Annex 3). ToRs are usually drafted at country office level, but technical assistance may be provided by Save the Children International Regional Office, Members or Global Initiatives.¹⁷ The ToR should specify the expectations for the evaluation, its design and the methods that will be used. The document must be approved by the steering group or reference group¹⁸, or relevant senior management member.

What information should the terms of reference (ToR) contain?

We suggest you use these main headings:

- Background
- Purpose of the evaluation
- Objectives and key evaluation questions
- Scope of the evaluation
- Evaluation design and methodology
- Organisation, roles and responsibility
- Outputs
- Timeline
- Budget/ resources
- Plan for disseminating the findings and sharing learning

¹⁷ Please contact the Regional Office M&E manager for further advice.

¹⁸ See section below on Managing Roles and Responsibilities (p34)

4.2 Involving children and other stakeholders

You need to identify all those who have been involved or affected by the project or programme – your stakeholders. You must ensure that children and young people from different backgrounds have the chance to participate meaningfully in the evaluation process.

Children and young people can participate in the evaluation in a number of ways:

- As **advisers** to the evaluation team, commenting on the relevance of the evaluation, its scope and content. They can help to draw up research questions and give input to the evaluation design. They could also be members of an advisory or reference group
- As **peer researchers**, helping to collect and analyse data
- As **documenters**, noting their own and other children's views on the project activities and giving their reflections, analysis and suggestions on what can be done better. They could use creative and audio-visual methods of documentation. They could also develop a child-friendly version of the final evaluation report
- As active **respondents**, using participatory methods and tools
- As **reviewers** of the draft evaluation report
- As **active change agents**, disseminating learning and following up on conclusions and recommendations, contributing their ideas and proposals about how to put recommendations into practice.

Whatever ways you choose to involve children and young people in the evaluation, their participation should always be meaningful, safe, voluntary and inclusive. Children and young people who take part in the evaluation must be able to give their informed consent – this is a basic ethical requirement. You may also need written consent from the child or young person's parent or carer in the case of a 'minor'. If the young person is estranged or separated from parents or guardians and provision is made to protect the young person's safety, security and wellbeing in the conduct of research, then the young person's consent may be enough.

To find out more, refer to the Child Safeguarding Policy and the *Practice Standards in Children's Participation*.



How to involve children?

You could organise a meeting or workshop to consult with children and young people on the relevance of the evaluation for them. Ask them how they would like to be involved and contribute to the evaluation process, then set up a timeline. They can also help you identify other key stakeholders who should be involved.

Remember to make sure that any materials you use are adapted to children's evolving capacities. There are lots of participatory, age-appropriate methods you can use to ensure children's meaningful participation, including pictures, games and role-plays. Find out more in Annex 7.

It is important to ensure that you involve representatives from different groups, particular attention should be paid to ensuring that girls' and women's voices are adequately sought, as well as children and young people with disabilities, from minority groups or speak a minority language. It is important to consider the possible roles children and young people can have in the evaluation process at the very beginning and ask them what role they would like to play.

4.3 Setting your budget

Earlier in the project design, you should hopefully have elaborated your evaluation budget in your M&E plan. Your budget will influence the scale of your evaluation and the methodology and tools you use. Generally, we recommend that 3–4% of the total programme budget is allocated for the evaluation. You should review that budget in light of the specific design you have chosen. You then need to draw up a more detailed budget that clearly identifies the allocation and sources of funds needed for all the evaluation activities. Your budget lines are likely to include external consultant(s), internal technical assistance, research and data collection, translation of tools, data cleaning and analysis, report writing and review and dissemination and publication of findings. As you draw up the terms of reference and the evaluation plan, you will have to identify and allocate adequate human and financial resources to implement the evaluation activities properly.

4.4 Agreeing a timeline

You need to ensure that the evaluation is carried out and the report completed in a timely fashion by agreeing a timeline with all stakeholders involved. This timeline should take into account new programmes or strategies that the evaluation findings may need to feed into. The complexity of the project, the methods employed, geographical and logistical considerations, availability of data and available resources are all factors that are likely to influence the timeline. You will need to plan adequate time to consult with and involve children and young people and other stakeholders, including producing child-friendly materials that may need to be translated. If children are to be meaningfully involved, they need sufficient time at each stage of the process.

4.5 Getting the right evaluation team

The composition of the evaluation team is crucial to ensuring a high quality evaluation. It is important that you consider different ways to include new or alternative perspectives on the team. Ideally, the following groups of people, preferably with complementary skills, will be involved:

- **An external independent consultant, with specialist experience of conducting evaluations, as well as knowledge of the thematic programme area and the country context**¹⁹

We advise that you use in-country technical resources whenever possible, as they will usually be more familiar with the context, more likely to speak a local language and more cost-effective. Academics from national or local institutions are also a good source of expertise. See box below for more guidance on using external consultants.

¹⁹ For real-time evaluations in emergency response programmes, the team will ideally be internal to Save the Children but external to the programme. See Monitoring and Evaluation document library on the SCI Xtranet

- **Staff, thematic managers, sector leads, M&E staff**

Inviting staff from different country programmes in the region to participate in the evaluation has worked very well in the past and should be considered if resources allow. Staff with specific expertise, like children's participation and/or gender and diversity, should be part of the evaluation team to ensure that these elements are assessed to the highest possible standard.

- **Local partners**

In many cases, you will manage and implement the evaluation together with your local partner(s). You will need to consider their capacity to manage or implement the evaluation and provide support accordingly. Inviting representatives from local partners will support capacity building and greater ownership of the process. It may also support advocacy efforts. You might also want to consider inviting one of your donors to take part. This will clearly depend on your relationship with them and the nature of the project.

Using external consultants

Save the Children International recommends that an external, independent team leader is contracted to lead the evaluation. This will largely depend on budget availability and donor requirements.

One of the principal reasons for contracting an experienced external consultant(s) to lead the evaluation process is that they are likely to be more objective and to have skills that may not be available internally. It will also add to the credibility of your evaluation. But whether you are able to do so will depend on a number of factors, including the length, budget, scope and scale of the project or programme and donor requirements.

The evaluation may be led by a single consultant, a team of consultants, or a research group, such as a university. The consultant(s) will need to work closely with the project or programme team, as their in-depth knowledge is crucial to ensure that the evaluation asks and answers the right questions.

Here, we summarise the advantages and disadvantages of using an internal or external evaluator.

Internal evaluator

Advantages

- May have better overall knowledge of the programme and possess informal knowledge
- Less threatening, as already familiar with staff
- Less costly
- More conducive to organisational learning

Disadvantages

- May be less objective or biased by involvement with the project and the organisation
- May be more preoccupied with other activities of the programme and not give the evaluation complete attention
- May not be adequately trained as an evaluator

External evaluator

Advantages

- More objective, offers new perspectives, different angles to observe and critique the process
- May be able to dedicate greater amount of time and attention to the evaluation
- May have greater expertise and evaluation skills

Disadvantages

- May be more costly and require more time for the contract, monitoring, negotiations, etc
- May be unfamiliar with programme staff and create anxiety about being evaluated
- May be unfamiliar with organisational policies, certain constraints affecting the programme

Children and young people

There are various ways that children and young people can be involved in the evaluation team, including developing evaluation questions, conducting data collection, triangulation/ analysis and communicating the findings (see Annex 7).

Women and men from the community, mothers and fathers and carers

Depending on the nature of your project or programme, you may feel it would be beneficial to involve community representatives as part of the evaluation team. They can usefully help with developing evaluation questions, data collection and triangulation / analysis. Please ensure that you involve both women and men from the community in equal numbers as appropriate.

4.6 Finding the right consultant to lead your evaluation team

As we have said, we strongly advise that you use an external consultant with the relevant expertise to lead the evaluation team. We suggest you get at least three proposals (technical and financial) or CVs from different individuals or organisations. In some instances, procurement may require a competitive tender process. In all instances, you should adhere to donor requirements.

It is good practice to draw up the criteria or person specification before you invite people to apply for the role of evaluation team leader. For transparency, these criteria should be included in the ToR. They may include previous experience with Save the Children, local experience, experience of involving children in participatory evaluations, value for money, track record and quality of the proposal and evaluation design. You should evaluate the proposals with one or more colleagues to reduce the risk of actual or perceived bias. Interviewing final candidates may assist, clarify or deepen your understanding of some aspects of the proposal.

4.7 Managing roles and responsibilities

In order to apply the principles of participation, transparency and accountability, your evaluation process should clearly define a management structure. The roles and responsibilities of the key players (see Table 3 below) should be set out in the terms of reference and the M&E plan.

- As well as the members of your evaluation team, different people will be involved in the *management process*.
- All evaluations should appoint an *evaluation manager*. This person will assume day-to-day responsibility for managing the evaluation process. Their role is distinct from that of the evaluation team leader, who will lead on content and implementation. The evaluation manager is also responsible for logistics and for facilitating the broader evaluation process.
- You may want to set up an advisory group or reference group that involves key stakeholders, including children and young people.²⁰ A reference group can make the evaluation process more accountable and can help to ensure buy-in from stakeholders, including donors. The reference group(s) might, for instance, input to the ToR and comment on your plans for the evaluation questions, design and methodology.

If you are evaluating an emergency response programme, it may be particularly useful to set up reference groups, especially when there are several Save the Children members with a large stake in the programme.

²⁰ Evaluation reference groups are defined as participatory spaces established for the systematic involvement of relevant stakeholders in the evaluation process (UNDP/UNIFEM Evaluation Unit, Guidance Note on Management Structures and Reference Groups for Evaluations, 2009)

Table 3: Who should be involved in your evaluation and why

Who	Why
Thematic programme manager/ Project manager	<p>The programme manager should have overall responsibility for managing the evaluation process and the evaluation team, including any consultant(s). He or she will be most familiar with the project and know what they need to get from the evaluation.</p> <p>It is important that the manager keeps sufficient distance so as not to compromise the integrity of the findings. There may be a conflict of interest if they have been too closely involved in the project. If that is the case, it needs to be declared and discussed by the evaluation team prior to implementation of the evaluation.</p>
M&E staff	<p>If the country office has designated M&E staff, it may be appropriate for one staff member to take on the role of evaluation manager, as they will have the best knowledge of the evaluation process and methodological issues. He or she should work closely with the thematic/implementing project manager in developing the ToR and reviewing the inception report and evaluation tools prior to approval (see below). M&E staff could also support the evaluation team, by being involved in data collection and analysis.</p> <p>If the country office has M&E staff but for some reason they are not appointed as the evaluation manager, they should at least take part in any reference group that is set up and have the opportunity to read and comment on the draft evaluation report.</p>
Country director/ Programme director/ Deputy director/ Senior management staff	<p>A senior member of staff should maintain oversight of the evaluation. They might also be the one who commissioned the evaluation. He or she should be present during the initial meeting with the evaluation team and at the debrief meeting. This ensures that any concerns or queries can be raised with the consultant(s) leading the evaluation team early on.</p> <p>A senior member of staff should also be part of any reference group that is set up.</p>
Technical advisers/ Member or Global Initiative at the Regional Office and Head Office levels	<p>Member and Global Initiative technical advisers are available to comment on the ToR, the evaluation plan / inception report, the design of the evaluation and draft reports. They might also be invited to join any reference group that is set up.</p> <p>Global Initiatives will want to engage with evaluations of shorter or smaller programmes if they are of particular relevance to wider advocacy, research or policy work, but this will be decided on a case-by-case basis. If your evaluation is complex and you would like the support of a technical adviser, please contact them well in advance to give sufficient time to negotiate this.</p> <p>They will also have a key role to play in supporting learning and sharing.</p>
Emergency advisers	<p>For all emergency programmes, Emergency Monitoring, Evaluation and Accountability Advisers are able to provide support. They will be involved in and may, on occasion, lead the process of selecting the evaluation consultant(s) or team. They will also draft, comment on or finalise the ToR and provide further support as required.</p>
Other Save the Children Members and Save the Children International Head Office/Regional Office	<p>You may want to invite other Save the Children members to be part of any reference group, especially if they have a stake in the project or programme, or can provide relevant technical assistance or expertise.</p>
Children, young people and other members of the community	<p>The target beneficiaries should be fully consulted about the evaluation process and their interests and information needs should be identified and considered. They might be represented on the reference group or through another feedback</p>

Who	Why
	mechanism.
Government/donors	Where projects or programmes are working closely with national or sub-national government structures, representatives of these structures should be invited to be part of the reference group. You may also want to include donor representatives in the reference group. Sometimes donors want to be part of the evaluation team. If this is the case, you should discuss what role they would like to play prior to the evaluation.
Partners/civil society organisations/national or international academic institutions	Implementing partners or other partners have strong knowledge of the context and are often a voice for civil society interests and needs. You could invite representatives of partner organisations to be part of the reference group. Academic institutions can help with the design of the questions, tools and methodology and they may be able to share relevant secondary data.

The pros and cons of involving many different stakeholders

From the outset, you need to be aware that involving more stakeholders in the evaluation process will have implications for how it needs to be managed.

On the one hand, the evaluation will benefit from the competencies and knowledge of a range of individuals and groups, which should promote accountability and strengthen buy-in, sharing and learning. On the other hand, coordinating and managing the evaluation process will be more demanding and time consuming. Large reference groups, for instance, are a good illustration of these pros and cons.

So, as we said before, you need to carefully plan and organise your evaluation according to the time, budget and staff available and the size and strategic importance of the project or programme.

4.8 Managing external consultants

In most cases, external consultants contracted to lead the evaluation team will sign a formal contract with Save the Children. This contract is key to the success of the evaluation process. A contract manager should be appointed, to be responsible for the day-to-day administration and management of the relationships with external consultants or ‘contractors’. The contract manager’s responsibilities are outlined in Annex 9, which also includes a contract review checklist. You will also find guidance on intellectual property (IP) rights and authorship – issues that need to be carefully considered by the programme or project manager.

Checklist – Planning your evaluation

- ✓ Have you drawn up clear terms of reference (ToR) that have been agreed by the key stakeholders?
- ✓ Have you set the budget and timeline for the evaluation?
- ✓ Have you identified an external consultant to lead the evaluation team?
- ✓ Have you identified the evaluation team?
- ✓ Have you considered the management structure for the evaluation?
- ✓ Have you defined the specific roles and responsibilities of key players?

- ✓ Have you consulted with children and young people to find out how they will be involved?
- ✓ Have you ensured that contracts meet SCI's standards regarding managing contracts with external parties?
- ✓ Have you gone through the contract review checklist?
- ✓ Have you agreed on authorship issues with the evaluation team?

5 Implementing your evaluation

This chapter explains the implementation phase of the evaluation, from the inception phase to dissemination of the final report.

5.1 Fine-tuning your evaluation plan

Before embarking on the evaluation process, you need to take some steps to ensure that the evaluation is conducted in line with its objectives. This is usually called the inception phase and you should allocate sufficient time for this before the evaluation is implemented. During this phase, you should share the proposed methodology and plan with the wider evaluation team. This is the best way to ensure a common understanding between the project owner/commissioner or sponsor, stakeholders and evaluators about the purpose and scope of the evaluation, the questions it must answer, the methodology, outputs, timeline, etc.

What is in an evaluation plan / inception report?

Before you begin implementing the evaluation, you should develop an inception report, or evaluation plan. It should include the following:

- Purpose and scope (from the ToR)
- Methodology
- Clarification, interpretation and elaboration of the evaluation questions
- Evaluation matrix (link evaluation questions with methods chosen)
- Existing sources of information, with special attention to monitoring data and Save the Children indicators where appropriate.
- Timescale and expected outputs
- Roles and responsibilities of evaluation team members
- Participation of stakeholders
- Resource requirements, including budget and logistics timeline
- Requirements for reporting and dissemination of findings
- Planned dissemination and any restrictions to confidentiality

Essentially, the **inception report** builds on the design and methodology you have chosen for the evaluation as set out in the terms of reference (ToR) (see Chapters 3 and 4). But it goes further, giving detail on the sampling schedule, timeline and the roles of each team member, for instance. The inception report also reflects the evaluation team leader's understanding of what questions the evaluation needs to answer and how it is best conducted. This will usually be based on a literature or desk review of relevant programme documents, reports, reviews, baselines and monitoring data, along with any relevant national or district-level data, evaluations and research studies. The evaluation plan will help the team fine-tune what they need to address during the evaluation and the methods and tools they need to use to give them the right information.

As mentioned previously, the evaluation plan should be shared with stakeholders and/or a reference group (see Chapter 4) for their comments and input. It should also be approved by the person who commissioned the evaluation (usually the country office or programme manager).

It is worth noting that while there is some overlap between the ToR and the inception report, the latter gives the evaluation team the chance to refine and revise the ToR if necessary. It also strengthens ownership of the process among stakeholders.

5.2 Facilitating children and young people's participation

As we have already said, children and young people from a wide range of backgrounds must be given the opportunity to participate meaningfully in the evaluation (see Chapter 4). There are however, other issues you need to consider to facilitate children's meaningful participation in the process:

- **Staff capacity**

You need to ensure there is capacity within your country or programme team to facilitate children's participation. Staff and partners need to be fully aware of the power relations between adults and children and between children themselves. Some staff and partner organisations may benefit from further training in participatory techniques. Body language and verbal language, the timing of activities, sensitivity to the particular challenges that some children may face in expressing their ideas – these are just some of the areas you need to consider when preparing field staff for their role as facilitators.

- **Tools and methods**

Consultants and/ or staff who are facilitating participatory tools need to be sensitive to a number of key factors, including cultural, social and economic differences, the different types of project and the evolving capacities of children and young people taking part. Facilitators should have good knowledge of the participatory tools and methods available and the skills to identify which to use in a particular setting. For example, visual methods may be particularly appropriate for use with younger children and those who are not confident in writing. Other creative tools like theatre exercises may be ideal in one place but cause great unease in another, particularly if introduced too early. You can find more practical examples and guidance in Annex 7 and in the *Kit of Tools for Participatory Research and Evaluation with Children, Young People and Adults* (Save the Children Norway 2008).

- **Organising and managing evaluation activities**

It is best to conduct evaluation activities with children in a familiar environment in which they feel comfortable. Always make sure that activities are planned so as not to interfere with children and young people's other commitments, such as school, work and household responsibilities.

Dividing a group into girls and boys may make it possible to discuss gender-specific issues, as well as making participants feel more comfortable to speak up. In any event, it is important to ensure that groups are composed in such a way that participants feel comfortable working together and exploring different ideas. Meetings are also likely to be most fruitful when children and young people are divided according to age. This is because young children and adolescents usually possess different types of information and are capable of very different levels of attention, reflection, discussion and analysis.

You will also need to consider whether you need an interpreter and to make sure that all materials are translated into the appropriate languages.

5.3 Designing and testing data collection tools

You will have decided on the methods you want to use to collect data during the design stage (see Chapter 3). You should already know the sampling methods you will use (eg, sample surveys, focus group discussions, or case studies) and have identified the sources of data needed.



How to involve children in collecting data

Make sure that you engage children in the pre-testing of data collection tools. Children will often have the commitment and capacities to be involved in data collection. You will need to ensure that they receive appropriate training to assist them in this role.

When designing and testing your data collection tools, make sure you base them on the evaluation objectives and questions set out in the design phase. Here are some other pointers:

- Bear in mind any other tools being used in baseline and ongoing monitoring of the programme or project, as well as existing literature (from Save the Children or external agencies), as there may be standardised questions for the indicators you are using.
- Use participatory exercises for the questionnaire design, involving thematic advisers/managers and other stakeholders.
- Make sure that you only collect the data you need (based on the defined scope of the evaluation, or the indicators under investigation).
- Make sure the data collection tools you want to use allow you to disaggregate data if need be – for instance, by sex or age group. Certain tools will be more appropriate for different sexes and ages.
- The data collection tools you select should be approved by the evaluation manager and reference group.
- Translate, back-translate and field test tools in selected communities or among selected beneficiaries. The level of effort required for this depends on the context (eg, if the evaluation is replicating a baseline survey, the questions should be the same as those used in the baseline questionnaire, so not much work should be needed for subsequent evaluations). You may also be able to use external human resources to do this (eg, university students or interns).
- Tools will need to be translated into local languages to enable data collectors and participants to use them properly.

5.4 Training

You will need to provide training on how to use the data collection tools and methods for enumerators and all those involved in data collection. Enumerators might be project staff or members of the target group (that is, children and young people involved in the project). They might be hired by the consultant or organisation undertaking the evaluation or they could be students or researchers from local universities. If an interpreter is required to facilitate the training, you should budget for this.

Although this is not an exhaustive list, training topics can include the following:

- Save the Children's guiding principles, vision, mission and ethical standards.
- Child Rights Programming and Save the Children's theory of change.
- Background on the programme/project and teams, noting any previous evaluations, baseline survey, monitoring data and appropriate Save the Children indicators.
- Reviewing the purpose and scope of the evaluation and key questions it needs to answer.
- The data collection methods or tools that will be used, allowing participants to review the questionnaires in detail. This can be done in a participatory way and any questions and discrepancies discussed with the group. Thereafter, exercises and role-plays can be carried out to begin familiarisation with the tool.
- Sampling methodology, focusing on the practicalities of carrying out the evaluation in the field (including training on the implementation of field sampling if relevant).
- Ethical issues related to data collection, including confidentiality, child safeguarding and where to access further support and guidance if necessary. It should also cover what to do in the event of any conflict of interest (if the enumerator has connections with the person or people being interviewed).
- Exercises, role-play and piloting/pre-testing interviews in nearby communities will help the evaluation team familiarise themselves with the tools and improve them if necessary. If you are using quantitative methods, a standardisation test can be carried out, where several researchers collect the same data, which can then be compared at the end of the training session. Common errors should be discussed in plenary.

5.5 Managing data

Data collection

You need to draw up a detailed schedule and plan for data collection activities, specifying locations, logistics and who will be responsible for collecting data in each area. This is usually included in the evaluation plan/inception report described above, which is reviewed and approved by the management team or reference group. They should pay particular attention to available resources and other constraints (eg, areas that are remote or otherwise hard to reach).

The evaluation plan should also include a system to address data quality issues (eg, checklists, observations, spot-checks). Where relevant, it should incorporate meetings with leaders of local communities to seek permission to carry out the survey in their villages.

During the data collection process, the evaluation team should review completed survey tools at the end of each day to identify any corrections, additions or clarifications needed. It is good practice for the evaluation team to prepare brief daily reports on interviews carried out. If possible, it is advisable to do the data entry on a daily basis for large quantitative data collection activities (this can be done in the field as long as the team have laptops or personal digital assistants (PDAs) that have automated programmes for data entry). This allows quick checks for any missing data or inconsistencies.

A note on incentives

While Save the Children does not encourage giving financial incentives to people taking part in data collection activities, we recognise the importance of other forms of incentives, such as drinks, snacks, etc. You may also need to reimburse participants for transport, childcare, lost earnings, etc. The evaluation manager and team need to give careful thought to this issue and decide what is appropriate for the particular context.

Data input, cleaning and analysis

The analysis of the evaluation findings, from desk review and primary and secondary data collection, will most probably be done by the evaluation team leader and the evaluation team. It is important that the project management team and members of the reference group understand the methodology that will be used to clean, validate and analyse the data.

To ensure timely availability of the collected data, preparations should be started early to develop a database (preferably using software such as Excel, Epi-Info, SPSS or Access) that includes data-entry check forms to verify data, as well as queries that can present the data in different forms. It is considered good practice to enter (a random part of) the data twice, checking for differences between both to verify the quality of the data. Cleaning the data is a process whereby you search for missing data or logical inconsistencies.

After data entry and cleaning, data can be analysed using software programmes similar to those mentioned above. The simple tabulation will then be complemented by more detailed information on the indicators used in the evaluation.

Quantitative analysis procedures can include descriptive statistics, trend analysis, cost analysis, charts, tables and graphs. To ensure that the precision of the evaluation can be objectively assessed, it is considered good practice to present statistics using confidence intervals that take into account the sampling design of the survey. This way, for example, indicators can be correctly assessed in terms of their significance compared with the baseline data.

For qualitative information, the analysis procedures can include content analysis, summaries from focus group discussions, etc. The evaluation analysis can be enriched by triangulating different sources of information (see page 26 and by contrasting different subsets of qualitative and quantitative information).

For both quantitative and qualitative data analysis, data should be disaggregated before being analysed. At a minimum, data should be analysed by sex, to highlight any gender differences. If the sample also includes boys and girls and men and women with disabilities, or those who are from a minority group, their responses also need to be analysed separately to obtain an understanding of the particular experiences of these groups of people.

The analysis should be summarised according to the evaluation objectives, preferably combining the quantitative and qualitative information collected.

The main findings and summary analysis can then be presented and discussed with the relevant staff and stakeholders during participatory feedback meetings. This not only serves the purpose of informing stakeholders, it also enriches the data with their input. While more detailed planning for using the results of the evaluation is covered in the next chapter, preliminary action plans (eg, for further qualitative research) can be suggested at this point.



Getting children's views on the data collected

You can invite children to participate in the analysis of data. Children have their own interpretations and perspectives based upon their experiences that will contribute to the quality of the data analysis.

Data storage and ownership

Your data collection plan should state how you will file the data (whether electronic or paper-based) and keep it secure. Evaluation data and records should be accurate, complete, authentic and reliable. The data should be recorded in a form that is adequate for verification purposes, but individuals' names should not be used. You must draw up guidelines on the ownership of and access to databases containing confidential information and ensure that all involved in the evaluation are aware of these and comply with them. If external consultants are working with datasets, the programme manager should make sure that Save the Children keeps a secure copy of the datasets. All evaluations must be in line with the intellectual property (IP) clause of Save the Children International (see Annex 9 for more on IP rights).

As per our legal policies, Save the Children International aims to be as transparent as possible, so all datasets should be available unless they identify individuals and are classed as confidential. For purposes of complementary data analysis or replicating analysis (eg, for subsequent evaluations after a baseline), the dataset should contain well-documented descriptions of variables, sampling points, statistical formulas, codebooks, computer data input files used to generate data, etc.

If needed, you should seek technical assistance on these matters.

5.6 Writing the evaluation report

When writing or reviewing the evaluation report, check back against the ToR and any donor requirements. The report should differentiate between the evaluation findings, conclusions and the evaluation team's recommendations.

The box below gives a template for an evaluation report. But it is only a suggested outline – it does not prescribe the format to be used for all evaluation reports. It sets out the content that should always be included, but, as always, donor or grant agreement formats take precedence. If there is flexibility on the format, keep in mind that a brief and concise report will be easier to use. This may mean putting more detailed information as annexes (eg, on the methodology, data collection tools and plans, training, a summary of costs, etc).

When writing up the report, you should include a section on the methodology used. It is important to describe which tools the team used to collect data and to give information on the sampling approach and sample size, as well as any potential sources of bias identified.

You should aim to document the evaluation findings in a format and language that all stakeholders can easily understand, respecting ethical standards such as being mindful of cultural, religious and gender issues. As part of communicating the findings of the evaluation and following up any recommendations (described more fully in the next chapter), you should carefully plan how to ensure timely dissemination and feedback of the report to project stakeholders and other audiences, both within Save the Children and more broadly.

Communicating statistics

You will obviously want to present some of the statistical information derived during the data collection process. Writing about statistics can be very dry, so you could try presenting your analysis in a way that tells a story about the data. The narrative should bring the data to life, making them real, relevant and meaningful to your audiences, external and internal. Any statistical information you include in the report should be clear, concise and accurate. For more on how to present the statistical information you have collected during the evaluation, see Annex 6.

5.7 Reviewing the draft report

The review stage is a critical opportunity to ensure that quality standards have been met. The draft evaluation report should be checked against the original ToR and relevant design documents. The following quality framework was designed for qualitative evaluations and is also suitable for a mixed method approach (www.natcen.ac.uk). The framework can be useful at the review stage and especially if it is introduced early in the design and/or report writing phase.

How to structure your evaluation report

Title page

Name and location of the programme or project, date of evaluation, one-line description of the evaluation and the name (and/or the institute) of the evaluator and acknowledgements

Contents page and list of acronyms

Including boxes, figures, tables, maps and annexes

Executive summary

All evaluation reports should include: a brief description of the context and the programme or project; the objectives of the evaluation; the methodology used; key findings and conclusions; recommendations and lessons learned

Background and description

A description of the programme or project goals, objectives, indicators and key activities

Rationale: scope and purpose of the evaluation

Explaining why the evaluation is being conducted and how the findings will be used, as well as clearly identifying the geographic and thematic coverage

Methodology

Rationale, sample selection, sources of information, a description of the tools (referencing annexes), data analysis (quantitative and/or qualitative)

Findings

A description of the findings against the objectives/indicators in the ToR

Conclusions

Conclusions derived from the findings that are consistent with the data and methodology used

Recommendations

Recommendations for the project/ programme based on evidence and analysis, that are realistic and achievable taking into consideration the context and resources available

Lessons learned

Both positive lessons and ongoing challenges, to share learning with other Save the Children country programmes

Annexes

These could include the project results framework or logframe, the ToR, the evaluation schedule, the list of people interviewed, data collection tools, bibliography and list of documents reviewed

Peer reviewers (both internal and external) can be very valuable in the review process. The focus should be on whether the evaluation is methodologically and analytically sound, transparent, practical and relevant.

As a guide, the role of a reviewer is to assess the report against key quality criteria:

- Are the research findings credible?

- Do they increase knowledge or learning?
- Does the evaluation report address the original evaluation questions?
- Is the research/evaluation design clearly justified and appropriate?
- Is the methodology used clearly explained?
- Is the analysis relevant to the context of the project/programme and addressed through the recommendations?
- Has a diversity of perspectives been considered?

It can be helpful if review criteria are given to the evaluation team leader before the evaluation so they know the basis on which the report will be reviewed. The success of the quality review depends on the willingness of reviewers to be fair and reasonable, to exercise their sound judgement and to adopt a balanced approach that takes into account the context of the programme or project. The author/s will be more likely to accept the reviewers' comments if they are constructive and help the author by pointing out areas that could be improved in a revised report.

Finally, you will need to manage the review process carefully to protect all parties and avoid unnecessary conflict. The person managing the review process should keep records of all review comments and any amendments made by the evaluation team leader in response. A quality review guide is set out in Annex 4. It can also be used for an internal quality review.

5.8 Approval procedures

The relevant country office thematic programme manager is *responsible* for the final report. She or he must ensure the following:

- That it incorporates feedback from stakeholders (external and internal, including country/regional office programme leadership, technical and M&E staff, the relevant global thematic advisers for any Global Initiative indicators and, for high-risk grants, the relevant Member HQ).
- That the final version is shared for approval with the country office leadership (country director or deputy director for programmes) who is *accountable* for its completion, along with an action plan detailing the recommendations and follow-up (see next chapter for further details).
- That the Member or manager who commissioned / sponsored the evaluation is both responsible and accountable for submitting the report to the donor, where relevant. Country office leadership and the relevant thematic programme manager, as well as the regional office, should be copied in or informed that this has happened.

In declared emergency responses, the emergency team thematic programme manager fulfils the responsibilities of the country office thematic programme manager outlined above, in consultation with the emergency team's M&E staff (rather than country/regional office M&E staff). Similarly, in these instances, the emergency team leader is accountable for those areas mentioned above in relation to country office leadership.

5.9 Planning for action and communication

When the final evaluation report is ready, you need to distribute it and communicate the findings and recommendations in an appropriate way. Ideally, you will have developed a plan for communicating and sharing the findings from the evaluation elaborated in your M&E plan. There should also be an action plan for implementing the recommendations, which clearly sets out who is responsible for certain actions and who is responsible for monitoring. The next chapter takes you through using the results of your evaluation. It focuses on communicating the findings and following up on recommendations.



Checklist – Implementing your evaluation

Data collection

- ✓ Have you finalised the evaluation plan/inception report with stakeholders and the reference group, including (where relevant):
 - How to involve children and young people and other stakeholders
 - How to test your data collection tools
 - How to train the evaluation team to use those tools
 - A detailed schedule and plan for the data collection activities
 - A plan for communicating findings
- ✓ Have you carefully considered the issue of incentives?
- ✓ Have you ensured that the evaluation is aligned with and does not contradict SCI's intellectual property clause and data security guidelines?

Data analysis, interpretation and drawing conclusions

- ✓ Have you thought about how children and young people can be involved in analysing and interpreting the data?
- ✓ Have you assessed the evaluation report against the ToR and donor requirements?
- ✓ Is the language used in the evaluation report clear and easy to understand?
- ✓ Does the report respect ethical standards such as being mindful of cultural, religious and gender issues?
- ✓ Have you conducted a peer review process?

6 Using the results

Communicating and disseminating the findings of your evaluation is a vital part of how we make Save the Children more accountable to the children and communities we work with. In line with the evaluation Management Operating Standards and Standing Operating Procedures, all evaluation results should be published in an appropriate and accessible format within three months of approval of the final report. Moreover, it should include a plan of how you intend to disseminate and communicate the findings, as well as an action plan for implementing the recommendations.

You should aim to disseminate the findings and give two-way feedback at the community, district and/or national level. You must make sure you report back to all those stakeholders (including children) who were involved in the evaluation and the project or programme and give them the opportunity to respond. You should also report back to Save the Children International's Head Office, Regional Office, Global Initiatives and Members who have an interest in the project or programme.

Evaluation findings can be a very powerful tool to support Save the Children's broader advocacy and policy work.

It is important that you tell children and young people the outcome of their participation and how their contribution has been used. Where appropriate, you should give them the opportunity to reflect on their involvement and to share their views on how evaluation activities could be improved in future. Children and young people should also be an integral part of your strategy for communicating the evaluation findings. They can present the findings in ways that are accessible to other young people and adults of all ages, whether they are literate or not.

Here, we set out the steps you need to take to communicate and disseminate the findings of your evaluation report and follow up any recommendations.

6.1 Preparing a management response to the evaluation report

The management response to the evaluation report should be a clear and comprehensive document consisting of the following:

- *Key recommendations or issues:* Are the issues and recommendations proposed by the evaluation team relevant and acceptable to the country office?
- *Key actions:* What are the proposed actions? Who are the key partners in carrying out these actions?
- *Implementation:* Who is responsible for follow-up and implementation of these actions? What is the time frame for implementation?

The management response is often a public document and gives the project or programme team the opportunity to respond to recommendations made by the evaluation team. This is also a chance to clarify any issues that were unable to be resolved prior to the report's approval and sign-off.

The management response should be prepared by the evaluation manager within one month of approval of the final report, to ensure that the recommendations and learning are fresh in everyone's mind. The response is often used to enable the country office or commissioning body to confirm their agreement (or disagreement) with the evaluation findings. This is done in consultation with key internal stakeholders and the senior management team in the country office.

You can use your evaluation findings to:

- support evidence-based decision-making
- improve the quality of the project or programme
- encourage collection and dissemination of good practice
- support the learning process
- improve management systems
- promote participation
- develop staff and stakeholder capacities
- improve transparency
- promote coordination and harmonisation among all those involved.

An example is provided in Annex 5.

6.2 Developing an action plan

An action plan is developed for internal purposes and should be much more detailed than the management response. There will be some overlap, so it is for the country office to decide if they need both documents. However, an action plan is a requirement of Save the Children International and the Country Director is accountable for its follow-up and implementation.

- All evaluation reports must be accompanied by an action plan that has been completed by the relevant staff team and signed off by the Country Director.
- The action plan should look at the lessons learned during the evaluation and any recommendations made and consider how these will be addressed in the future.
- It should identify resource requirements, financial and/or technical.
- The action plan must be reviewed by senior management, shared with the regional office (for example, using the quarterly reports) and it should make clear who is responsible and accountable for which actions.

The action plan should be developed in conjunction with local partners involved in the project, where appropriate.



Enabling children to come up with solutions

A next step would be to work with the children involved in your evaluation to identify solutions to any problems and to see if they can be active participants in making positive change in their communities – for instance, through campaigning, peer education, developing their own project. Quality and meaningful participation leads to more ideas, plans and initiatives and to positive change. Children will get more out of the experience in the long term if they can be part of that process of empowerment and action.

6.3 Sharing the findings with external audiences and stakeholders

You will have identified your external audiences at the planning stage, but a few more might have come up during the course of the evaluation. Children and young people, their families and carers and their communities will be among your key stakeholders.

Sharing your findings

There are a number of ways you can share your evaluation findings, depending on the audience. Think about how children can be involved to help you reach other children and young people. Your options include:

- producing a child-friendly version of the report
- translating the report, or key findings and lessons learned, into local languages
- producing publications, annual reports, newsletters or bulletins
- developing a brief with a concise summary in plain language that can be widely circulated
- publishing an article for an academic journal
- distributing leaflets and posters detailing the key findings
- writing a press release
- giving media interviews
- organising meetings with policy-makers and politicians
- presenting the findings at annual stakeholder dissemination meetings and annual learning reviews (see Standard Operating Procedure on Data for Decision-Making and Organisational Learning)
- presenting and speaking at public meetings or conferences
- thinking of innovative and accessible ways to present the findings, such as drama, video, plays, poetry or photography.

In some cases, there may be reasons why you do not want to publish your evaluation findings. Doing so could compromise someone's safety and security, or might have a serious reputational risk for Save the Children or any of our partners, or indeed risk damaging our relationship with the local community, local or national government, or a particular donor. Save the Children International recommends that all evaluations are published both internally and externally, as per the Standard Operating Procedure, to promote greater transparency and learning. So if you have any concerns about confidentiality or other sensitive issues, we recommend that you discuss these with the Regional M&E Manager or appropriate technical adviser in the first instance.

Once you have disseminated your findings through various activities, you should monitor feedback and measure the outcome of the dissemination strategy. How far has the learning from the evaluation changed or influenced decision-making within the project or programme? Were the methods you used to disseminate the findings accessible for their intended audiences? Did each audience find the information useful and relevant to their needs? What could you do better next time?



Involving children in disseminating your findings

You could give children the chance to design and produce their own version of the evaluation report, making it child friendly. They can communicate the findings back to their peer groups and others involved in the evaluation process. A good way of doing this is to ensure that you have an event planned for children and young people, with other stakeholders as appropriate, where the children or young people most closely involved in the process can facilitate the review meeting.

6.4 Sharing the findings with internal audiences

Save the Children International is continuously striving to develop a culture of learning and improvement. We should be honest and open about our failures and challenges, as this is the best way to learn. It is also important that we celebrate our achievements and successes, so that they can be replicated and scaled up, according to our theory of change.

The only way we will really be effective in replicating and scaling-up our successes is if we are better at disseminating and sharing learning.

The findings, learning and information generated through the evaluation process should not get lost. We must all strive to translate the learning derived from evaluations into practical lessons that can be shared with other Country Offices and Global Initiative so that we do not keep reinventing the wheel.

Here are some of the ways you can help to share your evaluation findings internally:

- Share the final evaluation report and action plan with Save the Children International's Regional Office M&E Manager and the Member or Global Initiative that was funding or managing the grant for the project or programme.
- Organise or take part in peer learning events organised by Members and Global Initiatives. These could be global or regional face-to-face meetings or webinars, where country offices share findings and learning through 'virtual' presentations.

- Get a story linked to your evaluation into an internal bulletin or newsletter. These are good places to share lessons learned and programme successes and challenges.
- Global Initiatives and Members may also produce meta evaluations of a number of different evaluations in a country or region, or on a particular theme, to support cross-sectoral or regional learning throughout the wider Save the Children family.
- Lessons from evaluations will also be used to improve tools and guidance for best practice as we continuously strive to improve the design and implementation of all of our projects and programmes.

6.5 Using your findings to improve programme strategy and implementation

Your evaluation findings, whether derived at the end of the project or mid-term, should be fed back into the project or programme design. Project teams should adapt their activities and strategies according to the lessons learned. The Standard Operating Procedure on Data for Decision-Making and Organisational Learning identifies a number of ways this can be done at country office level. They include project team meetings and senior management meetings.

Checklist – Using the results

Communicating findings

- ✓ How will you report back to all the stakeholders (including children) who were involved in the evaluation and the project or programme?
- ✓ How will you report back to SCI Head Office, the Regional Office, Global Initiatives and Members?
- ✓ Have you discussed any concerns about confidentiality or other sensitive issues with the regional M&E manager or appropriate technical adviser before publishing the evaluation findings?
- ✓ Have you considered working with children to produce a child-friendly version of the report?
- ✓ Have you thought about the different internal and external forums where you can share the findings?

Following up recommendations

- ✓ Have you considered your management response?
- ✓ Have you ensured that an action plan is developed and learning is fed back into the way the project or programme is designed and implemented?
- ✓ Have you considered how to involve children in identifying solutions to issues outlined in the recommendations?



Providing feedback to children

Remember to provide feedback to the children involved and to discuss what changes may or may not arise from the evaluation. Also remember to thank the children.



ANNEX I: DAC evaluation criteria

When evaluating programmes and projects, it is useful to consider the following DAC criteria (Development Assistance Committee of the Organisation for Economic Co-operation and Development), as laid out in the *DAC Principles for Evaluation of Development Assistance*. It is not appropriate to investigate every criterion in depth in every evaluation. However, you should provide an explanation of the criteria you have chosen to cover in your evaluation approach.

Relevance

The extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor.

In evaluating the relevance of a programme or a project, it is useful to consider the following questions:

- To what extent are the objectives of the programme still valid?
- Are the activities and outputs of the programme consistent with the overall goal and the attainment of its objectives?
- Are the activities and outputs of the programme consistent with the intended impacts and effects?

Effectiveness

A measure of the extent to which an aid activity attains its objectives. In evaluating the effectiveness of a programme or a project, it is useful to consider the following questions:

- To what extent were the objectives achieved/are likely to be achieved?
- What were the major factors influencing the achievement or non-achievement of the objectives?

Efficiency

Efficiency measures the outputs – qualitative and quantitative – in relation to the inputs. It is an economic term which signifies that the aid uses the least costly resources possible in order to achieve the desired results. This generally requires comparing alternative approaches to achieving the same outputs, to see whether the most efficient process has been adopted.

When evaluating the efficiency of a programme or a project, it is useful to consider the following questions:

- Were activities cost-efficient?
- Were objectives achieved on time?
- Was the programme or project implemented in the most efficient way compared to alternatives?

Impact

The positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended. This involves the main impacts and effects resulting from the activity on the local social, economic, environmental and other development indicators. The examination should be concerned with both intended and unintended results and must also include the positive and negative impact of external factors, such as changes in terms of trade and financial conditions.

When evaluating the impact of a programme or a project, it is useful to consider the following questions:

- What has happened as a result of the programme or project?
- What real difference has the activity made to the beneficiaries?
- How many people have been affected?

Sustainability

Sustainability is concerned with measuring whether the benefits of an activity are likely to continue after donor funding has been withdrawn. Projects need to be environmentally as well as financially sustainable. When evaluating the sustainability of a programme or a project, it is useful to consider the following questions:

- To what extent did the benefits of a programme or project continue after donor funding ceased?
- What were the major factors that influenced the achievement or non-achievement of sustainability of the programme or project?

The following additional three criteria are also useful. They were developed by the DAC for use in humanitarian evaluations, but can apply in other contexts.

Coverage looks at which groups are included in or excluded from a programme, and the differential impact on those included and excluded. (*Who and how many people are we reaching?*)

Coherence is the need to assess other policies and programmes which affect the intervention being evaluated.

Coordination includes assessing both harmonisation with other aid agencies and alignment with country priorities and systems. (*How are we working with others?*)

ANNEX 2: What you need to know about sampling²¹

1. What is sampling?

When you begin planning a data collection strategy, you have to decide whether it is possible or desirable to collect data from the entire population you intend to study. Collecting data from everyone in a defined study population is called a census. In most programme countries, we are unable to, or do not need to, collect data from everyone in our defined study population because the population can be very large. Instead we take a sample – a subset of the population. Depending on how we select a sample, we may be able to make inferences about the whole population based on our sample results – this is called generalising our results to the population.

Sampling is used for both quantitative and qualitative data collection methods. When selecting a sample, you should always consider the following key questions:

- What is the group of children and/or adults (study population) we are interested in, from which we want to draw a sample?
- How many children and/or adults do we need in our sample?
- How will these children and/or adults be selected?

The study population is very important and must be clearly defined (for example, according to age, sex, and residence). Apart from children and adults, a study population may consist of villages, institutions, records, committees, etc.

2. Sampling methods

Sampling methods are different depending on the purpose of your data collection and whether you are using quantitative or qualitative methods. There are two main methods for selecting a sample: random sampling and non-random sampling. Random sampling is usually associated with quantitative methods, and non-random sampling is usually associated with qualitative methods.

2.1 Random sampling (also called probability sampling)

Random sampling allows us to produce a sample that can be considered to be representative of the study population and so allowing us to generalise our findings. If the aim of our study is to measure variables distributed in a population (for example, school attendance rates, child labour rates) we need to use random sampling. Random sampling involves using random selection procedures to ensure that each unit of the sample is chosen on the basis of chance. All units of the study population should have an equal, or at least a known chance of being included in the sample.

The first step of random sampling is to list all study units (villages, schools, or children, for example) in the defined population. This list is called the sampling frame. Without a complete sampling frame, it is not possible to conduct random sampling. Once you have the sampling frame, there are five main ways of randomly selecting your sample from the list:

²¹ This guide is adapted from a sampling manual produced by INTRAC for Save the Children's programme in Pakistan. It draws heavily from the International Development Research Centre's (IDRC) module on sampling.

1 Simple random sampling

This is the simplest form of probability sampling. To select a simple random sample you need to:

- Make the sampling frame list
- Decide on the sample size (this is discussed below)
- Select the required number of sampling units, using a 'lottery' method (pulling numbers out of a box) or a table of random numbers.

Example of simple random sampling

A simple random sample of 50 students is to be selected from a school of 250 students. Using a list of all 250 students, each student is given a number (1 to 250), and these numbers are written on small pieces of paper. All 250 papers are put in a box, after which the box is shaken vigorously, to ensure randomisation. Then 50 papers are taken out, and the numbers are recorded. The students corresponding to these numbers will constitute the sample.

2 Systematic sampling

In this method, individuals are chosen at regular intervals (for example, every fifth) from the sampling frame list. Ideally, we randomly select a number to tell us where to start selecting individuals from the list.

Example of systematic sampling

A systematic sample is to be selected from 1,200 students of a school. The sample size selected is 100. The sampling fraction is:

$$\frac{100 \text{ (= sample size)}}{1200 \text{ (= study population)}} = \frac{1}{12}$$

The sampling interval is therefore 12. The number of the first student to be included in the sample is chosen randomly, by blindly picking one out of twelve pieces of paper, numbered 1 to 12. If number 6 is picked, then every 12th student will be included in the sample, starting with student number 6, until 100 students are selected: the numbers selected would be 6, 18, 30, 42, etc.

3 Stratified sampling

In this method, the population sampling frame is first divided into different sub-groups based on predetermined characteristics of interest (for example, children from urban and rural areas, or different religious or ethnic groups, working boys, girls in school and girls out of school). Then a random or systematic sample is selected from each sub-group.

An advantage of stratified sampling is that a relatively large sample can be taken from a small group in a study population. This allows a sample to be obtained that is big enough to enable valid conclusions are drawn about a relatively small group without having to collect an unnecessarily large (and hence expensive) sample of the other, larger groups. However, in doing so, unequal sampling fractions are used and it is important to correct for this when generalising findings to the whole study population through weighting.

Example of stratified sampling

A survey is conducted on household water supply in a district comprising 20,000 households, of which 20% are urban and 80% are rural. It is suspected that in urban areas, access to safe water sources is much more satisfactory. A decision is made to include 100 urban households (out of 4,000, which give a 1 in 40 sample) and 200 rural households (out of 16,000, which give a 1 in 80 sample). Because we know the sampling fraction for each group, access to safe water for all the district households can be calculated after the study (by multiplying the findings for the urban households by 40 and those for the rural households by 80, and then calculating statistics for the total sample).

4 Cluster sampling

It may be difficult or impossible to take a simple random sample of the units of the study population at random, because a complete sampling frame (eg, a list of every single school) does not exist. Logistical difficulties may also discourage random sampling techniques (eg, interviewing people who are scattered over a large area may be too time-consuming). However, when a list of groupings or clusters of study units is available (eg, villages or schools) or can be easily compiled, a number of these groupings can be randomly selected. The selection of groups of study units (clusters) instead of the selection of study units individually is called cluster sampling. Clusters are often geographic units (eg, districts, villages) or organisational units (eg, school management committees, schools).

Example of cluster sampling

In a study of the knowledge, attitudes and practices (KAP) of religious leaders and village headmen related to children's rights in rural communities of a region, a list is made of all the districts. Using this list, a random sample of districts is chosen and all religious leaders and headmen in the villages of those districts are interviewed.

5 Multi-stage sampling

In very large and diverse populations, sampling may be done in two or more stages. This is often the case in community-based studies, in which people are to be interviewed from different villages, and the villages have to be chosen from different areas. A multi-stage sampling procedure is carried out in phases and it usually involves more than one sampling method.

The main advantages of cluster and multi-stage sampling are that a sampling frame (a list, for example) of every single individual unit is not required for the whole population. Only within the clusters that are finally selected is there a need to list and sample the individual units. Also, the sample is easier to access than a simple random sample of similar size, because the individual units in the sample are physically together in groups, instead of scattered all over the programme area. However, the main disadvantage of this type of sampling is that compared to simple random sampling, there is a larger probability that the final sample will not be representative of the total study population. The likelihood of the sample not being representative depends mainly on the number of clusters that are selected in the first stage. The larger the number of clusters, the greater is the likelihood that the sample will be representative. Further, the sampling at community level should be properly random.

Example of multi-stage sampling

For an evaluation of a national education programme, it is decided to interview a randomly selected sample of 1,000 children from all over the country using the following method:

- 1 Decide that there will be equal numbers of children interviewed from each province the project is working in (say five provinces): 200 children in each.
- 2 For each province, select two districts by simple random sampling (10 districts in total).
- 3 For each district, decide on five villages, again by simple random sampling (50 villages).
- 4 You now need to find 20 children in each village to interview using a random system to select them. Since simply choosing households in the centre of the village would produce a biased sample, the following sampling procedure is proposed:
 - Go to the centre of the village.
 - Choose a direction in a random way: spin a bottle on the ground and choose the direction the bottleneck indicates.
 - Walk in the chosen direction and select every (or, depending on the size of the village, every second or every third) household until you have the 20 children you need. If you reach the boundary of the village and you still do not have 20 children interviewed, return to the centre of the village, walk in the opposite direction and continue to select your sample in the same way until you have 20. If there is nobody in a chosen household, take the next nearest one.

2.2 Non-random sampling

Non-random sampling methods should be used when you want to gather perspectives from specific people or groups, or when random sampling is not possible. Non-random sampling methods are usually associated with qualitative data collection methods. They are typically used when focusing on a smaller number of informants who are selected strategically because of their specific perspectives.

Non-random sampling should never be haphazard. It should be done with the same care and planning as random sampling. Selection rules should be developed to prevent the evaluator from sampling according to personal bias.

Non-random sampling can provide very interesting and rich data and is important for evaluations and studies that want to highlight different factors and different perspectives. However, it does not provide representative data for the total defined population. Representativeness is not usually of concern in qualitative data collection methods.

There are four main methods for non-random sampling:

1 Extreme case sampling

This method involves selecting individuals who are extreme cases or examples of study interest. This method is a rapid strategy to identify factors contributing to certain behaviours and to enable comparison between people or groups.

Example of extreme case sampling

In a study to learn about factors contributing to child labour, a project team decides to interview (i) girls aged 18 who have never attended school but have always worked, and (ii) girls who have attended 11 years of schooling and have never worked.

2 Quota sampling

If an evaluator wants to obtain a picture of a certain issue for specific groups of people, quota sampling can be used. This method involves identifying groups of people of interest and interviewing predetermined numbers of people in each group.

Example of quota sampling

In a study on attitudes towards children with disabilities attending school, an evaluation team decides to interview 25 each of: carers of children with disabilities; carers of children without disabilities; children with disabilities; teachers; and village health workers. This enables them to get a range of perspectives on the issue. The individuals in each group are chosen purposefully to represent a range of characteristics, such as rural/urban residence, age, and involvement in the programme.

3 Typical case sampling

This method involves identifying specific cases which are believed to be 'typical' for the group that are of interest. Examples include:

- The education and qualifications of a 'typical' teacher
- A 'typical' child whose education was interrupted by the earthquake
- 'Typical' health problems of children working long hours on carpet weaving
- 'Typical' facilities and equipment in rural schools.

Such descriptions or case studies give a good picture, but they cannot be generalised for the whole group. Typical examples can either be selected with the cooperation of key informants who know the study population well, or from a survey that helps to identify what is normal.

4 Snowball or chain sampling

This method is used when you do not know who to include in your sample, or when you have an ‘invisible’ or ‘hard to reach’ study population. In this approach, you start by identifying one or two information-rich key informants and ask them if they know other people (who know a lot about your topic of interest) who they recommend you talk to. You continue doing this for every new contact you make until no new suggestions are obtained.

3. Calculating the desired sample size

Once you have decided how to select your sample, you need to determine the desired sample size. The methodology for calculating a desired sample size differs depending on whether you are using quantitative or qualitative data collection methods.

3.1 Sample size in quantitative studies

There are many factors that affect the desired sample size for quantitative data collection methods. These are summarised in the table below.

Factors affecting the sample size		
Factor	Explanation	Influence on sample size
The purpose of the evaluation	Is this an exploratory study or are very precise statistical estimates required?	The more precise the required results, the larger the sample
Is the direction of the expected change known? (Will a one or two-tailed test be used?)	If the purpose of the evaluation is to test whether positive outcomes have increased, or negative ones have declined, then a one-tailed test can be used. If the purpose is to test whether there has been “a significant change” without knowing the direction, then a two-tailed test is required	The sample size will be approximately 40% larger for a two-tailed test
Is only the project group interviewed?	In some evaluation designs, only subjects from the project group are interviewed. This is the case if information on the total population is available from previous studies or secondary data. In other cases, a comparison group must also be selected and interviewed	The sample size will be doubled if the same number of people have to be interviewed in both the project and comparison groups
Homogeneity of the group	If there is little variation among the population with respect to the outcome variable, then the standard deviation will be small	The smaller the standard deviation, the smaller the sample
The effect size	Effect size is the amount of increase the project is expected to produce	The smaller the expected effect size, the larger the sample
The efficiency with which the project is implemented	While some projects are implemented in a very efficient way with all subjects receiving exactly the same package of services, in other cases the administration is poorer and different subjects receive different combinations of services. The quality of the services can also vary	The poorer the quality and efficiency of the project, the larger the sample required

Factors affecting the sample size		
Factor	Explanation	Influence on sample size
The required level of disaggregation	In some cases, only global estimates of impact are required for the total project population. In other cases, it is necessary to provide disaggregated results for different project sites, variations in the package of services provided, or for different socio-economic groups (sex, age, ethnicity, etc). As a rule of thumb, you should have at least 10 people for each sub-group you require	The greater the required disaggregation, the larger the sample
The sample design	Sampling procedures such as stratification can often reduce the variance of the estimates and increase precision	Well-designed stratification may reduce sample size
The level of statistical precision	When estimating whether the project had an impact “beyond a reasonable doubt”, this is normally defined as “less than a 1 in 20 chance that an impact as large as this could have occurred by chance”, that is, at the 0.05 confidence level. If more precise results are required, the 0.01 level may be used (less than 1 in 100 chance). For an exploratory study, the 0.10 level may be used (a 1 in 10 chance).	The higher the confidence level, the larger the sample
The power of the test	The statistical power of the test refers to the probability that when a project has “real” effect, this will be rejected by the statistical significance test. The conventional power level is 0.8, meaning that there is only a 20% chance that a real effect would be rejected. Where a higher level of precision is required, the power can be raised to 0.9 or higher	The higher the power level, the larger the sample
Finite population correction factor	If the sample represents more than say 5% of the total population, it is possible to reduce the sample size through the finite population correction factor	The greater the proportion the sample represents of the total population, the smaller the sample

A bigger sample is not necessarily always better. In general, it is better to increase the accuracy and richness of data collection (for example, by improving the training of interviewers or by better pre-testing of the data collection tools) than to increase sample size after a certain point. Also, it is better to make extra efforts to get a representative sample rather than to get a very large sample.

Once you have calculated the desired sample size, it cannot always be achieved for lack of resources such as time, human resources and money. This constraint applies to quantitative as well as qualitative studies. Therefore, the eventual sample size is usually a compromise between what is desirable and what is feasible.

3.2 Sample size calculations

In quantitative studies, where you want to show that your findings are statistically significant,²² you need to do a sample size calculation. There are websites that will do this for²³. You need to do this when you want to:

- Measure one single variable (for example, a mean, a rate or a proportion) in one group with a certain precision, or
- Demonstrate a significant difference between two groups. In this type of study, the sample size depends primarily on the estimated size of the difference between the two groups that are compared. The larger the difference, the smaller will be the sample that is needed to show this difference. Second, the sample size depends on how large we want the probability to be that we indeed will find a significant difference.

The formulae can only be used if you have a rough idea about the outcome of the study, which is not always the case. It is always advisable to call upon a statistician or an experienced researcher who can help you to choose and use the appropriate formulae.

²² Statistically significant means highly likely to be “true”.

²³ www.surveysystem.com - sample size calculator.

Example of a sample size calculation

In a certain province, to estimate the proportion of the child population that is engaged in harmful agricultural work. These are the steps:

- 1 Identify the key statistic you want to estimate (eg, % of children engaged in harmful agricultural work).
- 2 Estimate how big the proportion might be. You could use secondary data from official sources or other studies. Where no estimate is available, you should estimate at 50%.
- 3 Choose the margin of error you will allow in the estimate of the proportion (say ± 5 percentage points). This means that, if the survey reveals that 80% of the children are engaged in harmful agricultural work, the true proportion is probably somewhere between 75% and 85% in the whole study population from which the sample was drawn. The smaller your margin of error (also called confidence interval), the bigger your sample size needs to be. In general, you need more precision (or a smaller margin of error) if the estimated proportion is very small.
- 4 Choose the level of confidence at which you want to be able to state that the child labour rate in the child population is indeed between 75% and 85%. You can never be 100% sure. Do you want to be 95% sure? or 99%? A commonly used confidence level is 95%. If you would like to increase your confidence level from 95 to 99%, the sample would have to be bigger, but this will cost more.
- 5 Use an online web source to calculate your sample size based on the specifications decided in the steps above.
- 6 The website will use a formula such as the one below to calculate the desired sample size. Note that there are different formulae that should be used depending on whether you are estimating a single variable or trying to estimate the difference between two variables. Also, there are different formulae for estimating a mean value, a rate and a proportion. Always consult a statistician or experienced researcher for guidance on the appropriate formula to use. In the formulae below, the following abbreviations are used:

n = sample size

e = required size of standard error

p = percentage

The proportion of children working in harmful agricultural work is estimated to be 30% (p) from a national study from 2006. A new study using simple random sampling aims to provide updated data and wants to determine the percentage of children working in harmful agricultural work with a confidence interval of 25% to 35% (margin of error of ± 5 percentage points) and a precision of 95%. The standard error is therefore 2.5% (e) (half the size of the margin of error if a precision of 95% is required). The required sample size would be calculated using the formula for a single proportion as follows:

n (sample size) = $p(100-p) = 30 \times 70 = 336$ children

e^2

e^2

2.5^2

3.3 Sample size in qualitative studies

There are no fixed rules or formulae for determining the sample size in qualitative research. The size of the sample depends on what you try to find out, and who from (the different informants or perspectives you seek). One approach is to collect data until the point of saturation – ie, when no new data is coming up. This means that you get to a point where every new person that you speak to tells you something that you have already heard from other informants – no new messages or perspectives are expressed.

Example of a qualitative study sample size approach

If you want to explore how you can involve parents in your school catchment area more effectively in making sure their children attend school regularly, you might decide to conduct some focus group discussions to assess mothers' attitudes and practices with respect to school attendance.

You could start with two focus group discussions among low-educated mothers and two among mothers with more education. If the different data sets reconfirm each other, you may stop at this point and start a small-scale project; otherwise you conduct more focus group discussions until you reach the point when no new data comes up.

If your research objective is more complex than the example above, you may need a bigger sample that collects data from a number of different sub-groups. You might start with four focus group discussions (FGDs), two among boys and two among girls, subdivided according to their ethnic background. You might also want, say, another four FGDs with children who have some kind of disadvantage and whose voice is rarely heard: children with disabilities or older children who are starting primary school five or six years late, for example. Depending on how different their responses are, you might want to conduct a few more FGDs to explore whether the differences were just chance or confirmed by others. You could also run some more FGDs with children in rural areas and those in towns to see if there are differences there. In qualitative data collection, the sample size is therefore estimated beforehand as precisely as possible, but not determined. However, the richness of the data, quality of the data collected, and the skill of the evaluator in analysing data are more important in producing valid and meaningful information than sample size.

Sampling procedures and sample size should always be carefully explained in an evaluation report for both qualitative and quantitative methods in order to clearly identify the process and potential sources of bias in the findings.

4. Bias in sampling

Bias in sampling is a systematic error in sampling procedures which leads to a distortion in the results of the study or evaluation. Bias can also be introduced as a consequence of improper sampling procedures, which result in the sample not being representative of the study population. Bias can also occur as a result of faulty data collection tools, or poor accuracy in recording data.

There are several possible sources of bias that may arise when sampling. The most well-known source is non-response. Non-response can occur in any interview situation, but it is mostly encountered in large-scale surveys with self-administered questionnaires. Respondents may refuse or forget to fill in the questionnaire. The problem lies in the fact that non-respondents in a sample may exhibit characteristics that differ systematically from the characteristics of respondents. There are several ways to deal with this problem and reduce the possibility of bias:

- Data collection tools (including written introductions for the interviewers to use with potential respondents) should be pre-tested. If necessary, adjustments should be made to ensure better participation.
- If non-response is due to absence of the subjects, follow-up of non-respondents may be considered.
- If non-response is due to refusal to cooperate, an extra, separate study of non-respondents may be considered in order to identify to what extent they differ from respondents.
- Another strategy to account for the likelihood of non response and/or attrition may be to increase the size of your sample. Increasing the size of your sample can only be justified if the increased sample is representative of the study population.

The bigger the non-response rate, the more necessary it becomes to take remedial action. It is important in any evaluation to mention the non-response rate and to honestly discuss whether and how the non-response might have influenced the results.

Other sources of bias in sampling are:

- Studying volunteers only. The fact that volunteers are motivated to participate in the study may mean that they are also different from the study population on the factors being studied. Therefore, it is better to avoid using non-random selection procedures that introduce such an element of choice.
- Sampling of children in school only. To get representative information about what motivates children to attend school, or what factors keep them away, it is not sufficient to interview only children who are at school. You need to ensure that children who are working or kept at home are also interviewed.
- Missing cases of short duration. Short absences from a population are also important to follow up. For example, in a study of corporal punishment in schools, it is important to ensure that children who are off school on a particular day are not excluded from the survey. It may be that those are the children at greatest risk of harm and that they are away because they are afraid to go on that day or have injuries which prevent them attending.
- Seasonal bias. Weather, harvest season and migration patterns can affect who will be in a study population at a certain time. The cyclical movements of children and their families should be taken into account in any evaluation.
- Tarmac bias. Study areas are often selected because they are easily accessible by vehicle. However, these areas are likely to be systematically different from more inaccessible areas.

5. Further resources

5.1 IDRC (*International Development Research Centre*)

Provides more advanced sampling topics. See: http://web.idrc.ca/en/ev-1-201-1-DO_TOPIC.html

The website also includes other resources on evaluation, including descriptions of the technique of outcome mapping for assessing impact of a project.

5.2 *The Survey System*

This is a commercial website that companies use to do surveys for them. They give simple explanations of key information about survey design, questionnaires, the meaning of 'significance' in statistics and also: a sample size calculator.

5.3 *IFAD (International Fund for Agricultural Development)*

The International Fund for Agricultural Development (IFAD) website includes a very useful and comprehensive publication, *A Guide for Project M&E*, which includes information on sampling. See: www.ifad.org/evaluation/guide

ANNEX 3: A template for drawing up your terms of reference

The planning and design issues related to a mid-term review or evaluation can be summed up in a terms of reference (ToR). The ToR should cover WHAT we want to evaluate, WHY, and HOW it should be organised. A ToR can also serve as a management tool throughout the evaluation process and a reference against which the completed evaluation report and deliverables are assessed before approval.

The scope and level of detail of a ToR might vary; in some cases, the ToR will entail a plan for the entire evaluation process, while in others, ToRs can be developed for specific parts of the evaluation – eg, for the evaluation team or even individual team members. Furthermore, the scope and level of detail in the ToR should be consistent with the size, complexity and strategic importance of the project or programme being evaluated.

Below you will find a suggested inventory for a rather comprehensive ToR, covering the entire evaluation process. In some cases, a donor might also require other issues to be covered.

What a ToR should include:

1 Background

Description of the project or programme to be evaluated, including: objectives, indicators, baseline and monitoring data, logical/results framework, theory of change/intervention logic, target groups, partners, context, identified challenges, etc.

2 Purpose of the evaluation

Why are you evaluating the project or programme, what will the evaluation be used for, and who will use it? What additional information will an evaluation bring as a complement to the monitoring data?

3 Objectives and key evaluation questions

Two or three objectives (assess outcome/impact/document good practice/ understand causal relations/what works and why/formation of future advocacy, policy or programme) with key evaluation questions for each objective: what is it that you specifically want to know?

4 Scope of the evaluation

Which issues will be covered by the evaluation, and which issues or units fall outside its scope? Note selection criteria when focusing some geographical areas, particular target groups etc.

5 Evaluation design and methodology – Child participation is key

If it is an internal evaluation, the methodology and process steps should be quite detailed, especially if the internal evaluation team leader is involved already.

If it is an external evaluation, a methodology could be suggested, and key Save the Children principles/requirements stated but details left for the evaluation team to suggest. Make sure there is a clear requirement to ensure the meaningful and safe participation of children and young people. Other issues to cover could be: design workshops with stakeholders; review and approval process; data collection including translation, logistics; criteria for sampling (be it projects, project sites, partners, districts, etc)

6 Organisation, roles and responsibilities

- Evaluation team, internal or external, roles and responsibilities, criteria for external consultant(s)/researcher(s).
- Save the Children International's internal organisational set-up for the evaluation, with roles and responsibilities, and approval point including SMT involvement and other stakeholder involvement
- Access to documents and data
- Reference documents
- Data sources

7 Deliverables

- Inception report/detailed work plan for the evaluation to be approved by team leader or Steering Group
- Draft and final evaluation report(s), page limit, executive summary Debrief
- Presentation at workshops or conferences

8 Timeline

- For each step and deliverable.

9 Budget/Resources

- Human resources; budget; logistics.

10 Plan for dissemination and learning

- Feedback to stakeholders, particularly children, who were involved in the evaluation
- Plans for publication(s), briefing documents, lessons learned documents, communication materials, and presentation at workshops/conferences.

A plan for follow-up will be required within three months of approval of the final evaluation report; Global Initiatives Director/Country Director/Programme Manager responsible.

ANNEX 4: Quality review guide²⁴

	Appraisal questions	Quality indicators (possible features for consideration)
Findings	1 How credible are the findings?	Findings/conclusions are supported by data/study evidence (<i>ie, the reader can see how the researcher arrived at his/her conclusions</i>) Findings/conclusions “make sense” and have a coherent logic. Findings/conclusions are resonant with other knowledge and experience.
	Notes	
Findings	2 How has understanding been extended by the research?	Literature review summarising knowledge to date/key issues raised by previous research. Aims and design of study set in the context of existing knowledge/understanding; identifies new areas for investigation. Credible/clear discussion of how findings have contributed to knowledge and understanding; might be applied to new policy developments, practice or theory.
	Notes	
Findings	3 How well does the work address its original aims and purpose?	Clear statement of study aims and objectives; reasons for any changes in objectives. Findings clearly linked to the purposes of the study – and to the initiative or policy being studied. Summary or conclusions directed towards aims of study. Discussion of limitations of study in meeting aims.
	Notes	
Design	4 How defensible is the research design?	Discussion of how overall research strategy was designed to meet aims of study. Discussion of rationale for study design. Use of different features of design/data sources evident in findings presented. Discussion of limitations of research design and their implications for the study evidence.
	Notes	
	Appraisal questions	Quality indicators (possible features for consideration)
Sample	5 How well defended is the sample design/target selection of cases/documents?	Description of study locations/areas and how and why chosen. Description of population of interest and how sample selection relates to it (<i>eg, typical, extreme case, diverse constituencies, etc.</i>). Discussion of how sample/selections allowed required comparisons to be made.

²⁴ Adapted from L. Spencer, J. Ritchie, J. Lewis, and L. Dillon, 'Quality in Qualitative Evaluation: A framework for assessing research evidence', Government Chief Social Researcher's Office, 2003, see: http://www.civilservice.gov.uk/wp-content/uploads/2011/09/Quality-in-qualitative-evaluation_tcm6-38739.pdf

	Appraisal questions	Quality indicators (possible features for consideration)
	Notes	
6	How well was the data collection carried out?	<p>Discussion of:</p> <ul style="list-style-type: none"> • who conducted data collection • procedures/documents used for collection/recording • checks on origin/status/authorship of documents <p>Discussion of how fieldwork methods or settings may have influenced data collected.</p> <p>Demonstration, through portrayal and use of data, that depth, detail and richness were achieved in collection.</p>
Data Collection		
	Notes	
7	Contexts of data sources – how well are they retained and portrayed?	<p>Description of background or historical developments and social/organisational characteristics of study sites or settings.</p> <p>Explanation of origins/history of written documents.</p> <p>Use of data management methods that preserve context (<i>ie, facilitate within case description and analysis</i>).</p>
Analysis		
	Notes	
8	How well has diversity of perspective and content been explored?	<p>Discussion of contribution of sample design/case selection in generating diversity.</p> <p>Description and illumination of diversity/multiple perspectives/alternative positions in the evidence displayed.</p> <p>Evidence of attention to negative cases, outliers or exceptions.</p>
Analysis		
	Notes	

ANNEX 5: A template for your management response plan

Country name:

Prepared by:

Country Office:

Evaluation recommendation or Issue 1:				
Management response:				
Key action(s)	Time frame	Responsible unit(s)	Tracking	
			Status	Comments
I.1				
I.2				
I.3				
Evaluation recommendation or Issue 2:				
Management response:				
Key action(s)	Time frame	Responsible unit(s)	Tracking	
			Status	Comments
2.1				
2.2				
2.3				

ANNEX 6: How to communicate statistics²⁵

Being able to turn data into information or communicate statistical information accurately is vital for effective decision-making. Here, we provide guidance on how to write statistical commentary and use tables and graphs to communicate statistical findings.

Writing about data

Statistics help you to present your analysis in a way that tells a story about the data. In effect, statistical writing can bring data to life, making it real, relevant and meaningful to the audience. When communicating statistical information, you should ensure that the information you present is clear, concise and accurate. It is also important to provide contextual information and to draw out the main relationships, causations and trends in the data.

Here are some useful tips to help you write about statistics:

- Describe the context for the topic
- Present the complete picture to avoid misrepresentation of the data
- Accurately convey the main findings clearly and concisely
- Include definitions to support correct interpretations of the data
- Where necessary, include information on how the data was collected, compiled, processed, edited and validated
- Include information on data quality and data limitations
- Use plain, simple language and, where possible, minimise the use of jargon
- Ensure that information and data are accurate
- Where possible, avoid using data with quality concerns
- Use tables and graphs to present and support your written commentary.

Here are some other tips to help you ensure that the statistics you provide are accurate, and easy to read and understand:

- Avoid subjective language or descriptions (eg, slumped to 45%)
- Statements should be backed up by the data (eg, a greater proportion of 0–14-year-olds identified as indigenous than 15–24-year-olds (5.8% compared with 4.1%))
- Use proportions to improve flow and ease of comprehension (eg, nearly three-quarters (73%) of females)
- Use rates when comparing populations of different sizes (eg, age-specific death rate, crime rates)
- A percentage change (the relative change between two numbers) is different from a percentage point change (the absolute difference between two percentages)

²⁵ Adapted from the Australian Bureau of Statistics (2010), A Guide for Using Statistics for Evidence Based Policy, chapter on 'Communicate statistical findings', see: www.abs.gov.au/ausstats/abs@.nsf/lookup/1500.0chapter92010

- Be careful of percentage change and small numbers (eg, the region experiences a 100% increase in the number of reported crimes (from two reported incidences in 2002 to four reported incidences in 2003))
- Figures should always be written as numbers (eg, 45% instead of forty-five per cent)
- Comparison of large numbers can be improved by using a different scale
- Rounded figures are used in text, and raw data in tables.

Using tables, graphs and maps to communicate statistical findings

Whether you are writing a report or making a presentation, the story should be told by your evidence. A simple table, graph or map can explain a great deal, so this type of direct evidence should be used where appropriate. However, if a particular part of your analysis represented by a table, graph or map does not add to or support your argument, it should be left out.

While presenting statistical information in tables, graphs or maps can be highly effective, it is important to ensure that the information is not presented in a manner that can mislead the reader. The key to presenting effective tables, graphs or maps is to ensure they are easy to understand and clearly linked to the message. Ensure that all the necessary information required to understand what the data is showing is provided, as the table, graph or map should be able to stand alone.

Tables, graphs and maps should:

- relate directly to the argument
- support statements made in the text
- summarise relevant sections of the data analysis
- be clearly labeled.

Using tables to communicate statistical findings

An effective table does not simply present data to the audience, it supports and highlights the argument or message being presented in the text, and helps to make the meaning of that message clear, accessible, and memorable for the audience.

You may find this checklist useful when creating tables.

- Label each table separately
- Use a descriptive title for each table
- Label every column
- Provide a source if appropriate
- Provide footnotes with additional information required for understanding the table
- Minimise memory load by removing unnecessary data and minimising decimal places
- Use clustering and patterns to highlight important relationships
- Use white space to effect

- Order data meaningfully (eg, rank highest to lowest)
- Use a consistent format for each table.

It is also very important not to present too much data in tables. Large expanses of figures can be daunting for a reader, and can actually obscure your message.

Using graphs to communicate statistical findings

Graphs provide a way to visually represent and summarise complex statistical information. They are especially useful for revealing patterns and relationships that exist in the data and for showing how things may have changed over time. A well-placed graph may also be useful in improving readability by breaking up large chunks of text and tables.

There are a range of different graphs used for presenting data, such as bar graphs, line graphs, pie graphs and scatter plots. It is important to use the right type of graph for presenting the information.

Effective graphs are easy to read and clearly present the key messages. Here are some pointers if you are using graphs in your evaluation report.

- **Title:** Use a clear descriptive title to properly introduce the graph and the information it contains.
- **Type of graph:** Choose the appropriate graph for your message; avoid using 3D graphs as they can obscure information.
- **Axes:** Decide which variable goes on which axis, and what scale is most appropriate.
- **Legend:** If there is more than one data series displayed, always include a legend – preferably within the area of the graph – to describe them.
- **Labels:** All relevant labels should be included, including thousands or percentages, and the name of the ‘x’ axis if required.
- **Colour/shading:** Colours can help differentiate, but know what is appropriate for the medium you are using.
- **Footnotes:** These can help communicate anything unusual about the data, such as limitations in the data, or a break in the series.
- **Data source:** Where appropriate, provide the source of data you have used for the graph.
- **3/4 Rules:** For readability, it is generally a good rule of thumb to make the ‘y’ axis 3/4 the size of the ‘x’ axis.

Using maps to communicate statistical findings

A map can often convey a message more concisely than words. Using maps to present statistical information about a geographic area can provide a quick overview of what a dataset is showing and highlight the patterns and relationships in different regions.

When presenting statistical information in a map format, make sure that you label each map correctly: include a legend; provide a scale; and include all contextual information to assist with understanding the data, and any limitations there may be.

To produce a good map, you need to include:

- a prominent, clear title
- a clear, self-explanatory legend
- a neat, uncluttered layout
- if it is a thematic map, unobtrusive but useful topographic detail
- an explanation of the detail, accuracy and currency of the data
- an easily understood scale bar
- an acknowledgement of who produced the map.

ANNEX 7: How to involve children and young people in evaluations

This section outlines some ideas of how children and young people can be involved in the evaluation process. It is not intended to be a comprehensive step by step guide. Instead it is intended to highlight some of the possible entry points where children and young people can be involved. All the activities listed in this annex can either be lead by children and young people, or serve as a participatory activity for adults to gather information about a project from children. It is not suggested that an evaluation team should use all the activities listed in this annex, but to collaboratively decide which activities are best suited to the evaluation being conducted.

How to ... Prepare children and young people to be part of an evaluation

It is really important to ensure that time is spent with the children and young people to adequately prepare them to be part of the evaluation process. A good way to do this is through holding an Introductory Workshop, which can be run in parallel with the evaluation inception workshop. The Introductory Workshop should

- Give children and young people the opportunity to fully understand the project being evaluated
- Give children and young people an overview of what the objectives of the evaluation are and the different areas where they can be involved
- Give children and young people an opportunity to understand the themes of the evaluation and decide which themes they are interested in being involved in
- Enable the evaluation team to collectively decide which aspects of the evaluation process children and young people could be part of as well as the aspects that might be inappropriate for them to participate in
- Clarify the Child Safeguarding Mechanisms that will be in place throughout the evaluation
- Clarify expectations, and roles and responsibilities, both of the children and young people and the rest of the evaluation team
- Provide an opportunity for children and young people to get to know each other and the evaluation team
- Provide a space for children and young people to raise any issues or concerns

The Introductory Workshop could also be expanded to facilitate input from children and young people into the overall evaluation design and the terms of reference. One of the deliverables from this introductory workshop could be an evaluation plan that is age appropriate. This will need to be developed in collaboration with the overall evaluation team's inception report. Remember, that you are working with children and young people so encourage the use of different types of methods to develop and present the child friendly evaluation plan.

Some things to think about as you start to prepare to work with children and young people²⁶

- Be aware of Child Safeguarding at all times
- Ensure all work is in line with the Child Participation Practice standards
 - Children and young people need to understand why they are involved and exactly what the scope of their role is and is not
 - Ensure that a friendly, comfortable and inclusive atmosphere is created
 - Make sure that everyone is informed about the purpose of the discussions and that participation is always safe and voluntary
 - Remember to speak slowly and clearly, avoiding the use of jargon and using child friendly straight forward language. Listen to all views
 - Create a culture of respect and trust
 - Where possible support young people to facilitate or run sessions
 - Use participatory, interactive and fun methods
 - Ensure young people needing translation or additional support are able to fully participate
 - Ensure young people involved in the evaluation process have sufficient training to be meaningfully involved
 - Make sure that all activities happen at a time that suit children and young people
 - Use methods that allow all children to actively participate according to their age and abilities
 - Ensure all participants have given their informed consent to their involvement and they can withdraw this consent at any point.

How to ... Decide which data collection methods to use with children and young people

This activity is a way for the whole evaluation team to decide which data collection methods are most useful to use for the evaluation. Start by explaining to the group what each of the data collection methods are (for example, surveys, body map, spider tool, interviews, FGD, stories of most significant change, poetry, drama, drawing or painting, case studies, other reports, newspaper cuttings etc).

Building the matrix

As a group, list the main topics of the evaluation, each on a separate piece of large paper. These pieces of paper make the top of the matrix and are referred to as “headings”. Hand out smaller pieces of different coloured paper with a data collection method on each piece of paper.

²⁶ Adapted from J Howarth, D Hodgkiss, E Kalyva, and S Spyrou (2011) *You Respond: Promoting effective project participation by young people who have experienced violence: A guide to good practice through training and development* and from C O’Kane (2011) *Introductory Booklet on Monitoring and Evaluation with Children and Young people*

For each data collection method ask the group to discuss which of the headings it fits under – for example, is this method best used to gather information about children’s increased involvement in decision making or adults views or the research process? If a method (for example focus group discussions) fits under various headings then make extra copies of this card and place it under each appropriate heading.

Once the matrix is built, use a colour sticker (or star) to highlight tools which are most useful to use when gathering information with or from smaller children and/or with children with disabilities, children from minority groups etc. Encourage a discussion to reflect which tools have been selected and why as well as which tools have the coloured sticker (or star) and why.

This activity could result in many, and potentially all, data collection methods being chosen by the group. If this is the case, initiate a discussion where the group should rank the most appropriate methods to use for the evaluation.

An example from Bosnia-Herzegovina for PRESENTING the matrix exercise once completed to other participants²⁷:

The group hung up the matrix along a washing line (constructed with string hung up between flip charts, for example) and asked volunteers to explain method-by-method why the method was chosen and why it did or did not have a coloured sticker (or star). The group discussed why that particular method was useful for gathering information on that theme and which method had weaknesses. The group draw up a table of figures on a large sheet of flipchart paper which showed which method the group decided to use and which method were most useful and for what.

Data collection with children and young people

There are many fun, participatory activities that children and young people can use to collect data for an evaluation. Children and young people can lead or co-lead these activities. These activities can also be used for adults to gather information from children and young people for an evaluation.

Impact drawings can help to explore the many different outcomes that an activity can produce, and the relations between them. Give participants a piece of paper and ask them to draw a large upward curve and a downward curve. You could explain these curves are a smile and a frown. The curves should be back-to-back and join in the middle. Where the two curves meet the name of the activity being evaluated should be written. The participants then start to list all the positive things that have happened as a result of the activity on the upward curve and all the negative things that have happened as a result of the activity on the downward curve.

²⁷ Adapted from Save the Children Norway (2008) *A Kit of Tools for Participatory Research and Evaluation with Children, Young people and Adults*

Scale of Change thermometer can be used during participatory monitoring and evaluation sessions as a way to visually show perceptions from a group about the wider community who have been affected by the particular activity. The scale of change thermometer may be useful when exploring outcomes, including exploration of whether people other than those directly involved in the project or programme have benefitted in any way. Draw a large thermometer on a piece of flip chart paper with numerical markers within the thermometer. Encourage the children and young people to draw a red line up to the numbers of children they think have been affected by a specific activity²⁸.

Spider diagrams can be used to assess various aspects of a situation and then compare them with other situations over time. Draw a cross on the page and place markers up each line. Explain that the markers closest to the centre are the lower end of the scale and the markers towards the end of the line represent the upper end of the scale. Label each line, for example one line may be “how adults listen to children”. In one colour encourage the group to join up the markers, depending on how they rate the question **before** the intervention. In another colour do the same, but **after** the intervention. Bring the group together to discuss the results and what the results show.

The “H” method can be used to evaluate a particular activity or event by highlighting the positive and negative aspects or effects of a project and asking for suggested improvements. Draw a large H on a piece of paper. Explain that everything to the left hand side of the “H” highlights the positive effects of the project and everything on the right represents the negative effects of the project. In the middle of the H anything that can improve the project should be captured.

There are many other fun participatory activities that children and young people can use to collect data for an evaluation, these include²⁹:

- Body Map of Children’s Experiences
- Tree Analysis
- Risk Mapping
- Time line
- Wall of Wonder
- Drama
- Visioning Exercise
- Circle Analysis of Children’s Role in Peace Building
- Hot Air Balloon
- Preference Ranking

²⁸ Adapted from Claire O’Kane (2011) *Introductory Booklet on Monitoring and Evaluation with Children and Young People*

²⁹ For details of how to carry out any of these additional activities, see Save the Children Norway (2008) *A Kit of Tools for Participatory Research and Evaluation with Children, Young People and Adults*

How to ... Prepare children and young people to collect data for an evaluation

This section takes four examples – focus group discussions, surveys, key informant interviews, and stories of significant change – to show how to prepare children to start collecting data.

Preparing children and young people to run child led Focus Group Discussions

Start this exercise by explaining to the group what a Focus Group Discussion is and why the evaluation team will be using Focus Group Discussions as a tool to gather information for the evaluation. Explain that there are different types of questions that are appropriate at different moments in the Focus Group Discussions. These include:

- **Leading questions** such as “Why” or “What is it about.... What makes you say that”? These types of questions can help go into more detail from a specific topic.
- **Steering questions** get the participants back on track when they have strayed from the topic. For example, “Coming back to that point that you mentioned earlier...”
- **Factual questions** can defuse a group that is becoming tense. For example “How many children are in class 3?”
- **Anonymous answers** can help to generate responses from participants who may be reluctant to share information. For example, all participants can be given a piece of paper to write their answers down.
- **Asking questions** in a different way can help address the questions from the perspective of another person. For example “Why may a young person want to drop out of school?”
- **Silence** can be used very effectively to give participants time to contemplate and respond to more difficult issues.
- **Ending questions** prompt the participants to summarize their positions, provide feedback concerning the moderator's interpretation of the group results and seek any information that may have been missed.

Developing questions

Bring the group together and on a large piece of flip chart paper (even a couple of pieces of flip chart paper together) encourage the group to draw a stream, which becomes a river that leads to the sea. Encourage the group to draw rocks in the water or other landmarks along the water. Explain that the sea represents where the group wants to get to and they should write what they want to find out through the Focus Group Discussions over the sea – this can be called their **ultimate question**. Explain to the group that the group will all be taking an imaginary journey together from the beginning of the stream to the sea to find out what questions they should ask to get to their **ultimate question**. Each time the group get to a rock, they should develop a question that will help them get to their **ultimate question**. Each time a question is developed they should try to identify what type of question it is (see above). Once the group have come up with all their questions share the list of questions with the whole group. These questions will form the basis of the questions to be asked during the Focus Group Discussions.

With the group decide how many key questions will be asked during the Focus Group Discussion (it is advisable to have a maximum of four key questions, but this is dependent on time allocated to the Focus Group Discussion). With the group decide which questions from the activity will help them best answer their **ultimate question**.

Dividing the role of asking the key questions during the Focus Group Discussions

One approach could be to divide the questions up between the whole group so that one young person does not have to ask all the questions during the Focus Group Discussions. If the group decides to do this, explain that the group will need to nominate four lead people to ask one question each during the Focus Group Discussion. Encourage the group into four groups, with a lead for each group to ask the questions (see activity below for a fun way to divide the group into smaller groups).

A fun way to divide the group into smaller groups

Bring the whole group into the centre of the room and explain that they should imagine that they are on a ship. Explain that the ship is sinking so they all have to quickly get off the sinking ship into imaginary life boats. However, the life boats can only hold x number of people. Explain that you will count to three and then the group have to get into their (imaginary) life boats – with the right number of people. Repeat this a couple of times changing the numbers allowed in the life boat each time. Repeat this one last time, and say that the maximum amount a life boat can hold is x (the number of groups you need, in this case four). Once the group is in their four lifeboats, make sure that they nominate one captain for their lifeboat.

Prompt, prompt, prompt: delving deeper into a topic

This is an exercise to help children and young people come up with a list of prompt **questions** to ask during the Focus Group Discussions. Group participants into pairs and explain that one of them will be asking questions – this person can be the interviewer – with the aim of trying to get as much information as possible about a topic from their interviewee. Give each “interviewer” in each pair one question, for example “do you like school” which they then ask their partner – the interviewee. Each time the interviewer gets a reply they have to think of a **prompt question** and ask this **prompt question** to get more information. The interviewer keeps on asking **prompt questions** until they can get no further information. They make a list of all the **prompt questions** that were highlighted through the exercise and share with the wider group. The lead facilitator makes a list of all the **prompt questions** highlighted.

Some examples of prompt questions:

- Could you explain this further?
- Why is this?
- What is it about...that makes you say that?
- Could you share an example of what you mean?
- Could you tell me a little bit more about...?
- Please describe what you mean?

Who will participate in the Focus Group Discussions?

The facilitation team: with the children and young people, decide which roles are important for them to have for the Focus Group Discussion, and who in the group will do which role. Some roles for the group to think about include: One person to be the lead facilitator/co facilitator of the discussion; four young people to ask the key question each; one (or more) young person to ask the prompt questions; two people to take notes during the Focus Group Discussion.

Focus Group Discussion Participants: the types of participants to be involved can be developed in collaboration with the evaluation team. For a participatory way to develop a list of the types of participants refer to section: **Identifying who the key informants are (see below).**

Remember:

- Think about the length of the Focus Group Discussions and make sure that lots of opportunity for breaks and refreshments are factored in
- It is important to make sure that the young people have time to practice asking their questions as well as the prompt questions. This can be done through role play focus group discussion with the whole group
- Make sure that the physical environment to hold the Focus Group Discussion is welcoming and comfortable for all the participants. It should be neutral, private, free from distractions and easily accessible.
- Make sure that the children and young people are clear and confident about their roles during the Focus Group Discussions
- Make sure that the timing of the Focus Group Discussions does not overlap with the other commitments of children and young people – such as school

Preparing children and young people to carry out surveys³⁰

To introduce children to the idea of surveys, first conduct some simple surveys with the children and young people. These can start with very simple yes/no type of questions. To record the information, show the children how to make a tally chart. Each child can tick the box opposite their answer. Here is an example:

Do you like football?							
Yes	X	X	X	X	X	X	X
No	X	X	X	X			

³⁰ Adapted from Plan (2006) *Monitoring and Evaluating with Children*

Now introduce the children and young people to a tally chart that could have more than two possible responses. Here is an example:

What are your favourite colours?							
Red	X	X		X	X		
Blue	X			X	X		X
Green			X	X		X	X
Other		X	X			X	

Encourage the group to start to look at the findings. For example, ask the group which top three colours are the favourite amongst this group? Explore how the group can record their findings. Now the group is ready to conduct their own surveys amongst the group. They should prepare a tally chart to record the collected information. When they have completed their charts, they should analyse them and write two or three simple sentences identifying any trends they have identified. When they are familiar with the idea of a survey the group can begin to develop questions and a format for a real survey. Start by asking the group:

- What do we want to find out?
- Who do we need to ask to find out this information?
- How are we going to collect the information?
- How is the information going to be recorded?
- How are we going to display the information collected?
- What can we learn from the results of the survey?

Remember

- Test the survey questions before the survey is implemented
- Ensure that the group have plenty of opportunity to practise asking their questions
- Because of Child Safeguarding, if children and young people are going to be asking the survey questions, make sure that there is always an accompanying adult with the young person
- Think about different approaches for children and young people who may not have strong literacy skills such as using drawings or symbols
- Before the survey is implemented explore any sensitive issues that may be raised as a result of asking the survey questions as well as how young people should deal with any sensitive issues

Preparing children and young people to run Key Informant Interviews

Start by explaining to the group what a Key Informant Interview is and why the evaluation team could use it.

Identifying who the key informants are:

This part of the activity enables the group to identify and develop the type of people to interview.

- Prior to the session have a series of flip chart pages cut into the shapes of flowers (or with large flowers drawn on the pieces of flip chart paper)
- Label each flower with the name of an organisation or institution that children come into contact with, such as school, youth clubs, health facility etc
- Assign one person to each flower (this person should be someone who can write)
- Explain to the group that they are going to be bees that will visit each of the flowers that are around the room. When they get to the flower they should list all the types of people they think should be interviewed within that group
- Once all members of the group have been to every flower, bring the group back together and share the list the young people have identified
- Encourage the group to rank the participants in order of who will be most relevant to interview

Role play: interview skills

- Divide the group into groups and explain that each group will have 15 minutes to come up with a role play about what makes a bad interview
- Each group then presents their role play to the wider group
- After each group presentation encourage the wider group to shout out all the things that they observed from the role play about what they should not do. As the observations are shared, list them on flip chart paper. You may want to group the observations by area, such as Conducting the Interview, How to ask Questions, Interview Follow Up, How to build rapport etc.
- Once all the groups have presented share the complete list of what makes a bad interview
- On a separate piece of flip chart paper, explain to the group that you will then go through each observation and list the opposite to come up with a list of what makes a good interview

Developing questions for the interview

For a participatory way to develop the questions for the interview refer to the section **Developing Questions** (above).

Remember

- If children and young people will be directly interviewing members of the community, because of child safeguarding they should be accompanied by an adult that they feel comfortable with
- Ensure that the interview is recorded, and all members have given consent to the interview being recorded and the material used to inform the evaluation findings
- Before the interviews start, explore any sensitive issues with the children conducting the interviews that may be raised as a result of asking the questions as well as how young people should deal with any sensitive issues

Preparing children and young people to run Stories of Significant Change activities

The Story of Significant Change can be used individually or in groups. This tool requires special preparation. Ensure counselors are available to support children/young people in case issues arise that touch them profoundly. This tool is especially useful to document changes when it is used at different moments, e.g. with 3 – 6 months intervals. Most other tools can also be used for that purpose.

Start by explaining to the group/person what Stories of Significant Change is and why this is being used as a method to gather information for the evaluation. For more information on Stories of Significant Change see R Davis and J Dart (2005) *The Most Significant Change Technique: A guide to its use*.

First, start by preparing the children and young people to ask communities about what they think the most significant change has been as a result of the project being evaluated. Tell the group that you are going to ask them a question then they can choose how they want to share their answers to the question (Poetry, Song, Writing, Drawing etc).

The first question is: *‘Looking back over the last year, what was the most significant change in your life as a result of being involved with Save the Children?’*

Once the group has had a go at practising answering a significant change question, encourage volunteers to ask their own significant change questions to the whole group. Try to keep the questions around the same theme.

Explain to the group that all these significant stories will be important to inform the findings of the evaluation so they all must be really carefully recorded. Spend a few moments with the group deciding how all the stories of significant change can be recorded.

Some tips on documenting significant change stories

The Significant Change story itself should be documented as it is told. The **description** of the change identified as the most significant should include factual information that makes it clear who was involved, what happened, where and when. Where possible, a story should be written as a simple narrative describing the sequence of events that took place.

What is the most significant change?

If a few practice significant change questions have been asked to the group, there is probably quite a bit of information – flip chart papers/recordings/drawings - all jumbled up at this point. Explain to the group that you are going to start to categorize the information to help understand what the most significant change is. On a piece of paper, write all the different categories of significant change and place them round the room. Group all the recordings from the significant change discussions under the relevant categories. Let the young people decide which category they want to be in.

Explain that they are going to pretend to be magazine editors and that their category is the theme of the magazine. Explain that they have lots of articles (these are all the recordings of the significant change practice questions) and they need to make sure that the articles that are of the most interest and show the most significant change are placed on the front page of their magazine. The least significant change is placed on the back page of the magazine. The magazine editors then take it in turns to present their magazine to the rest of the group, highlighting how they chose their cover story and the sequence of all the other stories.

Once the group have practised the process for asking, documenting, recoding, ranking significant change stories, relate the activity to the project being evaluated.

Coming up with the significant change questions to ask the community

Bring the group together to develop the list of significant change questions about the programme being evaluated to ask the community. For an activity to develop the questions refer to section on **Developing Questions** (above).

Developing categories

Explain that when the evaluation team get the stories from the community, the stories need to be focused as the evaluation is not about capturing lots of changes that are not related to the project. Refer back to the intended project outcomes to keep the group focused. Highlight some possible significant change categories, for example

- Changes in the quality of people's lives
- Changes in the nature of people's participation in development activities
- Changes in the sustainability of people's organisations and activities

Choosing the most significant change story

The group may want to come up with their own process to select the most significant change story. Some common rules could be to ensure:

- Everybody reads the stories
- The group holds an in-depth conversation about which stories should be chosen
- The group decides which stories are felt to be most significant
- The reasons for the group's choice(s) are documented
- Stories that are filtered out should not be thrown away
- The results of a selection process must be fed back to those who provided the Significant Change stories
- Consent is always obtained to use the information within the stories

An activity to help children and young people practise asking tricky questions ³¹

The aim of this activity is to encourage the group to think about the kinds of questions that draw out information and to practise asking questions that will generate a more substantial answer.

How is this activity run?

Explain to the group that you are doing this exercise because sometimes when you ask questions it can be tricky as sometimes people do not tell you very much.

Organise the group into pairs and explain that you are going to play the Yes/No game. Explain that one person in each pair will ask a series of questions to their partner and that their partner is not allowed to say Yes/No to the question. For example, you could ask questions such as “Do you like your school?” to which then you must reply with a full answer and not Yes/No. If you hear a Yes/No you must clap loudly. Encourage the pairs to swap roles. Bring the group back together and ask them to share what kinds of questions get answers other than “yes” or “no”. These will be questions like “Can you tell me more about... “or “How does that happen” or “What do you like best about... and why?”

Ask the group to collectively develop a list of key phrases that they should use to frame the questions being asked as part of the evaluation. Encourage the group to make a note of the phrases.

How to ... Document the results with children and young people

Explain to the children and young people that whenever information is being generated for the evaluation, it is really important to ensure that the results are **documented**. Collaboratively come up with a list of all the ways information can be documented. These include:

- Drawing, photos, poetry, stories, collages
- Written notes, flip charts, photos and or film recording of children’s drama
- Tape/video recordings

Throughout the evaluation the children and young people may want to keep a diary to capture

- What happened (evaluation activities, incidents, events, significant conversations)
- Impressions and reflections about the possible significance of findings
- Thoughts, ideas about the evaluation process what works well, challenges faced, and ideas to strengthen the evaluation tools.
- Issues and questions

³¹ Adapted from Participation Works Taking Part in Making Decisions: Training for 8 to 12 year olds. Involving children and young people in Evaluation .See <http://www.participationworks.org.uk/resources/taking-part-in-making-decisions-training-for-8-to-12-year-olds-all-six-parts>

Remember

- To ensure all sessions are recorded by either audio/visual or written
- That all documents should be clearly identified with the date, session title, number of participants, gender and age and group name
- To ensure all documentation is kept safely and securely by locking documentation in a secure unit

How to... analyse and validate data with children and young people

Analysing the results can be the most exciting stage, but also the most challenging. You may find that it is difficult to reach consensus about the findings, with a number of contradictory findings, as well as unexpected results. It will be important to bring all the differing and contradictory information that has been collected through the participatory activities to an analysis workshop. If there is a lot of information you may want to split the group into smaller groups to look at the different data sets.

Analysis game ³²

Divide the participants into groups. Give each group a set of 12 objects (could include stones, leaves, plastic bottles, post it notes, scissors, note pad, tape etc).

- Ask the participants to divide the objects into two categories and to explain each category
- Now ask the participants to divide the objects into three categories and to explain each category
- For the last time ask the participants to divide the objects into four categories and to explain each category
- Ask the participants what they have learnt from this game?

Explain to the participants that categorising data into key themes - in order to present, explain and analyse – is an important part of the evaluation process.

There are many ways to interpret data and the children should be encouraged to give their views on each aspect. Everything should be noted down and kept for the report and for future reference.

When analysing the data some steps can be applied³³:

- Identify the issue
- Break it down into smaller elements
- Identify gaps and seek additional information if necessary
- Make a provisional diagnosis
- Check it out with interested people
- Form a conclusion

³² Adapted from Save the Children Norway (2008) *A kit of tools for participatory research and evaluation with children, young people and adults*

³³ Adapted from C O'Kane (2011) *Introductory Booklet on Monitoring and Evaluation with Children and Young People*

Validating data with children and young people

We need to systematically analyse our findings to see what they are telling us. We need to consider the validity and usefulness of the data that we have collected. We need to compare findings gathered from different people. We also need to compare findings gathered using different methods. This is called triangulation when we cross-check the information that we have collected through different methods. It helps us see which findings match up, and helps us determine which evidence is strong and which evidence is weak. It also can help us identify gaps, where we need to check additional sources of information.

How to ... Draft, review and disseminate the evaluation report

Once all the data has been collected, analysed and validated all the information can be brought together in a report. In line with the Save the Children International MOS, this section looks at how to develop a supplementary **child friendly version of the evaluation report**.

To gather the information together for the **child friendly evaluation report** a good game to play with the children and young people is **Reporters**.

Prior to this game, make sure that all the evaluation results are grouped by area or by how you intend the report to be structured.

Divide the group of young people into small groups and give them a section of the results each. You may want to have an adult within each group to encourage the group and provide clarification. Explain to the groups that they will be playing a game where they are reporters and they need to develop an article about a big news story. All the material for their article will be the evaluation results. Explain to the group that they can use whichever method they want to present their article. Some ideas include

- Role play/drama such as a TV report or interview
- Collage
- Magazine/newspaper page or article

Give the group lots of time to look at the results and decide how to develop their article. Once the groups have come up with their articles they can present them back to the wider group – either in a presentation or a gallery walk. One member of the group, or an independent member takes careful notes of the presentations of the reports. This information can then be pulled together to form the **Child Friendly version of the Evaluation Report**. Let the children and young people decide what form of media they want to use to present their **child friendly version of the evaluation report**. Do not assume that this will be a written report as it may also be a video, a collage, a play etc.

Disseminating the report

Sharing the results and developing recommendations, South Asia³⁴

Children who had been involved in a regional study on Children's Citizenship in South Asia supported by Save the Children UK, were included in a final presentation of the results. The presentation was attended by a wide variety of people from Save the Children, the UK agencies and NGOs. Children's representatives presented their own findings and feelings. Once the presentation had taken place separate groups of children and adults were formed to discuss the recommendations. Their comments and ideas were then fed back into the report

How to ... Feed back and following up with children and young people

It is important to get feedback from children and young people about their views of being involved in the evaluation process. One way to do this is through a **Gallery Walk**.

Before the session stick up/hang/pin all documentation from the processes that children and young people have been involved in. This can be done as a timeline with the first activities at the beginning of the room, and the final evaluation reports at the end of the room. Next to each activity have a little box called a suggestion box.

Give each young person some small pieces of red, green and yellow paper. Explain that they can walk around the Gallery and in each suggestion box post their thoughts

- On the green pieces of paper things that went well with the activity
- On the red pieces of paper things that did not go well with the activity
- On the orange pieces of paper overall things that can be improved with the activity

This exercise can be done by the whole evaluation team.

What next...?

In line with the Save the Children Child Participation practice standards, once the evaluation has been signed off it is important to provide feedback to the children and young people who were involved in the process. Feed back to the children and young people about what happened as a result of their involvement as well as what has happened to the evaluation report and how it will be used.

³⁴ Save the Children (1999) *Children and Citizenship - Regional Study in South Asia*

ANNEX 8: Gender equality audit

What is a gender audit?

A gender equality audit looks at the organisation, systems, procedures and staff, to get a picture of what is being done on gender equality, as well as challenges and needs to address them. It then develops actions on how to address them. It can be undertaken by the whole organisation or a selected part. The audit should seek to answer questions around levels of commitment, understanding and capacity and resources available and need to work on gender equality.

What is the purpose of a gender equality audit?

The ultimate purpose of a gender audit is to assess the organisation against a set of criteria related to gender equality and develop recommendations to address findings. However, the actual process of undertaking the audit is equally important. It provides a non-judgemental space for exploring the main issues, raising awareness, ensures staff engagement and ownership. For example, focus groups discussion is an opportunity to explore the understanding of concepts around gender equality to ensure that staff have a shared understanding.

The gender equality audit should help to answer the following questions:

- 1 What level of commitment is there in the team/office/among partners to promoting gender equality
- 2 How well do staff and partners understand terms such as gender, gender equality, women's rights, mainstreaming? Who is confused by the terms, who is opposed to them, who questions their meaning?
- 3 What skills, resources and capacity do staff have to do this kind of work, and what support do they need to do it better?
- 4 Is gender integral to all the work including the way the organisation works or is it missing from many documents?
- 5 How does the focus on gender equality fit with other organisational priorities?
- 6 Who has the responsibility for ensuring gender is integrated into the work at all levels and across all functions? How is that monitored?

A gender audit involves the following steps:

- 1 Developing the methodology, timeline and methods
- 2 Document review and scoring against gender criteria (see below)
- 3 Face-to-face work: individual interviews, focus group discussions,
- 4 Phone interviews with staff, partners and/or key external informants far away
- 5 Accurately recording the data
- 6 Data analysis to analyse the findings and produce recommendations
- 7 Face-to-face feedback on the findings in a workshop with staff
- 8 Report writing

- 9 An action planning workshop based on the findings to develop actions around:
- Training and capacity building needs of staff
 - Changes needed to guidelines, manuals, templates
 - Policy changes required to promote equal opportunities internally
 - Tackling areas of resistance and misunderstanding
 - Providing support to those motivated to take this work forward

What makes an audit work well?

In order to get the most from an audit it needs to:

- be supported by senior managers
- be highly participatory
- involve both staff and partners
- allow people to ask as many questions as they need, allow people to question and disagree,
- explore people's fears and worries around the issues as well as their commitment and motivation
- be given adequate resources, especially TIME
- set very clear ground rules of openness, confidentiality, respect for all opinions to ensure this is a safe space
- for staff and partners to explore issues and ideas³⁵

Criteria for assessing gender equality in document review

Criteria can be adapted to the needs of the programme or the organisation. The OECD/DAC gender equality policy marker is the closest there is to a sector-wide standard for gender mainstreaming. However, developed to assess donor efficiency in spending on gender equality it needs to be adapted for broader purposes. Nevertheless these criteria have been adapted and used for gender audits on different levels.

OECD/DAC gender equality policy marker: An activity should be classified as gender equality focused if it is intended to advance gender equality and women's empowerment or reduce discrimination and inequalities based on sex. Spending can be categorised as principally concerned with promoting gender equality, or not concerned with promoting gender equality. The gender equality policy marker was developed to enable donors to measure bilateral aid allocated by sector to activities which are intended to "advance gender equality and women's empowerment or reduce discrimination and inequalities based on sex". The DAC policy marker for gender equality is based on intentions at the design stage and provides a means for individual DAC Members to monitor their own progress in pursuing gender equality objectives. Most gender audit criteria are based on these principles³⁶.

³⁵ Adapted from TROCAIRE (2010) *Gender Audit Manual*

³⁶ <http://www.oecd.org/dataoecd/0/63/37461060.pdf>

DAC Gender Equality Marker

An activity should be classified as **gender equality focused** if it is intended to advance gender equality and women's empowerment or reduce discrimination and inequalities based on sex. Spending can be categorised as **principally** concerned with promoting gender equality, **significantly** concerned with promoting gender equality, or **not concerned** with promoting gender equality.

*Criteria for eligibility*³⁷

Gender equality is explicitly promoted in activity documentation through specific measures which:

- 1 reduce social, economic or political power inequalities between women and men, girls and boys, ensure that women benefit equally with men from the activity, or compensate for past discrimination; or
- 2 develop or strengthen gender equality or anti-discrimination policies, legislation or institutions.

This approach requires analysing gender inequalities either separately or as an integral part of agencies' standard procedures.

Examples of typical activities

Examples of activities that could be marked as **Principal** objective:

- legal literacy for women and girls;
- male networks against gender violence;
- a social safety net project which focuses specifically on assisting women and girls as a particularly disadvantaged group in a society;
- capacity building of Ministries of Finance and Planning to incorporate gender equality objectives in national poverty reduction or comparable strategies.

Such activities can target women specifically, men specifically or both women and men. Support to women's equality organisations and institutions scores, by definition, a *Principal* objective.

Examples of activities that could be marked as **Significant** objective:

- activity which has as its principal objective to provide drinking water to a district or community while at the same time ensuring that women and girls have safe and easy access to the facilities;
- a social safety net project which focuses on the community as a whole and ensures that women and girls benefit equally with men and boys.

³⁷ Source: OECD DAC GenderNet website: <http://www.oecd.org/dataoecd/4/23/39903666.pdf>

The Inter-Agency Standing Committee (IASC) for humanitarian assistance gender marker is an adapted version of the DAC gender equality marker launched in 2009/2010 in response to the recognition that evaluations of humanitarian effectiveness show that gender equality results were weak. The purpose of the marker is to improve humanitarian programming and make humanitarian response more efficient. They have developed a set of recourses to ensure gender is included in the design stage of a programme³⁸.

Gender Audit Manual, TROCAIRE

Score	Criteria
0	Gender is not mentioned as an issue
1	Gender is acknowledged as an issue: eg: there is recognition of gender inequality or discrimination as a development, poverty or human rights issue but this is not built into the rest of the document.
2	Gender is analysed as part of the context and problem: there is use of sex disaggregated data and gender analysis, for example, analysis reflects gender-based violence, gender and poverty, women's rights, etc, but not built into the response.
3	Gender is analysed in the context and problem and built into the response (for example, the target group reflects gender analysis; gender is incorporated into objectives and indicators, advocacy asks, policy recommendations etc.

³⁸ <http://oneresponse.info/crosscutting/gender/Pages/The%20IASC%20Gender%20Marker.aspx>

ANNEX 9: The contract manager's responsibilities³⁹

Good contract management helps us to manage the risks of a relationship with external parties, including consultants, and to ensure that the rights and interests of Save the Children are protected.

In general, the contract manager will act as the first point of contact for all correspondence and communication, will negotiate contract variations, and will maintain records of all interactions with the contractor. A successful relationship between Save the Children and the contractor is one where both parties recognise their mutual dependence and interest in a cooperative relationship with a shared understanding of the task, and have realistic expectations for the achievement of that task. Clear, open and effective communication is vital to achieve this.

Contract protection mechanisms such as performance fees, withholding instalments, and termination clauses may be appropriate. Any action will need to be consistent with the contract, should the final report or output not meet the requirements outlined in the terms of reference.

Contracts must comply with local laws and donor requirements. Consultants need to be provided with key policies and documents such as Save the Children's child safeguarding policy, ethical standards, and Standard Operating Procedures. The contract manager is responsible for ensuring that the consultant signs these and complies with them.

A note on intellectual property

Evaluation contracts entered into by Save the Children need to protect the intellectual property (IP) rights of Save the Children, especially with respect to publishing. Research/evaluation contracts often have a number of clauses that relate to research, specifically on IP rights and publication. IP rights should be clarified in the contract so that all materials associated with the evaluation (including data) belong to Save the Children. Where it is not possible to negotiate ownership of contract materials, you should try to negotiate for Save the Children to have the maximum access to and use of any contract materials and related data.

We recommend that you request Save the Children lawyers review contracts prior to signing. To speed up this process, follow these guidelines when you draft a contract and prior to the legal review.

All Save the Children contracts must:

- protect confidential and personal information
- align with and not contradict the IP clauses of Save the Children International (SCI)
- protect materials that are created by, and provided for the purposes of, the contract
- clearly outline procedures for publishing and authorship (see below)
- contain clear definitions of the above terms.

³⁹ Adapted from the Australia Government Solicitor, Managing Contracts in the Public Sector, Commercial Notes, Number 22, 9 May 2007.

Authorship

Authorship is another issue that should be carefully considered by the programme or project manager. Save the Children authorship guidance is in line with the Vancouver Protocol, an international standard applied across all disciplines in the world's top universities. Authorship of Save the Children evaluation/research is a matter that should be discussed between researchers at an early stage in a project and reviewed whenever there are changes in participation. Generally, the report will be authored by the evaluation team and Save the Children will own the intellectual property rights.

All persons designated as authors should qualify for authorship.⁴⁰ The order of authorship should be a joint decision of the co-authors. Because the order is assigned in different ways, its meaning cannot be inferred accurately unless it is stated explicitly by the authors. Authors may wish to explain the order of authorship in a footnote.

Group members who do not meet these criteria could be listed, with their permission, in the Acknowledgements or in an appendix. Courtesy demands that you should also acknowledge individuals or organisations that have provided facilities or other support.

Authorship criteria

- Conception and design, or analysis and interpretation of data
- Drafting the article or revising it critically for important intellectual content
- Final approval of the version to be published

Conditions 1, 2, and 3 must all be met. Participation solely in the acquisition of funding or the collection of data does not justify authorship. General supervision of the research group is not sufficient for authorship. Any part of an article critical to its main conclusions must be the responsibility of at least one author.

Checklist – Contract review

- ✓ Why are we entering into this contract?
- ✓ Who are the other parties to the contract? IS it clear what we are paying for? Save the Children should not be paying significant sums up front until we know that the other party will deliver, and ideally should only be paying for what we receive.
- ✓ What happens when they don't perform? It is important that Save the Children has express rights to terminate the contract (without paying any penalties) if another party does not do what they agreed to do, or becomes insolvent or otherwise unable to complete the contract. If the contract involves intellectual property, the termination provisions should deal with this.
- ✓ What are Save the Children's commitments outlined in the contract?
- ✓ Are there any other risks for Save the Children? Keep an eye out for the indemnity or other provisions that make Save the Children liable for losses to other parties, or for taxes, etc.

⁴⁰ Adapted from the Vancouver Protocol (1997), available at: http://www.research.mq.edu.au/about/research_@_macquarie/policies,_procedures_and_conduct/documents/Vancouver.pdf

- ✓ What question to ask? If you don't fully understand the contract, ask questions or seek advice. Contract provisions should include:
 - timetables and deadlines for key activities
 - contractor commitments
 - Save the Children commitments
 - Potential risk areas
 - Variations (formal/informal)
 - Optional sub-contracting arrangements. If another party tells you that a provision in the contract has a particular meaning, ask them to say that in the contract, or at least to confirm it in writing. Don't be influence by another party telling you that the contract is "standard", that others have just signed it, or that there's no time to negotiate changes to the wording