

Guide to the Evaluation of Psychosocial Programming In Emergencies

UNICEF
2009

GUIDE TO THE EVALUATION OF PSYCHOSOCIAL PROGRAMMING IN EMERGENCIES

ACKNOWLEDGEMENTS

UNICEF would like to thank the consultants from the Mailman School of Public Health at Columbia University who produced this guide, Neil Boothby, Alastair Ager and Wendy Ager. In addition, the input of many individuals and organizations who commented on early versions of this draft and have taken great interest in the process has been invaluable. Special thanks goes to those who participated in an expert workshop in early 2007 to review the first draft and to the participants in an early pilot of the guide in the occupied Palestinian Territory in March 2007. We would also like to thank ECHO who provided the funding for this initiative.

PREFACE

UNICEF has developed this **Guide to the Evaluation of Psychosocial Programming in Emergencies** to assist UNICEF, our partners and other organizations working in the field of psychosocial support in thinking through key issues in planning and implementing an evaluation.

There are challenges of conducting evaluations in areas of armed conflict and other humanitarian crises. However evaluation is a vital tool for improving current psychosocial programs as well as future planning, programming and decision-making. Evaluation provides the means to improve program performance and build inter-agency consensus on good and promising practices. Essentially the wider impact of well-documented, reliable evaluations will be the building of a stronger knowledge base for good psychosocial practice.

In recent years, psychosocial support has become an increasingly common part of development and humanitarian programming. Psychosocial and mental health programming has included a range of very different approaches and activities which have often led to diverse and at times, inconsistent approaches. Differing approaches and understanding have also tended to polarize debates on effective and appropriate programming.

Recently there has been significant progress in bridging some of these gaps and building consensus on effective, ethical programming. In particular, the Inter-Agency Standing Committee Taskforce (IASC) on Mental Health and Psychosocial Support has finalized guidelines that outline appropriate minimum responses in emergencies and which bridge the traditional divide between mental health and psychosocial programming. These guidelines represent the consensus of the international humanitarian community on appropriate psychosocial support and mental health programming in emergencies.

Such consensus represents a major achievement in developing more effective and ethical psychosocial and mental health programming. However, there is a widespread recognition among both practitioners and researchers that there remains a need to build a stronger evidence for this work. More extensive and better quality evaluations and research are required to develop a better understanding of what approaches to psychosocial support are the most effective, in what situations.

A desk review of existing psychosocial assessments and evaluations by the Mailman School of Public Health, with the support of UNICEF, found that some effective psychosocial evaluations have been conducted and a wide range of tools exist. However, it also identified a number of widespread problems that led to questionable or inconclusive results including:

- 1) Lack of clear and appropriate project objectives;
- 2) A number of common methodological weaknesses in evaluations; and
- 3) Lack of appropriate quantitative tools for assessing psychosocial wellbeing.

These problems appeared to stem from both the complex nature of the subject as well as a lack of capacity and accessible guidance in this area. A lack of agreement about the objectives of psychosocial programmes and what exactly constitutes psychosocial wellbeing have been contributing problems, as have methodological difficulties such as the challenges of measuring wellbeing and of evaluating the impact of psychosocial programme. Capacity to address these issues remains a problem - while some useful tools and approaches for psychosocial evaluation exist, field colleagues are often overwhelmed by the difficulties of understanding this complex programming area and have not had sufficient access to appropriate technical guidance and support.

Recent reviews by UNICEF and partners have suggested that these problems are also indicative of a broader problem of knowledge generation and management. For instance, the recent Machel+10 review of children affected by armed conflict found that there is a need for greater strategic investment and capacity building in data and knowledge on children affected by armed conflict. This is particularly the case for new or recent programming areas, such as psychosocial support. One crucial recommendation has been to have a more strategic approach to evaluation – for instance, conducting fewer but better quality evaluations that are coordinated across countries.

In the case of psychosocial programming, the resulting lack of evidence has contributed to divergent and at times contradictory approaches to psychosocial programming that are common in the field. In order to build the evidence base for psychosocial programming, one important step is to develop understandable, usable guidance and tools for the field on how to conduct psychosocial evaluations, and to build capacity on using these tools through training and

technical support. Strengthening interagency consensus on these key issues is important as building the evidence base for psychosocial programming is by necessity an interagency effort.

The aim of this manual is to provide basic guidance on psychosocial evaluation that can be used by UNICEF, our partners and other organizations conducting psychosocial programmes. The guide aims to provide an overview for those designing and managing the implementation of psychosocial evaluations at the field, regional and headquarters level. It does not attempt to give detailed guidance on specific skills or techniques required to conduct psychosocial evaluations, but refers the reader to appropriate guidelines/tools for this information.

The current draft of this manual has been developed through wide consultation with organizations and experts working on psychosocial programming and evaluation. This draft is available for piloting testing by UNICEF and other interested organizations through 2008. All organizations are encouraged to use, adapt and provide feedback on this draft. Based on these experiences and feedback, UNICEF will subsequently revise the manual.

Amanda Melville

December 2007

CONTENTS

- 1 The aim of the guide**
- 2 Planning for evaluations**
- 3 Overview of psychosocial programming**
- 4 Relationship to general UNICEF guidance on evaluation**
- 5 Objectives of psychosocial programming**
- 6 Indicators**
- 7 Evaluation design**
- 8 Preparing for an evaluation**
- 9 Methodology**
- 10 Implementing an evaluation: case examples**
- 11 Writing up and sharing findings**
- 12 Annex A: A Guide to Developing Indicators**
 - Annex B: A Step-By-Step Guide to Conducting an Evaluation**
 - Annex C: A Program Officer's Responsibilities in Evaluation of Psychosocial Programs**
 - Annex D: Guidance on Sample Selection**
 - Annex E: Useful Resources**
 - Annex F: Summary of Major Recommendations**

1. The Aim of this Guide

This guide provides practical guidance for the evaluation of psychosocial programs in emergencies. Although not the main objective of this guide, we recognize that the tools and methods described may also be relevant to non-emergency contexts, such as those impacted by HIV-AIDs and circumstances of severe violence, abuse and exploitation. The guide is written primarily for UNICEF staff with responsibility for psychosocial programming and is a supplement to existing UNICEF guidance on programming and evaluation. It is also intended to be of value to other organizations implementing psychosocial programs.

'The guide is written primarily for UNICEF staff with responsibility for psychosocial programming... [but] is also intended to be of value to other organizations'

The two main goals of UNICEF's psychosocial response are:

- To promote psychosocial wellbeing by promoting an environment that provides appropriate care, opportunities for development and protects children from exposure to situations are harmful for their psychosocial wellbeing, and
- To respond to psychosocial problems by strengthening social and psychological supports for children who have been exposed to situations harmful for their psychosocial development.

This guide focuses on the evaluation of psychosocial programs that are aligned with these two goals. It does not include consideration of clinical services and other specialized mental health programs.¹

The guide aims to provide concise, clear guidance in an accessible format, using real world examples throughout to illustrate how psychosocial programs can be evaluated in the field.

2. Planning for Evaluations

Evaluations set out to understand whether or not programs have achieved their goals and what has been learned in the process. More specifically, evaluations usually fulfill one or more of the following purposes:

1. To provide accountability to stakeholders (including beneficiary communities as well as funders etc.) regarding the achievements of programming.
2. To provide information to develop and improve programming in subsequent phases of implementation, and to identify any unintended negative consequences of programming.
3. To help develop a more effective evidence base for psychosocial programming in other situations and settings.

From this it should be clear that evaluation is a central feature of project design and 'project cycle management'². Ideally an evaluation strategy should be considered right from the start of the project cycle. It is much more difficult to do an evaluation if it is tacked on towards the end of a program. When evaluation is considered from the outset, it can help clarify objectives and promote the engagement of local communities in the design and planning of the program. Quite often, however, evaluation is not planned in advance and time and resources are limited. Even in circumstances of budgetary, time, data and political constraint, evaluation can provide the opportunity to capture valuable lessons³.

UNICEF and implementing partners inevitably need to make decisions regarding how much resource to commit to evaluation processes. Some evaluations are clearly more intensive than others. Decisions about evaluation are part of a wider Monitoring and Evaluation strategy. There are three components to be considered (A, B and C below):

A: **Monitor** the inputs and outputs of the project

This is a reporting requirement for all projects and is usually effectively addressed by project teams. As a result, this guide does not provide much guidance regarding project monitoring methods.

Project inputs are the money, materials, equipment, staff and other resources 'put in' to project activities

Project outputs are the planned achievements 'put out' in the process of implementing a project (such as newly trained staff or improved services or facilities) that signal that work is on track

B: Evaluate the outcomes of the project

This component is also essential but, as noted in the Preface, is often ineffectively addressed. This guide seeks to provide concrete assistance in planning for effective evaluation of project outcomes. In emergencies such evaluations may help in the adjustment of programming in light of rapidly changing circumstances (sometimes referred to as 'real time evaluations').

Project outcomes are the results that 'come about' during the course of a project as a result of the outputs achieved

Project impact is the overall change in the lives of children (and their families and communities) that results from a project

C: Evaluate the impacts of the project

Although all projects need to assess if they have achieved some change – that they can report some outcomes of their work – for the psychosocial field at this time, we need rather more. We need to show that these outcomes result in real changes in the lives of children, their caregivers or their communities. The rationale for projects often begins by identifying major needs regarding children's well-being and circumstances that need to be addressed. Psychosocial programming must be able to demonstrate that project outcomes do lead to real changes in children's well-being and circumstances – that is, that they have the impact intended.

To build the evidence-base of outcomes of psychosocial projects leading to such change, some longer-term impact evaluations need to be conducted. These may be at country, regional or global level and may be particularly important in circumstances of large scale emergencies or in considering key interventions in emergency settings (e.g. Child Friendly Spaces, counseling).

Getting help

For those evaluations monitoring progress and usually focusing on outputs (see A above), it is most likely that this will be an internal exercise. It should be a process built in to the running of the program, and the responsibility of program staff.

Outcome evaluations (see B above) should be something that program staff feel able to engage with, perhaps with the support of technical assistance from within the organization or external consultants. Both can be good routes to bringing in relevant expertise. However, in some circumstances, external consultants can promote a degree of independence in the evaluation that lends credibility to findings.

Impact evaluations (see C above) will usually need the expertise of external consultants.⁴ However, programme managers will need to be able manage and

oversee the design and implementation of these evaluations, and therefore need a strong understanding of how they should be conducted.

Within UNICEF there are a number of ways of bringing in external expertise:

- Bringing in UNICEF staff with a Monitoring & Evaluation (M & E) remit at country or regional level country officers
- Drawing on relevant headquarters staff in the Child Protection and Evaluation sections for technical support in psychosocial evaluation
- Engaging local NGO specialists (for example, with expertise in psychosocial evaluation)
- Contracting external expatriate and/or local consultants

The costs involved in conducting evaluations, including where appropriate contracting external consultants or NGOs, is an expense that should be anticipated at the stage of program planning. The resources that are available to support an evaluation will determine its scale and effectiveness. Given the importance of establishing the effectiveness of psychosocial interventions, committing resources to evaluation is an essential part of good programming.

3. Overview of psychosocial programming

Introduction

Exposure to the disruption, loss, and violence associated with emergencies places significant psychological and social strain on children, adolescents, their families and communities. The way in which children and families experience and respond to conflicts and disasters varies greatly, yet with the right support the majority will be able to overcome these difficult experiences. It is essential that social and psychological issues are not ignored while homes are rebuilt, social services re-established and livelihoods recommenced. It is now widely accepted that early psychosocial interventions must be an integral part of humanitarian assistance⁵.

A shift in emphasis from children's vulnerabilities to a view of children as active agents in the face of adversity has been reflected in a shift from trauma based models of service delivery to those which recognize and strengthen resilience and local capacities. A resiliency-building approach to psychosocial wellbeing and child protection therefore focuses on the following kinds of objectives:

- Reducing risks to children's safety and emotional well-being while promoting an environment conducive to positive development, effective coping, and resilience
- Promoting children's holistic development and age-appropriate physical, cognitive, and emotional competencies

- Fostering a secure and stable environment for children
- Strengthening family and community care-giving structures for children
- Supporting children's and youth's voice and full participation in all phases of child protection programming
- Strengthening local networks that enable child protection, care, and well-being

This approach is reflected in the IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings⁶, summarizing current consensus on best practice in psychosocial support and mental health programming in emergencies.

Domains for psychosocial evaluation

The term 'psychosocial' emphasizes the close connection between psychological aspects of our experience (that is, our thoughts, emotions and behavior) and our wider social experience (that is, our relationships, traditions and culture)⁷.

There are many examples of psychosocial practitioners using different language and concepts to describe their work. Although this is confusing, across different approaches two principles seem to consistently emerge. Firstly, psychosocial programs are concerned with psychological and social aspects of children's lives. Secondly, programs don't just focus on children as individuals, but include their families and/or caregivers and also take account of the place of children in the wider community.

Acknowledging subtly different emphases across agencies, this guide suggests that the following **three domains** as the most helpful to evaluate how well UNICEF's work affects the lives and experiences of children:

- **Skills and knowledge**
e.g. life skills, using culturally appropriate coping mechanisms, vocational skills, conflict management etc.
- **Emotional well-being**
e.g. feeling safe, trust in others, self-worth, hopeful for the future etc.
- **Social well-being**
e.g. attachment with caregivers, relationships with peers, sense of belonging to a community, access to socially appropriate roles, etc. resuming cultural activities and traditions

Psychosocial programming is generally related to one or more of these domains – skills and knowledge, emotional wellbeing and social wellbeing. They may be reflected in different ways in different cultures but they represent the common core of most psychosocial work.

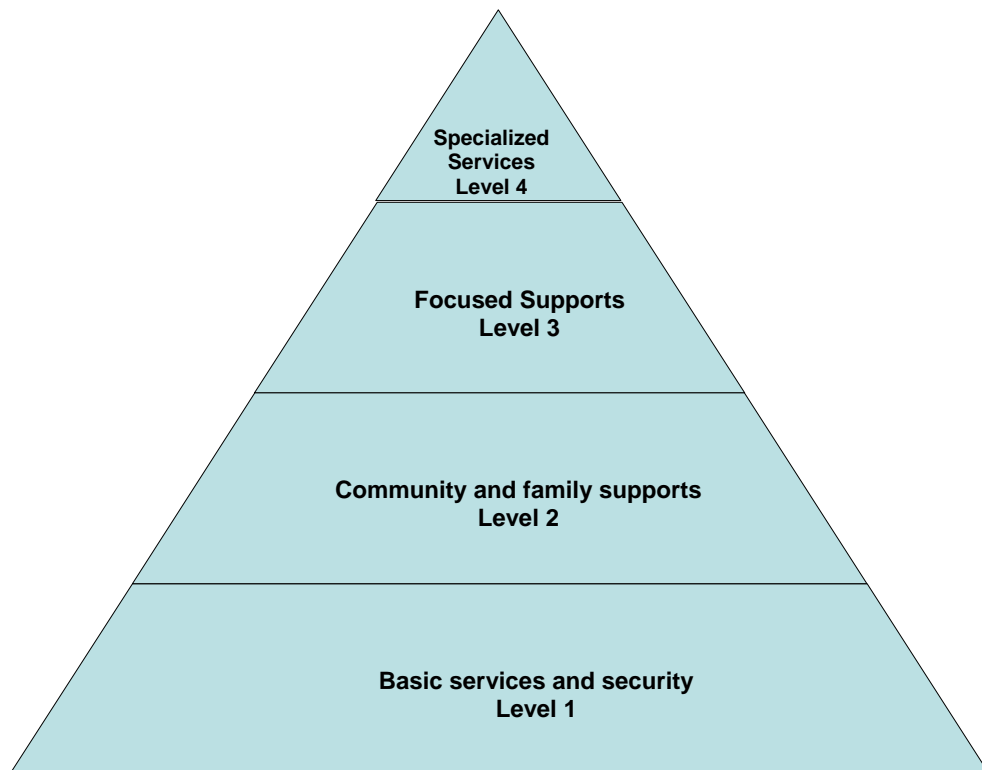
These domains draw on the framework developed by the Psychosocial Working Group⁸ that sees psychosocial wellbeing as reflecting three inter-related issues:

- Human capacity - the physical and mental health of people, as well as their knowledge and skills
- Social ecology – the social connections and support that people share
- Culture and values – the specific context and culture of communities that influence how people experience, understand and respond to circumstances

In sections 5 and 6, these core domains are used to define objectives for evaluation.

Core psychosocial activities

Agencies get involved in many different types of activities in their quest to make a difference to the wellbeing of children. A helpful representation of the range of work undertaken in this field is provided in the IASC guidelines:⁹



Core psychosocial work within UNICEF is typically at levels 1, 2 and 3, that is advocacy for basic security and services, strengthening community and family

supports and, providing focused supports. Children with psychological disorders are referred to specialized mental health resources, where they exist. The IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings (2007) provide more information on these multi-layered supports (see pages 11 to 13).

Briefly, services offered at these four levels are as follows¹⁰

1. Basic services and security. The foundation for well-being is via the meeting of a person's basic needs and rights for security, adequate governance, and essential services such as food, clean water, health care and shelter. Advocacy with other sectors can focus on ensuring that these services and protections are put in place, and that this is done in a way that prevents psychosocial problems, and supports well-being e.g. by ensuring families are not separated or discriminated against in the way aid is distributed.

2. Community and family supports. Community mobilization is an essential primary activity to strengthen social support networks, and help people resume functioning. This may include funding educational and vocational projects, supporting community based children's activities, or promoting social support networks.

3. Focused supports. A smaller number of people will in addition require supports that are more directly focused on psychosocial well-being. These are for children or adults who having difficulty coping with their existing support network, but who are not suffering from a clinical mental disorder. This may include activities to help deal with the effects of particularly distressing events e.g. support groups for victims of rape or torture. These are typically carried out by trained social or community workers, or health care professionals

4. Specialized services. At the top of the pyramid is additional support for a small percentage of the population whose suffering, despite the aforementioned supports, is intolerable and/or who have great difficulties in basic daily functioning – that is, those who have severe clinical mental health disorders such as psychosis, drug abuse, severe depression, anxiety, or harmful to themselves or others etc. This assistance could include psychological or psychiatric supports for people with mental disorders when their problems cannot be adequately managed within primary health services.

Principles of psychosocial support

In the course of evaluating psychosocial programs, it is crucial that key psychosocial principles are observed at every stage of the process. The principles defined here are, firstly, psychosocial principles, grounded in the Convention on the Rights of the Child¹¹ and, secondly, ethical principles for conducting psychosocial evaluations¹².

General psychosocial principles¹³:

- ***Best interests of the child***

The best interests of the child should be the primary consideration for all activities, taking into account what will be the impact for children, and avoiding doing harm. For example, groups for separated children may be designed to support them, but may also cause discrimination if these children are seen as different.

- ***Child, family and community participation and empowerment.***

The most effective and sustainable approach for promoting psychosocial well-being and recovery is to strengthen the ability of families and communities to support one another. Relief efforts can make matters worse if they reinforce a sense of powerlessness by treating those affected as helpless victims. Girls, boys, women and men should be active partners in decisions that affect their lives e.g. via involvement in relief efforts, older children working with younger children, parent committees.

- ***Build capacities and strengthen resilience***

Successful psychosocial programmes integrate into and build the capacity of community structures, civil society and governmental organizations. This means focusing activities on building strengths e.g. via training, awareness, community support groups, partnerships with local structures. Provision of direct support to community members by those not deeply familiar with the context, or stand-alone services or activities that deal with only one specific issue (such as post-traumatic stress disorder) should be avoided.

- ***Structure and continuity in daily life***

Programmes should attempt to bring some 'normality' to daily life by re-establishing family and community connections and routines, enabling children to fill the social roles that are customary for children, strengthening predictability in daily life, and providing opportunities for affected populations to rebuild their lives. For example, schooling for all children should be re-established at the earliest stage.

- ***Understanding of cultural differences***

Cultural practices give people a sense of meaning and continuity with the past, which are considerable sources of psychosocial support. Grounding all psychosocial interventions in the culture, except where it is not in the best interests of the child, is both ethical and more likely to produce a sustained recovery. Those who wish to help with psychosocial healing should have a deep understanding of and respect for the societies in which they are working. Aside from the basic principles of child development and local beliefs about children, they should also understand local cultural beliefs and practices. This includes the rites and rituals related to becoming an adult as well as those associated with death, burial and mourning.

- ***Appropriate training in working with children and families***

Exploring sensitive issues with children requires skills, local knowledge, and experience. This kind of work risks tearing down a vulnerable child's defenses and leaving him/her in a worse state of pain and agitation than before. Any such work should only be carried by trained and experienced staff who can ensure appropriate support and follow-up, and work within agreed standards. In addition, any counseling related work should take place in a stable, supportive environment with the participation of care-givers who have a solid and continuing relationship with the child.

Ethical principles for conducting psychosocial evaluations¹⁴:

- ***Define the purpose of the evaluation***

Ensure the evaluation activity is necessary and justified, with a clearly defined purpose: careful advance planning is crucial – evaluators are responsible for thinking through all possible consequences and for anticipating the effect on children, families and communities.

- ***Coordinate the evaluation***

Coordinate evaluation activities with other organizations so that children, families and communities are not subject to repeated questioning covering the same or similar issues.

- ***Clarify aims and procedures***

Design the evaluation activity to get valid information: develop protocols to clarify aims and procedures for collecting, analyzing and using information.

- ***Ensure the evaluation is a participatory and collaborative process***

Ensure that the evaluation activity is a participatory and collaborative process with stakeholders and affected populations: include diverse sections of the affected population; make every effort to ensure participation is voluntary; clarify limits and consequences of the evaluation to avoid raising unrealistic expectations.

- ***Conduct consent and interviewing procedures appropriately***

Conduct consent and interviewing procedures appropriately with children or other groups: children should give their agreement to participate, but consent is also required from appropriate adults; interview procedures should reflect the need to protect children's (and other groups') best interests; interviewers should have appropriate skills and experience; ensure that functional support systems are in place to assure the wellbeing of participants.

- ***Respect privacy and confidentiality of participants***

Privacy and confidentiality of participants should be respected: information that could identify individuals should not be disclosed publicly; confidentiality is

defined as 'conditions under which the information revealed by an individual participant in a relationship of trust will not be disclosed to others without permission'; confidentiality should be breached however to ensure the immediate safety of a child or vulnerable adult.

4. Relationship to general UNICEF guidance on evaluation

For all UNICEF projects there are the standard evaluation criteria of relevance, efficiency, effectiveness, impact and sustainability¹⁵. This guide is not seeking to offer alternative evaluation criteria. It aims to provide concrete guidance on how these criteria can be assessed in the specific case of psychosocial programs.

Relevance concerns the extent to which programs have addressed important needs, and have done this according to current policy guidance. For this field, this means the extent to which an intervention has addressed the psychosocial domains of skills and knowledge, emotional well-being and social well-being for children, their families and communities.

Efficiency is generally the number of people a program has reached in relation to the resources expended. It can be seen as a measure of how well **outputs** have been achieved, given the inputs made.

The major focus of this manual is guidance on the means of evaluating the *effectiveness* and *impact* of interventions. **Effectiveness** needs to be measured in terms of the **outcomes** of a program – what has come about as a result of the programme that has made a change for children, their families and their communities? In the following sections of the manual we provide detailed guidance on how such outcomes can be measured in a way that the effectiveness of a program can be evaluated.

Impact refers to evidence that such outcomes have brought about real change in the lives of children and their communities. This is the sort of change that justified the planned intervention – has the central goal of the project been met? We make suggestions later of core indicators that can serve as criteria for this across different settings. We also describe evaluation designs that can suggest the extent to which observed changes in children and their lives are attributable to the work of the program.

Where longer-term changes reflect new or restored capacity within communities – or the services that are available to them – we begin to address the issue of **sustainability** of change.

Table 1: Key Evaluation Questions related to Psychosocial Programming prompted by the general UNICEF Evaluation Criteria

General UNICEF Evaluation Criteria	Key Questions in Context of Psychosocial Programming
Relevance	<ul style="list-style-type: none"> • Did the program articulate objectives related to changes in children’s wellbeing and lives, and that of their family and community? • Were clear needs defined with respect to required ‘levels’ of psychosocial support? • Were potential beneficiaries involved in developing programming? • Is program response relevant to identified needs?
Efficiency	<ul style="list-style-type: none"> • Have activities been delivered cost-effectively? • Has programming reached an appropriate number of beneficiaries, given program costs? • Was the program implemented in a timely manner?
Effectiveness	<ul style="list-style-type: none"> • Have stated program outcomes been achieved? • What difference has come about for children in terms of skills and knowledge, emotional well-being, and social well-being? • What difference has programming made to the skills, capacities or attitudes of families and other caregivers, and communities?
Impact	<ul style="list-style-type: none"> • Has the central goal of the project – the needs that provided the rationale for intervention – been met? • What enduring changes can be identified in the lives of children, caregivers and the wider community’s engagement with children related to programming?
Sustainability	<ul style="list-style-type: none"> • What new capacities within services or communities have been established or restored? • Are these capacities being actively used in the psychosocial support and development of children?
Coverage	<ul style="list-style-type: none"> • Has programming reached all geographical areas targeted? • Have potentially vulnerable or marginalized children and communities been reached? • Have the needs and capacities of different age groups been appropriately addressed?
Coordination	<ul style="list-style-type: none"> • Have agencies worked well together towards the common goal of improved psychosocial well-being amongst children?
Coherence	<ul style="list-style-type: none"> • Has work been consistent with the stated approach of the IASC guidelines on mental health and psychosocial support?
Protection	<ul style="list-style-type: none"> • Does the project contribute to protecting children by strengthening the ‘protective environment’?

In emergency settings, four additional criteria are used: coverage, coordination, coherence and protection.

In psychosocial work, **coverage** will mean the proportion of affected children (and communities) that have been reached by an intervention, focusing both on geographical coverage and the intervention 'reaching' sub-groups of a population who may be particularly vulnerable (e.g. children with disabilities, adolescents).

Coordination will usually mean the effectiveness of collaboration and communication amongst agencies delivering psychosocial support and other services to a community. This includes ensuring that the work of one agency neither disrupts nor duplicates the work of another, establishing common programming guidelines and strategies, coordinating geographical distribution of programmes, establishing referral mechanisms and sharing of resources and information.

Coherence means that work has been consistent with the approach and principles set down in current policy. In emergencies, this means that psychosocial programming should be consistent with the IASC Guidelines on Mental Health and Psychosocial Support in Emergency Situations.

Finally, in terms of **protection**, the issue for psychosocial programming is whether activities have worked to strengthen the 'protective environment'¹⁶ supporting children. Work related to levels 1, 2 and 3 specified by the IASC guidelines will normally be expected to have achieved this, by strengthening the capacities of children, their caregivers and families, and the wider community.

In subsequent sections, this guide explains the basic steps of putting together an evaluation of a psychosocial program, looking specifically at the two crucial issues of **effectiveness** and **impact**. For more general UNICEF guidance on evaluation, please consult existing UNICEF documentation¹⁷, including materials on results-based management¹⁸. Other resources are listed in Annex E.

5. Objectives of Psychosocial Programming

Evaluations examine how successful programming has been in achieving what it set out to do and what the consequences are for children's wellbeing. The stated objectives of the project should provide the clearest definition of what programming is seeking to achieve.

Staff responsible for psychosocial programming should be able to specify the objectives of their work in relation to the three domains of psychosocial well-being presented in Section 3: skills and knowledge, emotional well-being and social well-being. This will ensure that programming addresses an appropriately broad range of issues influencing children's well-being.

Project documents usually feature different types of objectives. These will normally include the most

Evaluations sometimes feature 'outputs' - such as the number of children involved in activities and what the activities were - and incorrectly describe them as 'impacts.'

immediate project 'outputs', expected project 'outcomes' and a clear statement of the changes in the lives of children that the project seeks to encourage: project 'impact' It is important to fully understand the differences between these terms, especially as their usage is often confused and frequently interchanged.

To strengthen psychosocial programming, evaluations need to always look at what has been achieved at the first two of these 'levels' and, as described earlier, seek opportunities for collecting evidence at the third, that of impact.

Level One: Outputs

Definition: project outputs are the planned achievements 'put out' in the process of implementing a project (such as newly trained staff or improved services or facilities) that signal that work is on track.

To identify project outputs ask: What are the visible signs that the process of implementing the project is going to plan?

Most projects will list the outputs that are expected during the course of programming. Depending on the nature of programming these might include, for example, safe play areas being constructed, teachers trained in the use of a new psychosocial curriculum, youth having attended a district sports event etc. Keeping track of these outputs is an important part of monitoring a project, for example, to see if it is falling behind its planned schedule.

Psychosocial Support through Schools for War-Affected Youth in Sierra Leone (1)

This project involved delivering a structured curriculum of activities for those in the last year of elementary school aimed at assisting their coming to terms with conflict-related events and facilitating their transition into productive work and community roles through vocationally-related, community-based activities.

The **Outputs** specified for the above project were:

- Teachers trained in the delivery of structured psychosocial curriculum (K, E, S)
- Psychosocial curriculum delivered in schools across District (E, S)
- Youth mentors trained and mentorship scheme established (K, S)

The initials in the brackets indicate the psychosocial domains to which these outputs are most relevant. Note that at the **Output** level, objectives are typically relevant to more than one domain.

An evaluation at this level is simply looking at whether the project has done what it set out to do in terms of

strengthening knowledge, services or facilities. Because such evaluations are considering the processes of delivering the project, this is sometimes referred to as 'process evaluation.' Whether the outputs achieved by a project have had any influence on the lives of children, their families and communities is the focus of the next level: outcomes.

Level Two: Outcomes

Definition: Project outcomes are the results that 'come about' during the course of a project as a result of the outputs achieved.

To identify project outcomes ask: What changes have come about for children, their caregivers or communities?

For psychosocial programs, projects usually seek changes relevant to one or more of the 'domains' of children's lives, described earlier in section 3. This might involve children learning new skills, gaining new knowledge, or having new relationships. It can also involve changes in behavior, attitudes or dynamics of children's families, or of the wider community, that are assumed to support the psychosocial well-being of children in the longer-term.

An evaluation at this level would reflect on what differences have occurred as a result of children, their families and/or communities participating in a project. It is not sufficient to assume that taking part in drama, for example, automatically increases self-worth. A measure of changes in self-worth would be needed to claim this as a project outcome.

Psychosocial Support through Schools for War-Affected Youth in Sierra Leone (2)

The **Outcomes** for this project – described in the previous box - can be clearly related to the three psychosocial domains defined earlier:

Skills and Knowledge

- Increased knowledge amongst youth about influences on well-being
- Vocational skills acquired by participating youth

Emotional Well-Being

- Increase in positive behavior and decrease in conflicts between youth
- Greater sense of hope for the future expressed by youth

Social Well-Being

- Effective mentoring relationships established for participating youth
- Community acceptance of war-affected youth enhanced

Level Three: Impacts

Definition: Project impact is the overall change in the lives of children, their families and communities that results from a project

To identify project impacts ask: What is the ultimate change relevant to children's lives that was hoped for when funding the project?

It is extremely important to see if project outcomes lead to real benefits in the lives of children. Project documentation usually begins with some statement about the current circumstances of children that are a cause for concern. Impact is a measure of the extent to which those concerns have been reduced as a result of programming. In the worked example described in this section, although achieving outcomes such as acquiring vocational skills, gaining a greater sense of hope and establishing mentorship relationships was welcome, when the project was funded these were seen as a means to an end, not an end in themselves. The 'end' – the goal – is expressed in terms of targeted Impacts, as below.

Psychosocial Support through Schools for War-Affected Youth in Sierra Leone (3)

The **Impacts** targeted this project can again be linked to the three psychosocial domains defined earlier. These **Impacts** clarify the changes that the program aimed to achieve by securing the preceding **Outputs** and **Outcomes**. These **Impacts** reflect the stated goal of the project to assist youth in their coming to terms with conflict-related events and facilitating their transition into productive work and community roles.

Skills and Knowledge

- Youth using vocational skills to support livelihoods of their households

Emotional Well-Being

- Youth seen as well-adjusted and positively engaged in life of community

Social Well-Being

- Youth assume valued social roles within community

Evaluating impact means measuring the change in the lives of children, families and communities that is attributable to a project. Impact evaluations are thus concerned with detecting both the intended and unintended consequences of psychosocial programming. For instance, while children associated with armed forces might have gained skills from a programme that targeted them, when they return to their communities this targeted approach can lead to resentment and stigmatization from their peers who did not have the chance to participate in the programme.

It can be difficult to measure such impact, as it usually requires following-up on children some time after the end of the project. But it is very important in judging

the real worth of our programs. As noted earlier, the psychosocial field is at sufficiently early a stage of development that a significant proportion of program evaluations need to look at the evidence that achieving project outputs and outcomes have led to real, positive, sustained impacts in children's lives.

Organizations face many challenges in trying to demonstrate the impacts of their work. One of the major reasons for difficulty lies in the project planning stage of most psychosocial programs. Most projects simply do not establish a comprehensive set of clearly defined objectives¹⁹. If objectives remain vague or perhaps are implicit, it will be difficult to gauge progress of any kind. Thinking through what is targeted as the long term impacts of programming is crucial. With such a clear goal in mind, it will be easier to set objectives for project outputs and outcomes that should lead towards its achievement.

Impact evaluations require considerably more resources than outcome evaluations.²⁰ Resources such as funding, staffing, time, technical assistance will need to be identified for impact evaluations and it is vital that this planning begins at an early stage. Based on the recommendations of Section 2, this means identifying the necessary resources for their completion within the budgeting and planning cycle. Without explicit identification of resources in this way it is unlikely that sufficient resources will be available.

In Ethiopia, Save the Children, Norway found that the placement of separated children in foster homes (a planned outcome of their programming) resulted in the exploitation of girls after they were adopted (clearly, an unintended impact). The girls' experiences were only discovered during an impact evaluation ten years after the project for separated children had finished and the girls were adults.

'It can be difficult to measure impact...but it is very important in judging the real worth of our programs'

Setting Objectives for Children, their Families and Communities

It is very important that the objectives of psychosocial programming – whether these are outputs, outcomes or impacts - are not decided upon by project staff alone, but with active engagement with beneficiaries and other relevant stakeholders (refer back to section 3 for a discussion of principles that should inform such participative working).

When objectives are discussed in this manner it is usual for them to address not just aspects of the lives of children themselves, but also issues related to their families, and the wider community, that affect the circumstances of children. The domains of skills and knowledge, emotional well-being and social well-being can

again be used to prompt discussion about a suitable range of objectives. At family and community levels the ‘dividing line’ between such categories can be hard to draw, but the aim is not to worry so much what category an objective belongs to, as to ensure that a suitable range of aspects of psychosocial well-being are addressed in one way or another. Table 2 summarizes the potential objectives of psychosocial programming that were identified in the course of a consultation with a range of psychosocial programming staff working to strengthen psychosocial support to children and youth in Palestine.

Table 2: Sample Objectives across the Core Domains resulting from a Palestinian Consultation

For Children	Objectives
Skills and Knowledge	Increase in the use of locally defined life skills by children (e.g. adolescents are able to make informed decisions about vocations and career pathways)
	Increase in the percentage of children who are able to say with concrete examples what they plan or would like to be doing next year
	Increased engagement of children in school and in community activities
Emotional Wellbeing	Increase in the self-confidence, playfulness and sense of security of children attending activities
	Increased sense of locally defined purpose and meaning
Social Wellbeing	Children are more cooperative and less violent in relations with peers
	Increase in children’s ability to assume socially appropriate roles (locally defined)
	Increase in the number and quality of relationships with supportive adults (primary caregivers or community adult role models)
For Families	
Skills and Knowledge	Improved communication skills among family members
	Increase in ability of families to materially support themselves
	Improved parental ability to address conflicts non-violently
Emotional Wellbeing	Increase in percentage of parents who actively discuss problems with their children
	Decrease in parents/caregivers use of violence
	Increased ability of families to cope with external stressors
Social Wellbeing	Increase in engagement of parents in activities that support children’s development
For Communities	
Skills and Knowledge	Increased open debate on psychosocial issues in public forums (media, community meetings etc.)
Emotional Wellbeing	Increase in access to psychological and social services by vulnerable/marginalized groups
	Reduction in levels of distress of front-line workers
Social Wellbeing	Community involvement in social and cultural activities
	Decrease of interfamily violence and decrease of casualties as a result of interfamily conflicts
	Increased number and quality of social support mechanisms in the community

6. Indicators

One of the dangers of trying to do outcome and impact evaluation without the necessary resources and expertise is that questionable methods for assessing may be adopted. For example, an evaluation might say that a project had a positive impact because children who were ‘sad’ before an activity began were now ‘happier’. This might be measured by an evolution in their drawing – choosing frowning and smiling faces to represent how they were feeling before and then after the intervention. But we do not know if such a trend reflects a real change in the lives of children in terms of their skills and knowledge, emotional wellbeing or social wellbeing in settings outside of the project. We need to be clear in advance, what we would count as ‘evidence’ of targeted change. This takes us to the issue of indicators.

Definition: an indicator is a simple, clear statement that helps measure and communicate change.

For all objectives we need to define how we would measure results. This involves identifying **indicators** for such objectives, i.e. what it is you want to measure and how you will measure it.

Core Indicators

Core indicators are indicators which should be used for all evaluations. Using core indicators in all UNICEF psychosocial programs will help in attempts to establish a more robust evidence-base for psychosocial programming. Over time it should then be possible to compare results across programs and countries.

However, given the different ways that psychosocial wellbeing is reflected in different societies it is not possible to suggest a number of specific ‘one size fits all’ indicators. Rather, we suggest the choice of indicators of local relevance that link to the critical

‘Core indicators are indicators which should be used for all evaluation’

‘Using core indicators in all UNICEF psychosocial programs will help in attempts to establish a more robust evidence base for psychosocial programming’

domains of **skills and knowledge, emotional wellbeing and social wellbeing** which were highlighted in the introduction to psychosocial programming.

For most UNICEF psychosocial programming it should be appropriate to specify outcome and impact indicators for each of these domains as follows:

Domain	Core Indicator
Skills and knowledge	Some measure of acquisition of skills
Emotional wellbeing	Some measure of improved emotional adjustment
Social wellbeing	Some measure of improved social functioning

To measure achievement against such indicators we need to use methods that are valid, accounting for cultural variations in understandings of what defines children’s well-being. We recommend a ‘mixed method’ approach to such measurement, drawing on both quantitative and qualitative methods (see Section 9). We can also use the availability of existing information to help define relevant and practicable measures of change.

Using Existing Sources of Information to Define Indicators

Although evaluation involves developing some means of collecting information about programming, we should not ignore the potential value of information that has already been collected by others. Existing sources of information can often provide valuable insight into the experience of children and their communities.

In developing indicators to assess the achievements of a project, we recommend considering adding indicators that would be informed by existing sources of information.

The examples below are illustrative of how existing information can be used in this manner. The use of existing information in identifying indicators clearly depends on what the project is aiming to do, and the nature and quality of information available.

- a program focused on parent education might consider as an indicator the percentage of children attending clinics for immunization (calculated from data routinely collected by health workers)
- a program addressing girl empowerment might specify as an indicator the proportion of girls enrolled in school (which could be obtained from school records)
- a program addressing non-violent conflict resolution could use the number of police reports of offences by youths as a measure of impact on local conflict.

Developing An Indicator Framework

With project objectives (a) at output, outcome and impact levels; (b) across the domains of skills and knowledge, emotional well-being, and social well-being; and (c) potentially targeting children, their caregivers and/or the wider

community, the task of identifying indicators may seem a very complex one. However, the basic rule is simply that if there is an objective there should be a means of measuring if it has been achieved.

Table 3 shows an attempt by UNICEF staff working in Sri Lanka to begin to define indicators for their psychosocial work in the east of the country. Notice how the different levels of objectives, and the different domains of psychosocial well-being, have been used to specify what changes are expected as a result of their psycho-educational projects. At the output level – where the delivery and coverage of the drama and theater is the main measure of progress – it was not considered meaningful to list separate indicators related to each domain. Notice also that a number of the impact indicators make use of existing data that is routinely collected.

Table 3: Developing Indicators for Psychosocial Programming in eastern Sri Lanka

Domain	Skills and Knowledge	Emotional Well-Being	Social Well-Being
Indicator			
Impact	Quality of care provided to children by families improved <i>[measured by interview with children and parents, and by reduced incidence of reports of child abuse]</i>	Strengthened referral networks between communities and service providers <i>[measured by increased case loads of community referrals]</i>	Reduction in number of conflicts between IDPs and host communities <i>[measured by community interview, and by reduction in police reports of social disturbances involving IDPs]</i>
Outcome	Increased awareness of parenting skills and children's rights <i>[measured by increase of knowledge in parents and children]</i>	Appropriate referral of children requiring specific support <i>[measured by increase of referrals through District Child Protection Committee]</i>	Strengthened relationships between IDPs and host communities <i>[measured by frequency of community activities featuring both IDPs and host communities]</i>
Output	Psycho-educational program delivered through drama and theatre activities for children and families in all IDP camps and host communities in Batticoloa <i>[measured by number of program sessions delivered, number of sites for delivery]</i>		

Refining Indicators

We defined an indicator earlier as ‘a simple, clear statement that helps measure and communicate change’. In the case example above, the team will need at

some stage to ask ‘what do we mean by increased knowledge of IDP rights and services’? Or ‘how much reduction in IDP-host community conflict would represent a good impact?’ There are a number of sources of guidance available to help project evaluation teams develop unambiguous, measurable indicators of this sort²¹. However, as with many other issues, these are best addressed at the planning stage of a programme, when they can define key milestones and targets for programming.

Section 10 provides worked examples, and Annex A a template for developing and refining objectives and indicators for a psychosocial program, using the ideas in this section.

7. Evaluation Design

Evaluations need to be carefully designed so that you can reliably draw conclusions from them. Measuring the extent to which an objective has been achieved is usually not enough to show the value of a project because the change may have come about because of something other than the project. If we find out there has been progress towards some outcome, we will generally be pleased. But we may not know if this progress would have happened anyway, without the intervention. Or maybe another intervention would have made far more progress. Having some sort of comparison group helps us interpret our evaluation findings. Designs help us make an appropriate comparison.

‘To show that programming has made a real difference we need to make some sort of comparison’

Baselines

A baseline is a measure of something before programming begins. Good baseline measures provide a basis for measuring these same things at the end of the project, and seeing what change has occurred. This sounds deceptively simple. However many programs do not establish adequate baseline data on children’s wellbeing and so there is no reliable way to measure changes over time.²²

‘Many programs do not establish adequate baseline data on children’s wellbeing and so there is no reliable way to measure changes over time’

It is very important to establish baselines on relevant indicators before beginning psychosocial programming. Evaluation designs that have a clear measure of how things were before the start of a

In Sierra Leone, an IRC evaluation did a baseline assessment in 2002 followed by an interim evaluation in 2004. Measures of confidence, prosocial behavior, depression, anxiety and hostility were taken.

program are much stronger than those that don't have such measures. Needs assessments/situation analyses may have been conducted before initiating programming (as is good practice) and if done correctly, then this data may provide a relevant baseline. Otherwise, it is recommended to always invest time in establishing a clear baseline.

Section 9 provides guidance on collecting baseline information, and how many people (whether adults or children) to collect information from. If such work is tackled in a focused way there is little risk that interventions will be significantly delayed – a common fear for agencies working in emergency settings. And conducting a baseline potentially provides a hugely useful base for program planning and development, as well as evaluation.

The American Red Cross put in place a comprehensive baseline survey of psychosocial support before their post-Tsunami interventions in the Maldives and Sri Lanka.

Comparison Groups

A comparison group that receives no intervention (sometimes referred to as a 'control group') is a group of children who in every way – except their not receiving an intervention – are as similar as possible to the children receiving the intervention through the project. There is a normal path of healing after disaster in individuals and communities. It is critical to be able to attribute the results of interventions by also measuring results that occur from instinct and adapted community and family behavior.

There is often a fear that conducting any assessment creates expectations of delivering an intervention. This discourages agencies from approaching people in an area where there is no immediate prospect of intervention – but similar levels of need – as a source for a comparison group. However, the risks of creating harm by falsely raising expectations need to be balanced against the risk of creating harm by delivering an 'untested' intervention (see section on principles of psychosocial support p.13). Providing a comparison group is established sensitively, potential harm should be minimized and this will be better than having no comparison group at all.

A study looking at the experience and needs of children abducted by the LRA was considerably strengthened by looking at a 'comparison group' of children not abducted. This established that the former typically had a year less education, were twice as likely to report family difficulties, were three more times likely to have a physical impairment, but were little different in terms of psychosocial needs.

One way of establishing a comparison group arises in the common practice of rolling out programs over time. In emergencies this may happen, for example, when a program is initially introduced in one refugee camp and then it is planned to repeat it in a second camp. It would be possible in these circumstances to compare a group of children in the first camp with that in a second. Alternatively it may be possible to set up comparison groups between children currently enrolled in a program with those waiting to be enrolled.

In a phased-in education intervention in 75 Kenyan schools, an evaluation was able to capture the effects of the program by comparing those schools currently participating with those awaiting their turn.

Establishing a comparison group should be given the highest priority by program managers in program planning decisions. Interventions are very rarely 'rolled out' simultaneously in multiple areas. Careful planning at this stage will often provide access to an effective comparison group. This is an important way of establishing the effectiveness of an intervention.

Another way of making some form of comparison between those who receive an intervention and those that don't is by using information about local norms. If there is reliable government data already available about girls' rate of enrollment in school, for instance, this could be used by a program seeking to foster girls' access to schooling as a basis for comparison.

The evaluation of programming with former child soldiers in Mozambique reviewed in Section 10 compared the circumstances of former child soldiers with local norms for social adjustment and economic well-being.

A final strategy for having some basis of comparison is to compare two interventions with each other. If it is not clear which approach is more effective in work with children – and that is often the case with psychosocial programming with children – then this can be an effective way of learning what works best. This strategy should not be used as a way of avoiding the question of how children fare without any intervention, however. Identifying a non-intervention group would also be preferable in this circumstance.

One of the most powerful designs is one where we combine the features of baseline and comparison groups. We take measurements at the start of the project (i.e. a baseline) and at various points during the project, accessing those who receive the intervention and those who don't (or a proportion: see section 9). At the end of the project we are then in a strong position to conclude the changes can be linked to their being in the project intervention.

This is the design that should be adopted for the majority of UNICEF evaluations of psychosocial programs. To achieve this design program managers need to set down the foundations for subsequent evaluation during the program planning stage, and certainly before intervention begins. Where because of limited funding or programme scale it is not possible to reach all those affected, it is still possible to use this design.

Annex C identifies the key actions that program managers can take at the stage of program planning to make subsequent evaluation more effective and informative.

'This is the design that should be adopted for the majority of UNICEF evaluations of psychosocial programs'

In Indonesia, UNICEF made comparisons between children who participated in psychosocial programmes, those in EMDR (eye-movement desensitization) programmes and those who because of limited funding had not participated in any programmes. These comparisons provided evidence of the benefit of psychosocial programmes.

Looking Backwards

Sometimes we have to try to evaluate a project when there were no good baseline measures taken. This is a weaker design, in the sense that it is harder to be able to show that an intervention has made a difference – but it is not impossible.

Essentially, we have to 'reconstruct' some sort of baseline. There are a number of ways of doing this²³:

Using existing documents

'Secondary data,' such as data from health and education agencies, government surveys, school enrolment and attendance records, project records, can be used but should be assessed for their reliability and validity.

Using recall

'when there were no good baseline measures taken... we have to 'reconstruct' some sort of baseline'

It is not a good idea to expect children, families, communities or workers to be asked to remember how they felt or behaved some time in the past. This information is not usually very reliable. However, it is sometimes possible to define key events (like being enrolled in school, or getting paid work) that people can recall that can be used to identify changes that have happened between a point in the past and the present. Reliability and validity of this method can also be tested.

Asking key informants

By asking key informants for information, it may be possible to verify data from other sources, thus establishing some consistency about the baseline you are reconstructing. Key informants may be people such as community leaders, teachers, nurses and doctors.

In evaluating its programmes with girls who had been abducted by military groups, CCF in Sierra Leone was able to develop a 'local calendar' of key events in the community since the end of the war. With respect to these events, girls were then able to identify when they had returned from the bush, received traditional cleansings, married and taken other steps towards reintegrating within their community.

To conclude this section, we repeat that UNICEF evaluations should normally plan to use a design with a baseline and comparison groups. Where this is really not possible – post rapid onset emergencies like the tsunami, for example - then establishing adequate comparison groups should be the priority.

8. Preparing for an evaluation

Ensuring participation

Evaluations should as far as possible be a participatory process involving beneficiaries. In promoting communication between those involved in a program, the resulting evaluation should reflect the aspirations and concerns of beneficiaries and also identify emerging needs. In essence, by involving beneficiaries in assessing and reviewing program aims and objectives a more meaningful level of accountability is possible.

The participation of children is vital:

- **It is their right**
- **It leads to better programming and evaluation**
- **It strengthens their psychosocial well-being**

Article 12 of the Convention on the Rights of the Child states that all children have the right to participation in decision-making processes that may be relevant in their lives and to influence decisions taken in their regard within the family, the school or the community.²⁴ Children should be involved in the planning, implementation and evaluation of programs. Their involvement adds authenticity to all these processes.

‘Encouraging activities in which children take responsibility, help and support one another according to principles of tolerance and non-discrimination are particularly valuable.’

Involving children in program evaluations may itself have significant benefits for children’s wellbeing. Participation may empower children by valuing their opinions and enhancing their knowledge and skills. Opportunities for social interaction also contribute to children’s psychosocial wellbeing being linked with social and emotional development.²⁵ Encouraging activities in which children take responsibility, help and support one another according to principles of tolerance and non-discrimination are particularly valuable. There are of course ethical issues related to the participation of children. The principles listed in section 3 for ensuring that the risk of harm is minimized provide guidance in this area.

Although this guide is about evaluation, the participation of beneficiaries should be considered at all stages of a program, from planning to implementation and evaluation. Actively promoting and sustaining participation is a process whereby relationships are built and forged over time. It is not a one-off event. It should be built formally into program planning. This will require flexibility and commitment to make it work, especially in complex emergencies.

Objectives can be set to ensure that participation processes are included in the program, for example:

PLANNING	IMPLEMENTATION	EVALUATION
establish a baseline with parents and children regarding the changes they want to see	run a feedback ‘game’ with children at the end of every activity and record findings	children and parents help define objectives and indicators children/families/communities provide local understanding of wellbeing

There are of course different levels of participation. Arnstein’s ‘ladder of participation’²⁶ provides a useful prompt to understanding the way we might be ‘involving’ children. The ladder has 8 rungs spanning participation, with ‘manipulation’ at one end and ‘citizen control’ at the other, and ‘consultation’ and ‘partnership’ being in the middle.

For the most part, children are 'consulted' in the course of evaluating a psychosocial program. They may for example be asked their opinions about the activities they have been involved in and take part in exercises to measure the progress of the program. They are less likely to be part of the evaluation team, working as facilitators with to review a program, for example, and, in Arnstein's terms, working in 'partnership' with adults to complete an evaluation.

If in the course of evaluation, children are most likely to be 'consulted' as respondents, what is their role in terms of *planning* for evaluation? From the few examples of their participation at this stage it looks like they rarely have a role. However it is possible to achieve this so that children's views of what are measures of success are included in, and inform, the evaluation process. All evaluations should have in place as a minimum the following features of children's and community participation:

- determining objectives and appropriate indicators for the project
- determining local definitions of wellbeing
- providing their views on the project
- providing feedback on draft results and the implications for their community/future similar projects

Moving The Goalposts, Kilifi

This Kenyan NGO working with girls and young women through the medium of football, worked with a research team to develop a participatory monitoring and evaluation strategy²⁷. The research team was made up of an experienced researcher, two research assistants who were members of the NGO and were nominated by their peers, plus a female translator. In addition, a planning group of 15 girls and young women formed as a self-selected sub-group of the NGO's 'girls' committee.' The research team and the planning group worked together to formulate a list of indicators which could be used for evaluation.

For example, as indicators of 'self esteem' the group suggested:

- feeling good about yourself (feeling strong, fit and active; feeling good about doing something) and
- feeling that you're OK despite what others might say about you (having a positive body image; being confident to try a new thing; ignoring the bad things people say about you)

In their discussion of the process, the researchers noticed how participation was influenced by power relations, culture, attitudes and skill levels. For example, the planning committee combined girls and young women of ages 11 to 21 years and of different educational experiences. In discussions, girls tended to wait to hear what older or more experienced members said before giving their opinion. However the process overall demonstrated that, with support and opportunity, youth-led evaluations can be achieved.

Approach

Thinking through carefully how you will introduce yourself, how you will explain the work, and how you will deal with participants in a sensitive and courteous way is essential. You also need to decide how to ensure that ethical considerations (e.g. ensuring appropriate confidentiality and minimizing any risk from participation) are addressed (see section 3 regarding ethical principles). All methodologies should be gender sensitive, culturally appropriate, and child focused. The Population Council's guide 'Ethical Approaches to gathering information from children and adolescents in international settings – guidelines and resources'²⁸ provides extensive detailed guidance in this area.

We also highlight here some general points about working with children²⁹. They are very fundamental and demonstrate practically how to engage with children in respectful ways:

- Facilitators should choose activities where children enjoy themselves and have fun. Keep sessions to a reasonable length so that children are not tired or bored. Provide drinks and biscuits and plenty of time for physical games.
- Facilitators should group children together in age ranges that allow children to work together well, and in groups of a size that will not prevent each child from taking part. Choose spaces where children will have room to do activities so that they don't copy each other
- Facilitators should be sensitive and have good listening skills, and address tensions or disputes that have been prompted by activities
- Facilitators should not feel that there is a correct or specific answer but that all children's ideas and opinions are respected
- Facilitators should start with easy questions and activities which help children relax, before moving on to more sensitive topic, where appropriate.³⁰

Equipping Staff

We also need to consider the capacity of project staff to engage effectively in evaluations. Tasks are likely to involve frontline staff as well as UNICEF country officers. This means that training will be needed to prepare all those involved, and that supervision and support will need to be in place.

There are a number of resources that may be helpful in training and development of staff in evaluation work.^{31,32} Here is a list of core competencies identified after a review of frontline staff's successes and difficulties in implementing evaluations³³. The list describes the skills needed to work with children in psychosocial programs such that staff can plan, engage, record and analyze activities to evaluate the effectiveness of those programs:

- Facilitation, communication and negotiation skills in order to manage the monitoring and evaluation activities (i.e. explain the tools clearly to the children, encourage participation, explore/elicit the children's views, manage group dynamics);
- Conceptual and practical knowledge related to aspects of psychosocial well-being, to be able to understand the meanings/implications of children's life experiences and their responses to these;
- Activity planning skills, in order to implement monitoring and evaluation activities coherently;
- Analytical capacity, problem-solving skills and a capacity for creativity/innovation in order to interpret and synthesize the outputs from methods or adapt these to the needs of a specific circumstance;
- Note-taking, process-recording and observational skills, to ensure that (...) information is accurately recorded;
- Confidence / pro-active attitude / caring and warm attitude towards children.

9. Methodology

This section provides guidance on the kind of tools you can use to do an evaluation. We recommend a 'mixed methods' approach which involves both qualitative and quantitative methods of collecting information. This section also provides guidance about sampling i.e. the number of people needed as participants to ensure that the evaluation is able to draw valid conclusions.

We recommend that evaluations use both 'qualitative' methods (i.e. those more focused on description) and 'quantitative' methods (i.e. those focused on 'numbers'). Both have their particular strengths, and information from one can usefully complement the information provided from the other. Evaluations are improved if the same issue is considered from a range of methodological perspectives. If accounts from different methods produce a similar picture, it increases confidence in findings. It also potentially deepens analysis. Collecting information from different sources in this way is known as 'triangulation': viewing something from different perspectives helps build a fuller picture.

In an assessment of Palestinian children, a mixed method of focus groups discussions and quantitative data analysis allowed themes identified by children to be ranked. 95 focus groups were held. At the end of each focus group facilitators recorded the 3 most prevalent responses for each question. The combined information across all the groups showed that the majority of children maintained high hopes of developing themselves personally and/or academically in order to be able to meet the needs of the future.

Qualitative Methods

Using qualitative methods can provide vivid insights into the experience, perceptions and beliefs of program beneficiaries and other relevant stakeholders. Used appropriately such methods can provide robust, insightful information to support an evaluation. There are a range of different methods and tools which encourage the active participation of adults and children.

'Using qualitative methods can provide vivid insight into the experience, perceptions and beliefs of program beneficiaries.'

These kinds of methods are valuable in identifying the resources and strengths people have in dealing with adversity, as well as the challenges that they face. Most participative methods tend to be collective rather than individual, based on interaction and collaboration in groups. Usually a variety of methods and tools are used, based on the principle of triangulation noted above. Methods and tools aim to be culturally sensitive and valid, drawing on local understandings, resources and contexts.

Although the activities described may appear unobtrusive and unthreatening, any work of this kind with children may trigger strong emotions. This needs to be anticipated and adequate support put in place to support children in distress. This should reflect the principle of 'do no harm' and use of the ethical guidelines discussed earlier.

Focus groups

Focus groups can be very helpful in encouraging participants to express their thoughts and experiences, without being too obtrusive. A series of questions is explored in a systematic way by a focus group, with the facilitator posing the questions and being responsible for recording the responses. The facilitator may probe certain key issues or concepts that emerge in the discussion. The key feature of focus groups is that group members build on each others' responses by, for example, adding detail or correcting one another, and in this way the information gathered is likely to be more accurate.

Assessing Afghan Children's Psychosocial Well-being: A Multi-Modal study of Intervention Outcomes³⁴

In focus groups of 8-10 people, in age groups 7-13, 14-18 and over 21 (separate for boys and girls), four main questions were addressed:

What are children's main worries?
Which are the most severe worries?
What do children do to cope or to manage their situation?
Which strategies work best to solve their worries?

Scenarios were included as a means of providing concrete situations for participants to respond to:

An Afghan boy is upset. Why? When are Afghan boys upset?

An Afghan boy is on his way to school and an older boy stops him and takes his books. What does the younger boy do in this situation?

Focus groups can be used in their own right or in conjunction with another tool to crosscheck information obtained. In the above example given from Afghanistan, differences between adult and children's perspectives were found through convening focus groups with people of various ages. The focus groups identified gaps in some villages between what adults said children worried about and what children themselves said they worried about. For example, none of the men who were interviewed in one village indicated the lack of water and toilets at school as a significant worry for young boys. The young boys themselves, however, ranked this among their top three worries. In another village women said young girls worried about being poor and having no access to a clinic. But the girls themselves said their main worries were getting sick from sun exposure, being yelled at by teachers and being injured in traffic. The report noted that these differences prompted further program planning around parent-child communication.

Key informant interviews

Key informants such as community leaders, teachers, caregivers, doctors, local government agencies and NGO staff can be interviewed individually to provide information about a particular community or issue. Key informants do not have to be people in positions of authority; it is good to get a range of perspectives and thereby access different sources of information about the matter in hand.

Mapping tools

Mapping is a generic term for visual information, which is sometimes literally presented as a map. Mapping is usually a good starting point for participatory work because it involves children in drawing a map of some kind which is then used to generate discussion. There are many different kinds of mapping tools. Children may be asked to draw things such as: the locations and activities of their day; the people they spend time with in the course of their day; the places where they perceive risks or fears in their community.

In the example given from Kabul, the children involved were living in an internally displaced persons' camp. They were asked to draw their immediate surroundings and all the places of physical danger they could identify.

The ranking of 'dangerous places' was arrived at by listing the places children drew in order of their frequency. In this case the mapping exercise showed that children were not overly occupied with memories of past distress, but instead were impacted by their current surroundings and wellbeing. The report stated that 'although children do have bad memories of the coalition bombing, the repression of the Taliban and fleeing from the war, the past is experienced more in the way it impacts children's current surroundings, relationships and well-being'.

Another mapping tool is the *risk and resource map*. This involves children drawing a map of their immediate surroundings and community and other areas they frequently visit. This method identifies the things that children find threatening and the things/people/institutions they see as sources of support and protection in their daily lives.

A risk and resource map was administered with children in six Badulla Road villages in Sri Lanka³⁵. This enabled comparison of the circumstances of children

For children in Kabul, the most frequently drawn places of physical danger were:

- ***traffic on the streets***
- ***destroyed houses***
- ***mined areas***
- ***'places where mad dogs are'***
- ***military posts***
- ***airport***
- ***mountains for firewood***
- ***high walls***
- ***electricity lines***
- ***'places where mad people are'***
- ***open wells***

in different communities, allowing the identification of both the crosscutting issues that affect all children in the region as well as those that were village specific. For example, snakebites featured as a major source of fear in all of the villages, whereas traffic accidents were only mentioned in one village. In this way, the maps indicated the extent of an issue.

In their evaluation of this tool, the team said it was easily understood and generated a lot of very useful information and discussion. However they noted that the resources children identified tended not to have a particular bearing on the risks they named. They suggested linking a focus group discussion with this exercise to discuss the possible links between risks and resources with the children.

Free listing

The aim of free listing is to identify the criteria by which wellbeing is understood in a particular culture or community.

In Northern Uganda, for example, this method was used to identify what children, parents and teachers perceived to be the characteristics of a 'resilient child' (a phrase that had meaning and significance in *Acholi*). This involved children developing a long list of the qualities they associated with resilient children. The listed qualities were then discussed and grouped these under 6 major headings: playful & sociable; intelligent; happy; respectful; responsible; and healthy. The evaluation was then focused around these local 'indicators' of resilience. A similar exercise was conducted with parents and teachers which revealed – despite some overlap – subtly different perspectives on what signaled children's wellbeing.

'Free listing...was used to identify the characteristics of a 'resilient child' ... as indicators of wellbeing'

In another example³⁶, a 'Wellbeing Exercise' was adapted for use in Sri Lanka. Participants were asked to think of a child they knew who, in their view, was doing well in life. They were then asked to think of the things about this child that indicate to them that he or she is doing well. The characteristics that emerged were then used as indicators of wellbeing.

Creative self expression

Drawing, drama and other forms of self expression can be used to gain insight into children's needs, concerns, ideas and opinions. Drawing, for example, provides information in terms of the content of the drawing itself as well as through the child's explanation of the drawing. This is a popular technique but It is crucial that the needs, concerns, ideas and opinions gathered in this way are explained and interpreted by the children involved themselves and not by the adults facilitating such activities.

Particular forms of drawing can be used to address key issues in the lives of children. For example, with a ***timeline*** children are asked to draw the important events and changes that have occurred in their community over a certain time period. When all the drawings are complete, the children present their timelines to one another and discuss the different events and when they happened.

In a ***lifeline***, children draw major events in their own lives across a horizontal line, placing positive events above the line and negative events below the line. A lifeline is drawn by joining all the events together from left to right.

A review of the use of these techniques in Sri Lanka suggested that 'Timelines provide extremely useful information about the historical and environmental forces that have an impact on children's lives. The advantage of using a timeline over recording children's individual life histories is that it is a less invasive and threatening way of gathering potentially sensitive information.'³⁷ Hart and colleagues review a number of other methods using creativity as a route to gain insight into the perspectives of children on particular issues.³⁸

In evaluating psychosocial programmes addressing the needs of children affected by conflict in the Balkans, Kostorova-Unkovska and colleagues have used the medium of drawings and drama for children to describe the impact of activities on them and their way of thinking.

'Timelines provide ... a less invasive and threatening way of gathering potentially sensitive information'

Developing Local Measures

With quantitative methods there are two main approaches. One is to develop a questionnaire or interview guide or whatever 'from scratch'. This can have the advantage that it reflects the core concerns of the project, and attempts to use ideas appropriate to the local culture. But it can be very time consuming, can be done very badly, and means that your findings will not be comparable to those in other situations.

If this first approach is taken, the aim should be to produce a simple listing of indicators that reflect local understandings of well-being or adjustment. This builds upon the ideas of 'free listing' discussed earlier. Local participation informs how these indicators are defined, integrating local community values and ways of understanding into the process of evaluation. The resulting listing then becomes an 'agenda' against which the performance of the program can be judged. Checking this 'agenda' against the core psychosocial domains is a way of making sure that key issues have not been missed.

Using Existing Measures

The other major approach with quantitative methods is to use established measures that have been used in other studies and evaluations. For example, some organizations use standardized psychosocial assessments as measurements before and after programs take place. This has the advantage of building on the work of others, and being able to make comparison with other situations. But, done carelessly, a measure may be used that is meaningless in the specific culture of the project or misses crucial culturally relevant aspects of psychosocial wellbeing. Tools imported from one culture to another without careful consideration can result in misleading comparisons.

Pre-existing psychosocial assessments should only be used when there is clear evidence of their validity

Communities in Sri Lanka were asked to specify the characteristics of a child that was 'doing well'. Each characteristic was recorded on a single card, and the resulting cards sorted into piles of what were seen as related characteristics by a number of different villagers. This identified a number of 'domains', and a number of specific 'indicators' within each. The domains included characteristics related to skills and knowledge, emotional well-being and social adjustment – each defined in relevant local terms.

The Child Behavior Checklist (CBCL) has been used in many cultural contexts, and there are 'norms' for many settings. Similar 'norms' are also available for other measures – such as the Hopkins Symptom Checklist-25 (HSCL25) and the Impact of Events Scale for Children (IES-R) – though these are best suited for work focused on relief of symptoms of mental ill-health.

in a given cultural context. This usually means that they have been used before and that 'norms' (typical scores for children in that context) have been established. In practice this leaves very few options for appropriate use of such measures. The Child Behavior Checklist³⁹ is one of the few measures that meet these criteria. Other widely used assessments – such as the Hopkins Symptom Checklist and the Harvard Trauma Questionnaire – may often have local norms, but generally address clinical symptoms rather than the broader psychosocial needs which are the focus of UNICEF programs.

Sampling

Definition: a sample is a 'subset' of a population which we aim to be representative of the entire population.

Sampling is perhaps the most frequently ignored part of evaluations. We need to make sure that we don't select people who were, for example, particularly pleased with the project, or who were unrepresentative of beneficiaries because they live particularly close to where an intervention was delivered. We also need to make sure that we speak to enough people to feel confident that we can draw general conclusions about the experience of program beneficiaries. These kinds of selection issues are relevant whether we are collecting information through interviews or focus groups or any other method.

'Sampling is perhaps the most frequently ignored part of evaluations'

Sample Size for Quantitative Methods

For some evaluations it may be possible to collect information from all those who are, or were, beneficiaries of a project. This may be possible for small projects, but for interventions on a larger scale we have to choose to collect information from a sample of beneficiaries. And even where it is possible to contact everyone, it is not advisable as it is unnecessary and is not an efficient use of resources.

For quantitative research methods – discussed below – there are mathematical formulae to work out the

size of the sample required⁴⁰. The key factors that decide the required sample size are the **amount of change** we can reasonably expect a successful intervention to produce and the **degree of precision** we need in our findings.

Perhaps surprisingly, the total number of beneficiaries who have received the intervention (generally known as the ‘population’) doesn’t greatly influence the size of the sample required.

You can calculate the number required for a specific evaluation using the relevant formulae if you wish. But based on assumptions that will usually apply for psychosocial programs⁴¹ **a sample of between 80 and 100 people will be adequate for most evaluations**. Targeting at the upper end of this range allows for ‘drop outs’ (that is, participants withdrawing before all information has been collected from them) without threatening the validity of the evaluation.

This number applies to each distinctive ‘population’ of beneficiaries. If an intervention is delivered throughout all schools in a District, all the children attending school potentially comprise a single ‘population’. But if an intervention is delivered in two refugee camps with very different resources and cultural practices, it would be appropriate to consider each refugee camp the basis of defining a separate population. In this case you would need to get a sample of 80-100 children who have received the intervention from each camp.

Also if a comparison group is being used, a similar number (i.e. 80-100 children) needs to be sampled from those not receiving the program being evaluated.

Sample Selection for Quantitative Methods

The other key factor in sampling is how potential participants are selected. We want to ensure that everyone has an equal chance of being selected for the sample.

Random Sampling. The best way to ensure that a sample isn’t ‘biased’ in some way is to randomly

‘We recommend that evaluations [using quantitative methods] seek a sample of between 80 and 100 people’

In the example given in section 10, the evaluation enrolled as participants children who were to receive in psychosocial programs in various sites across Gaza and the West Bank. Given the differences between the situation in these two parts of Palestine, 150 children (well above the 100 recommended here, but allowing for significant ‘drop outs’) were sampled in each area. However, only 50 children from each area were recruited as a comparison sample (those not receiving services). This number is below the 80 minimum recommended here and put the evaluation at risk of having an inadequate basis on which to judge the difference between those receiving the intervention and those not.

select participants from a comprehensive list of those that could potentially be involved.

For example, if we are looking at the impact of a classroom-based psychosocial curriculum, we want a list (probably based on the school register) of all children who received the intervention in a school. From this list we can then randomly select children to interview. If there are a number of schools that have used the curriculum, we would take a list of all those schools, and randomly select from that list. Choosing at random in this way removes the risk that we will (consciously or unconsciously) choose the most enthusiastic children to speak with, or the 'best' schools to visit. Doing this would mean that our findings would not represent the 'typical' experience with the program and thus be misleading.

Annex D gives detailed guidance on how to follow-through such 'random' selection in a systematic way. This may initially sound complicated, but if a full list of beneficiaries is available, it should take no longer than 15 minutes to work out. And you end up feeling confident that the children that will be participating in the evaluation are genuinely representative of all those who have received the intervention.

Cluster Sampling. Identifying lists of people becomes unrealistic when you are talking about populations of beneficiaries that run into thousands and are spread over a wide geographical area. In this situation, some sort of 'cluster sampling' can be helpful. This means selecting at random a number of 'clusters' where the evaluation can focus. These may, for example, be IDP camps (from a list of all camps) or schools (from a listing of all schools in a District). In each of these clusters a number of participants is (preferably randomly) selected. There are a number of useful guides to such procedures.⁴²

Quota Sampling. These two methods are recommended as the best way of getting 'typical' experiences of beneficiaries. If it is impossible for practical reasons to do either, then 'quota sampling' is a substitute. This method is also appropriate if there are sub-groups that you are particularly interested in

In an evaluation of a classroom-based intervention in schools in Gulu District of N Uganda, eight of nineteen participating schools were chosen at random (by rolling a dice) to be visited. Within the schools classes were selected in a similar fashion. In this way a manageable sample was obtained that was representative of all those children and schools that had participated in the program.

The three approaches to sampling with quantitative evaluation methods are:

Systematic random sampling: getting a list of everyone, and choosing at random a way of selecting from the list

Cluster sampling: selecting sites or 'clusters' from a list at random, and then within each cluster, sampling – at random – the required number of people

Quota sampling: deciding on 'quotas' of participants by criteria such as gender, age, economic status etc.

collecting information from (such as children with disabilities, or children from certain minority ethnic groups). Here, you select your sample to fit in with a 'quota' of beneficiaries defined by certain criteria. For example, in a sample of 80 children in a school, you could decide to sample 20 young girls, 20 older girls, 20 young boys and 20 older boys. Although less than 5% of children may come from a particular ethnic minority you may want to set a 'quota' of 20 children from this group, as with random or cluster sampling you are likely to recruit only 4 or 5 such children in a sample of 80-100.

Sample Sizes and Sample Selection for Qualitative Methods

When using qualitative methods we are faced by the same issues of wanting to access information from a suitable range of beneficiaries and be confident that we have a valid picture of changes brought about by the intervention. However, for this the numerical methods of deciding on sample size used are not appropriate.

The key idea that guides the number of interviews, focus groups or other methods we use is **saturation**. Saturation refers to collecting data until further data collection adds little to the 'picture' that has already been established. It is obviously hard to know in advance how quickly this will happen. However, particularly if a number of different methods of data collection are being used (see the discussion on triangulation in the earlier section), **twenty interviews** or exercises (whether involving individuals or groups) with any particular methods will usually be sufficient to produce saturation.

The issue of how to select participants remains as crucial for qualitative methods of collecting data as it does for quantitative methods. Too often evaluations are weakened by a failure to ensure that participants represent the range of beneficiaries that have received an intervention. If, for example, focus groups are run only with youth in the most accessible refugee camp that has received services, it is likely that their

'Saturation refers to collecting data until further data collection adds little to the 'picture' that has already been established'

In Northern Uganda CCF arranged focus groups to discuss issues around sexual violence against women and girls. Although this was a qualitative method, the groups were convened at locations within a refugee camp that had been selected at random from a listing of all 'blocks', and used a 'sampling interval' in a similar fashion to that used with quantitative household surveys

experience will not be representative of others staying in more inaccessible camps.

Three methods are of value here. If good records are available then again beneficiaries can be selected at **random** to participate in qualitative discussions. More typically participants may be selected to fulfill a **quota** of beneficiaries defined by a particular set of characteristics (e.g. boys, in school, aged under 11; or girls, not in school, aged over 11). Or '**snowballing**' can be used, using initial participants/informants to identify others that fulfill certain criteria for inclusion.

10. Implementing an Evaluation: Case Examples

Here are two case studies that illustrate the principles outlined in earlier sections of this manual. The first example is an evaluation of programs promoting the psychosocial well-being of children in Palestine.⁴³ The second example is an evaluation of a program supporting the reintegration of children formally associated with fighting forces in Mozambique⁴⁴.

Evaluation of Psychosocial Programs Supporting Children in the West Bank and Gaza

This example considers an evaluation of programs of psychosocial support by Catholic Relief Services (CRS) and Save the Children³⁹. These programs were developed to address the needs of children in the West Bank and Gaza related to the political conflict across the Occupied Palestinian Territories. An external evaluation team was called in at the time of program design, before implementation, which helped establish a coherent process of developing objectives and indicators for the programs. Discussions (between the evaluation team, the program team and the beneficiary groups with which they had contact) resulted in defining the objectives and indicators listed overpage.

The objectives and indicators chosen reflected a number of issues. Recent work had suggested that children were under considerable stress and considered that relationships with their parents were increasingly strained by the prevailing situation. The principal means to address such needs was to be the creation of facilities and skills within communities to facilitate play and cultural activities. It was hoped that 'safe' activities for children would influence children's well-being and provide the opportunity for stronger relationships with parents to be established. The program did not so much focus on delivering play and

recreational activities, as enabling partner community organizations to deliver these.

Objectives	Objectives	Indicators	Source of Information
Impacts	Enhanced children's well-being Stronger parent-child relationships Enhanced capacity of local community organizations for work with children	<ul style="list-style-type: none"> Improved emotional wellbeing Improved social wellbeing Children report closer relationship with parents Organizations develop wider range of activities utilizing skills acquired 	<ul style="list-style-type: none"> Child Behaviour Checklist (Palestinian Version) Parental Checklist Interviews with organizations Organizational records
Outcomes	Children engaged in play in a safe environment Children participate actively in cultural activities	<ul style="list-style-type: none"> High proportion of local children regularly attended safe play area Children active in play Parents report positive engagement by children Children report active engagement with activities 	<ul style="list-style-type: none"> Program records Site visits Site visits Focus groups with parents Participative activities with children
Outputs	Safe play areas established and equipped Training conducted in delivery of cultural activities	<ul style="list-style-type: none"> Number of safe play areas built or refurbished Equipment delivered to play areas Number of trainings offered in facilitating cultural activities Number of persons trained 	<ul style="list-style-type: none"> Program records Site inspections Program records

Consequently, the evaluation was able to focus both on issues of program outcome and of wider impact. The outcome indicators chosen reflected the objectives to get children engaged in safe play activities, and in a wide range of cultural activities. This phase of the evaluation used information from qualitative interviews with children and parents, site visits and analysis of program records. Focus groups and participatory exercises such as those discussed in Section 9 were used.

On completion of the planned program, the evaluation team aimed to see if wider impacts on children's well-being and relationships with parents could be identified. Here the sources of information were mostly quantitative, predominantly structured questionnaires completed by children and parents. The team took advantage of the fact that the Child Behavior Checklist, which measures aspects of emotional and social well-being, had been validated for use in Palestine. A questionnaire that had been developed and used by staff at the University of Bethlehem provided a measure of parental support. If neither of these had been available, the evaluation would have had to consider developing a questionnaire based on issues that parents and children came up with in a form of 'free-listing' exercise.

A key feature of the design of the evaluation was its use of **baseline** measures and **comparison groups**. As a baseline, information on children's well-being (and relationship with parents) was collected from 300 children (150 from the West Bank and 150 from Gaza) before they began to engage with the programs. This enabled the evaluation team to measure differences in scores in these same children a year or so later when the project had delivered its intervention. Scores suggested that there was an overall improvement in well-being and relationships for many of the children. But could this change simply be that conditions in Palestine had improved a little over that time? A comparison group of 100 children (50 from another community in Gaza and 50 from another community in the West Bank, all of whom had not had access to the intervention) enabled the team to examine this question. In fact there had been improvements in these comparison communities also, but generally much less than in the communities where the program had been delivered. This represents good evidence that the programs had real impact on the lives of children and their parents.

Evaluation of Psychosocial Support to Former Child Soldiers in Mozambique

In the late 1980s Save the Children initiated one of the first internationally subsidized rehabilitation and support programs for child soldiers.⁴⁵ The opportunity arose some sixteen years later to follow-up those who had participated in programming. This is an unusual example, being so long after the end of the project, but it usefully illustrates a number of issues about impact evaluation.

Monitoring and evaluation during the lifetime of the project suggested that a number of outputs and outcomes had been met. However, did the project have – as was its goal - real benefit on the lives of these children when they returned to their communities? Those children that had been involved in the program were now – some sixteen years later - mature adults. No impact objectives and indicators had been formally defined at the time of project implementation, so these had to be devised by the evaluation team. They did so by talking with local communities about the 'signs' that someone had re-adjusted to life well after the

war. Using the structure of the three core psychosocial domains, they came up with this list:

Sixteen Year Post Program Evaluation (2003-2004)

Domain: Skills and Knowledge

Indicators: ability to be economic providers for household, as measured by: standard of housing; farming activity; off-farm income

Domain: Emotional Well-being

Indicators: presence of fears, nightmares and other symptoms related to war experiences; self-esteem, sense of community acceptance and belonging

Domain: Social Well-being

Indicators: ratings of social functioning by spouses, parents and neighbors

By interviewing beneficiaries of the project, their families and members of their local communities, the impact evaluation indicated that those that had received support as youths were generally functioning very well as adults.⁴⁶ They had incomes above the average, assumed significant responsibility within their communities and were well respected within them. Many reported emotional symptoms related to their experiences during the war (e.g. fears, bad dreams etc.), but they generally coped with them well.

In terms of methods, there are two major observations from this evaluation. Firstly, there were initially plans to identify a comparison group of adults who, while having been involved in the civil war as child combatants, had not been beneficiaries of the psychosocial support program. In practice, however, it proved very difficult to sensitively and ethically identify those who had fought in the war as children, if they were not already known to the evaluation team through their being project beneficiaries. The evaluation team included members who had been involved in the original program, and who therefore were trusted by beneficiaries. In this situation, as discussed earlier, 'local norms' were used as a basis for comparison. This is obviously a weaker design than a comparison group. The evaluation team could thus say 'those who participated in the program are doing better than is typical in this area'; but they could not say 'those who participated in the program are doing better than those who did not participate'.

Secondly, in terms of deciding how 'doing better' is defined, 'free listing' proved to be one of the more important methodological tools. The most often cited characteristic of a "good and successful adult" in rural southern Mozambique was someone who consistently helped neighbors in need. It was placed far above

individual achievements, for example. The other most commonly cited characteristics included the ability to economically support a family household, to be a good spouse and to be a good parent.

11. Writing Up and Sharing Findings

This section provides a brief guide to writing up evaluation findings in a report, and developing a presentation for communication of findings.

Report writing

Comprehensive guidance on standards for UNCEF evaluation reports is already available to UNICEF staff.⁴⁰

Each evaluation report should include

- a three page Executive Summary;
- the full report.

A 1 page summary suitable for non-professionals is also a useful tool for dissemination.

This means that findings are accessible and can be used for different purposes and for a variety of audiences. For example, the one page summary covers key findings and is a 'quick read' version of the whole report.

We also want to encourage the building up of a methods database and so we are looking for full descriptions of methodology including tools used in evaluation reports. This will also make peer review more possible.

'Evaluation reports should be prepared.. that are accessible and can be used for different purposes and for a variety of audiences'

Sharing findings

Guidance is provided here on effective ways of sharing your findings. The IASC guidelines lists 'collate and disseminate assessment results' as one of four key actions in the assessment, monitoring and evaluation of mental health and psychosocial issues⁴⁷. They suggest that key findings are distributed to relevant stakeholders – government,

coordination bodies, the affected population, and usually funders – to ‘facilitate reflection, learning and change.’

It is usually the case that findings are shared and discussed with donors; but it is important to remember UNICEF’s accountability to other groups too, especially to affected populations.

There are some basic questions to think through before you do anything:

- What are the objectives of dissemination?
- Who are the target audiences?
- What are the priorities?
- What resources are available in terms of time, skills and experience, existing contacts and relationships?

When you look at your findings, you will need to consider⁴¹:

- the source
make sure the source of the work is firmly established and credibility and quality is ensured
- the message
make sure that findings are presented consistently; that they are clear and can be understood; that they are relevant to the audience
- the medium
make sure that findings are presented clearly and attractively; that they are user-friendly with ‘quick read’ options; that they can physically reach the intended audience; that they are circulated promptly on completion
- the user
target audiences may have different needs and these needs will influence how findings are ‘packaged.’ The format and level of information needed and the way findings are presented will need careful attention. Thinking about the perceived relevance of findings to the current needs of the audience is also crucial.

The reports of the program Survey of War-Affected Youth (<http://www.sway-uganda.org/>) provide a good example of well-presented, easily accessible material. Note how the implications of the information presented for policy and practice are clearly highlighted.

‘the passive distribution of written materials ...is not likely to have much impact’

Remember that the passive distribution of written materials or presentations about findings with no opportunity for discussion is not likely to have much impact.

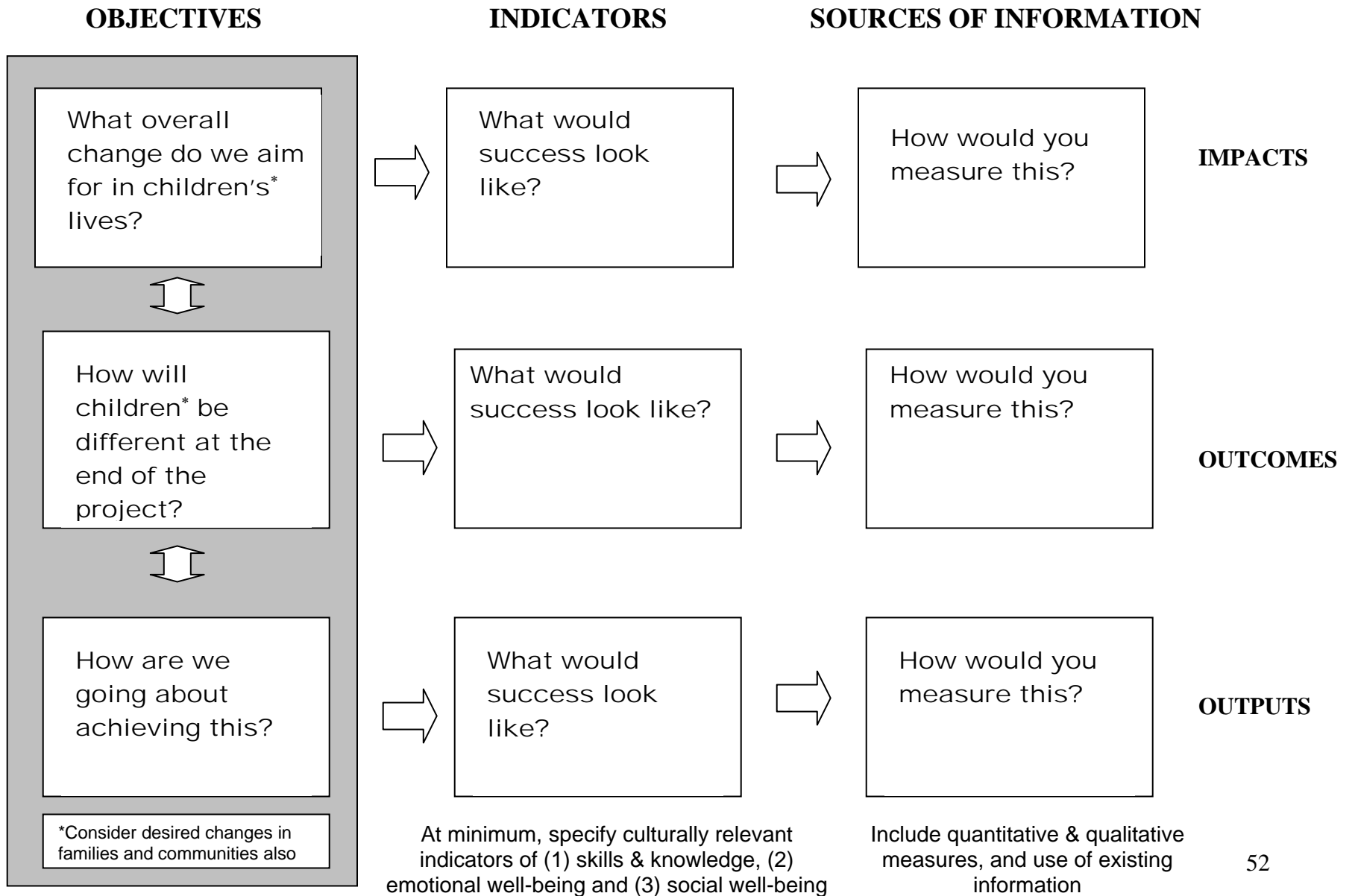
Using workshops to discuss findings or arranging a launch party or showing a DVD featuring main findings is more likely to engage attention.

Developing interaction and ongoing links with practitioners, NGOs and policy makers through informal and formal partnerships will also provide opportunities to share work that you have done.

Make opportunities to further develop skills among program staff and policymakers in using findings from evaluation in order to build up the evidence base concerning psychosocial programming

Crucially, reviewing findings should be linked with planning a response. One way to do this is to develop an action plan with UNICEF management that outlines steps to be taken in response to the recommendations of an evaluation and then monitors implementation of that action plan.⁴⁸

ANNEX A: A GUIDE TO THE DEVELOPMENT OF INDICATORS



ANNEX B: A STEP-BY-STEP GUIDE TO CONDUCTING AN EVALUATION

Step*	Task	Section reference	Comments
1	Engage Beneficiaries and Staff in Evaluation Planning	2, 3, 4, 5, 6, 8 & 10	Review scope of evaluation and resources available/ required for its completion.
2	Confirm/Develop Program Objectives and Indicators	3, 5, 6, 8 & 10	Best if original program plan develops these in collaboration with affected population. If not, involve key stakeholders in developing key questions for evaluation. Develop 'core indicators'.
3	Identify Basis for Comparison	7 & 10	Identify comparison groups (through phased roll-out or other mechanism).
4	Identify Sources of Existing Information	6, 7	Identify pre-existing baseline information – and other sources of routinely available information - relevant to chosen indicators
5	Choose Methods for Collecting Information	6, 9 & 10	Use indicators to identify required methods of quantitative and qualitative data collection.
6	Plan Baseline Assessment (or 'Reconstruct' Baseline, if required)	7, 9 & 10	Projects should prioritize good baseline assessment as a basis for programming. However, where there is no baseline and programming has begun, use suggested guidance.
7	Train Staff As Necessary for Collecting Information	8 & 9	Equip staff for range of methods, and broader issues around ethics and sensitivities of collecting information.
8	Access Communities	8	Arrange with sensitivity and participation of all relevant authorities (drawing on earlier engagement). Secure appropriate consents.
9	Collect Information	9 & 10	Arrange focused, supervised program of collecting information from beneficiary and comparison settings;
10	Analyze and Share Findings	11	Ensure prompt analysis of information, so that findings can be shared and applied. Ensure multiple versions of findings are prepared for range of stakeholders.
*Steps are indicated in sequence, but earlier steps will often need to be 'revisited' subsequent to later decisions.			

ANNEX C: A PROGRAM OFFICER'S RESPONSIBILITIES FOR EVALUATION OF PSYCHOSOCIAL PROGRAMS

Key responsibilities within the yearly cycle of programme planning:

- Organizations should identify – at least once every two years – opportunities for impact evaluation of completed programs
- Identify staff training needs in support of effective monitoring and evaluation of psychosocial programs

Key responsibilities at the time of program planning (prior to implementation):

- Ensure development of clear statement of program objectives with relevant participation from children and communities (including identification of appropriate indicators at output, outcome and impact levels)
- Make appropriate budget provision – in negotiation with funders as required – for outcome evaluation
- Plan 'roll out' of intervention in a way that explicitly identifies comparison groups for evaluation purposes

Key responsibilities in support of evaluation activity:

- Develop a clear plan for the evaluation (following a structure like that provided in Annex B)
- Identify requirements for contracting external assistance to support the evaluation
- Continue – through times of times of pressure due to time, limited resources, competing priorities and skepticism – to serve as a 'champion' for evaluation as a key ingredient supporting quality programming of confirmed value to children

ANNEX D: GUIDANCE ON SAMPLE SELECTION

A Procedure for Ensuring Unbiased Selection of A Sample

1. Divide the total number of beneficiaries in the listed population by the required sample size to produce a number called the 'sampling interval' (e.g. with a population of 800 children receiving a school-based intervention, divide by 100 to get a sampling interval of 8).
2. Choose a random starting point between 1 and the sampling interval (e.g. by writing the numbers 1 through 8 on pieces of paper and picking out one piece of paper by chance, say, the number 5)
3. The first person in the sample is then defined by this number (e.g. in this example, the fifth child on the list).
4. Subsequent members of the sample are selected by picking those that fall according to the sampling interval (e.g. in this example, we select the 13th (5+8), 21st (5+8+8), 29th (5+8+8+8) child, all the way up to the one hundredth member of the sample (who turns out to be the 797th on the list).
5. Make clear rules for how you deal with those who are due to be in your sample, but are not around when you come to meet them. You can afford to lose one or two people from your sample, but if there are many absentees you will need to make 'substitutions' (e.g. in the above example, if the child 21st on the list is not available for interview on more than two occasions when you visit the school, you replace them with the next child on your list i.e. the 22nd child listed)

ANNEX E: USEFUL RESOURCES

Action Aid International. 'Participatory Vulnerability Analysis: A step-by-step guide for field staff'.

<http://www.actionaid.org.uk/wps/content/documents/PVA%20final.pdf>

Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP). Chapter 6 'Monitoring'; Chapter 7 'Evaluation'. In *Participation by Crisis-Affected Populations in Humanitarian Action: A Handbook for Practitioners*, pp.193-227.

<http://www.globalstudyparticipation.org/index.htm>

Bamberger M, Rugh J and Mabry L (2006). 'Real World Evaluation.' Sage: California.

Bolton P. and Tang A.M. (2002). 'An alternative approach to cross-cultural function assessment'. *Social Psychiatry and Psychiatric Epidemiology*. 37: 537-543.

http://www.who.int/mental_health/emergencies/mh_key_res/en/index.html

Bragin M. (2005). 'The community participatory evaluation tool for psychosocial programmes: A guide to implementation'. *Intervention: International Journal of Mental Health, Psychosocial Work and Counseling in Areas of Armed Conflict* 3, 3-24.

http://www.interventionjournal.com/downloads/31pdf/03_24%20Bragin%20.pdf

Horizons Program/Population Council and Family Health (2005) 'Ethical Approaches to gathering information from children and adolescents in international settings – guidelines and resources.'

<http://www.popcouncil.org/horizons/childrenethics.html>

IASC (2007) 'IASC Guidelines on Mental Health and Psychological Support in Emergency Settings.'

<http://www.humanitarianinfo.org/iasc/content/products/docs/IASC%20MHPSS%200guidelines%20Feb%2025%202007.pdf>

International Institute for Environment and Development (IIED). 'Participatory Learning and Action (PLA Notes)'. (see specifically notes 31 and 42, which cover PRA monitoring and evaluation)

http://www.iied.org/NR/agbioliv/pla_notes/backissues.html

Oxfam GB (2007) 'Impact Measurement and Accountability in Emergencies: The Good Enough Guide'. <http://www.oxfam.org.uk/publications>

Psychosocial Working Group (2005) 'Psychosocial Intervention in Complex Emergencies: A Framework for Practice.'
<http://www.forcedmigration.org/psychosocial/papers/A%20Framework%20for%20Practice.pdf>

Perez-Sales P. (2006). 'Repensar Experiencias. Evaluación de programas psicosociales y de salud mental. Metodologías y técnicas'. *Ed Popular*. (www.psicosocial.net) ('Rethinking experiences. Assessment of mental health and psychosocial programmes. Methods and techniques', in Spanish)

Refugee Studies Centre Working Paper 30 (2004). 'Piloting Methods for the Evaluation of Psychosocial Programme Impact in Eastern Sri Lanka.'
<http://www.rsc.ox.ac.uk/PDFs/rrpilotingmethods04.pdf>

Sphere Project (2004). 'Humanitarian Charter and Minimum Standards in Disaster Response'. <http://www.sphereproject.org/handbook>

UNICEF East Asia and Pacific Office and Regional Emergency Psychosocial Support Network (2005). 'Handbook on Psychosocial Assessment Of Children And Communities In Emergencies.'
<http://www.crin.org/docs/Handbook%20new%20update.pdf>

UNICEF (2005). 'Understanding Results-based Programme Planning and Management.'

UNICEF (2006). 'Child Protection Information Sheets.'
[http://www.unicef.org/protection/files/Child_Protection_Information_Sheets_\(Booklet\).pdf](http://www.unicef.org/protection/files/Child_Protection_Information_Sheets_(Booklet).pdf)

ANNEX F: SUMMARY OF MAJOR RECOMMENDATIONS

1. The IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings (2007) represent the consensus of the international humanitarian community on appropriate psychosocial support and mental health programming in emergencies. These guidelines should be used as the core policy framework for developing psychosocial projects in emergency settings.
2. The following three domains are the most helpful to evaluate how well UNICEF's work affects the lives and experiences of children:
 - **Skills and knowledge**
e.g. life skills, using culturally appropriate coping mechanisms, vocational skills, conflict management etc.
 - **Emotional well-being**
e.g. feeling safe, trust in others, self-worth, hopeful for the future etc.
 - **Social well-being**
e.g. attachment with caregivers, relationships with peers, sense of belonging to a community, access to assume socially appropriate roles, etc. resuming cultural activities and traditions
3. This framework can be used to develop both objectives for psychosocial programming, and indicators that signal whether these objectives have been met.
4. All projects should ensure that evaluations consider the outputs and outcomes of psychosocial programming. To build the evidence-base of outcomes of psychosocial projects ultimately leading to changes in children's well-being and circumstances, impact evaluations also need to be conducted.
5. Impact evaluations require considerably more resources than outcome evaluations. Resources such as funding, staffing, time, technical assistance need to be identified for impact evaluations and it is vital that this planning begins at an early stage. This means identifying the necessary resources for their completion within the budgeting and planning cycle.
6. In the course of evaluating psychosocial programs, it is crucial that key psychosocial principles (and related ethical principles) are observed at every stage of the process.
7. All evaluations should aim to feature both baseline measures and comparison groups.

8. All evaluations should have as a minimum the following features of children's and community participation:

- determining objectives and indicators for the project
- determining local definitions of wellbeing
- providing their views on the project
- providing feedback on draft findings and the implications for their community and/or future similar projects

9. Evaluations are strengthened by the varied perspectives gained from using both 'quantitative' and 'qualitative' methods. These methods bring different insight, but can also inform each other. Quantitative data can inform the selection of participants, and the focus of enquiry, for qualitative work. Qualitative work can help develop locally relevant quantitative measures.

10. Rigorous evaluations can generally be conducted with a sample of between 80 and 100 participants (with a comparison group of a similar number).

11. Careful attention should be given to the selection of samples for evaluations. Random, cluster, quota and snowballing approaches may each be appropriate, depending upon circumstance.

12. Evaluation reports should be accessible and suitable for a variety of audiences. We encourage the building up of a methods database, detailing descriptions of methodology, including tools used.

13. Reviewing findings should be linked with planning a response. One way to do this is to develop an action plan with UNICEF management that outlines steps to be taken in response to the recommendations of an evaluation and then monitors implementation of that action plan.

END NOTES

¹ See 'Project 1 Billion Book of Best Practices: Trauma and the Role of Mental Health in Post-Conflict Recovery' (available <http://www.hprrt-cambridge.org>) for guidance in this area.

² Understanding Results-based Programme Planning and Management: Tools to Reinforce Good Programming Practice. UNICEF, May 2005.

³ Bamberger, Michael et al. Real World Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage, Thousand Islands, California, 2006. Section 7 of this guide also looks at these issues in more detail.

⁴ Impact evaluation consultants require expertise in research design and quantitative and qualitative methods, as well as experience in crisis affected situations.

⁵ UNICEF Child Protection in Emergencies Training and Resource CD.

⁶ Inter-Agency Standing Committee (IASC) Guidelines on Mental Health and Psychosocial Support in Emergency Settings 2007

⁷ Psychosocial Working Group Working Paper: A Framework for Practice. Available: www.forcedmigration.org/psychsocialA%Framework%20for%20Practice.pdf

⁸ Psychosocial Working Group Working Paper: A Framework for Practice. Available: www.forcedmigration.org/psychsocialA%Framework%20for%20Practice.pdf

⁹ Inter-Agency Standing Committee (IASC) Guidelines on Mental Health and Psychosocial Support in Emergency Settings 2007

¹⁰ UNICEF Child Protection in Emergencies Training and Resource CD

¹¹ Convention on the Rights of the Child 1989. Office of the United Nations High Commissioner for Human Rights, Geneva, Switzerland.

¹² Inter-Agency Standing Committee (IASC) Guidelines on Mental Health and Psychosocial Support in Emergency Settings 2007 pp 41-43

¹³ UNICEF Child Protection in Emergencies CD ROM pp 21- 22. UNICEF: New York.

¹⁴ Horizons Program/Population Council and Family Health (2005) 'Ethical approaches to gathering information from children and adolescents in international settings: guidelines & resources.' <http://www.popcouncil.org/horizons/childrenethics.html>

¹⁵ These criteria are based upon standard OECD/DAC criteria for development assistance. Further details are given in Monitoring & Evaluation Quick Reference: Extracts from the Programme Policy and Procedure Manual Revised, May 2005. UNICEF. Available at: www.unicef.org/

¹⁶ UNICEF 2006. Child Protection Information Sheets.

¹⁷ Monitoring & Evaluation Quick Reference: Extracts from the Programme Policy and Procedure Manual Revised, May 2005. UNICEF. Available at: www.unicef.org/

¹⁸ Understanding Results-based Programme Planning and Management, UNICEF 2005

¹⁹ Assessment and Evaluation of Psychosocial Programming for Crisis-affected Children, Prepared by Program on Forced Migration and Health, Mailman School of Public Health, Columbia University USA, UNICEF 2006

²⁰ Children in Crisis: Good Practices in Evaluating Psychosocial Programming SCF, 2004

²¹ Children in Crisis: Good Practices in Evaluating Psychosocial Programming SCF, 2004; Bamberger, Michael et al. Real World Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage, Thousand Islands, California, 2006

-
- ²² Assessment and Evaluation of Psychosocial Programming for Crisis-affected Children, Prepared by Program on Forced Migration and Health, Mailman School of Public Health, Columbia University USA, UNICEF 2006
- ²³ Bamberger, Michael et al. Real World Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage, Thousand Islands, California, 2006. pp 92-93
- ²⁴ Technical Notes: Special Considerations for Programming in Unstable Situations, UNICEF 2003, p 412
- ²⁵ Hart et al 2004 p.22 Piloting Methods for the Evaluation of Psychosocial Programme in Eastern Sri Lanka. Refugee Studies Centre, University of Oxford.
- ²⁶ Arnstein, S. (1969) 'A Ladder of Citizen Participation,' Journal of the American Planning Association, Vol. 35, No. 4, July 1969, pp. 216-224
- ²⁷ Holte-McKenzie, M., Forde, S. and Theobald, S. Development of a participatory monitoring and evaluation strategy Evaluation and Program Planning 29 (2006) 365-376
- ²⁸ Available at: www.popcouncil.org/horizons/childrenethics.html
- ²⁹ After De Berry et al 2003. The Children of Kabul: Discussions with Afghan Families. Save the Children USA and UNICEF p 3-4 and Hart et al, 2004 p17
- ³⁰ This final bullet point addresses a complex area. As indicated in section 3, 'exploring sensitive issues with children requires skills, local knowledge, and experience. Any such work should only be carried by trained and experienced staff who can ensure appropriate support and follow-up, and work within agreed standards.'
- ³¹ Bamberger, Michael et al. (2006) Real World Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage, Thousand Islands, California.
- ³² Impact Measurement and Accountability in Emergencies: The Good Enough Guide Oxford: Oxfam 2007
- ³³ From Hart et al. 2004 p.50.
- ³⁴ Assessing Afghan Children's Psychosocial Well-being: A Multi-Modal Study of Intervention Outcomes. CCF, Oxford University and Queen Margaret University College, Edinburgh 2005.
- ³⁵ Available: www.rsc.ox.ac.uk/PDFs/rrpilotingmethods04.pdf
- ³⁶ see above.
- ³⁷ see above.
- ³⁸ see above.
- ³⁹ Achenbach, T. M. 2001. Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, Vermont.
- ⁴⁰ Bamberger, Michael et al. (2006) Real World Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage, Thousand Islands, California.
- ⁴¹ A confidence level of 95% (i.e. accepting that there is a 5% chance that there is a change created by the intervention that we won't detect, or vice-versa) and an effect size of between 0.20 and 0.30 (representative of the sorts of change that effective psychological and social interventions can have).
- ⁴² The Food and Nutrition Technical Assistance (FANTA) Sampling Guide 1997. Available at: www.fantaproject.org/publications/sampling/shtml
- ⁴³ Evaluation of the psychosocial intervention for Palestinian Children undertaken by the CRS and the Save the Children Federation by Loughry, M & Ager, A on behalf of the Gates Foundation. Refugee Studies Centre, University of Oxford.

⁴⁴ Boothby, N, Crawford, J, and Halprin, J. The Life Outcomes of Former Mozambican Child Soldiers, Global Public Health An International Journal for Research, Policy and Practice, Routledge Publishing, Issue One, January 2006, 87-104.

³⁹ Evaluation of the psychosocial intervention for Palestinian Children undertaken by the CRS and the Save the children Federation by Loughry, M & Ager, A on behalf of the Gates Foundation. Refugee Studies Centre, University of Oxford.

⁴⁵ Boothby, N, Crawford, J, and Halprin, J. The Life Outcomes of Former Mozambican Child Soldiers, Global Public Health An International Journal for Research, Policy and Practice, Routledge Publishing, Issue One, January 2006, 87-104.

⁴⁶ See above.

⁴⁰ Monitoring and Evaluation quick Reference: Extracts from the Programme Policy and Procedure Manual, revised may 2005. UNICEF Evaluation Report Standards. Available at: www.unicef.org

⁴⁷ Inter-Agency Standing Committee (IASC) Guidelines on Mental Health and Psychosocial Support in Emergency Settings 2007 p 43 and p 48

⁴¹ Drawn from notes from Research Unit for Research Utilization, St Andrews and University of Edinburgh 2005.

⁴⁸ Bamberger, Michael et al. (2006) Real World Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage, Thousand Islands, California. p 402