ACTION CONTRE LA FAIM

CARE AUSTRIA

CHILD FUND INTERNATIONAL

HEALTHNET TPO

INTERNATIONAL FEDERATION OF RED CROSS AND

RED CRESCENT SOCIETIES

INTERNATIONAL MEDICAL CORPS

MEDICOS DO MUNDO

REPSSI

SAVE THE CHILDREN

TERRE DES HOMMES

UNICEF

WAR CHILD HOLLAND

WAR TRAUMA FOUNDATION

WORLD VISION INTERNATIONAL

INTER-AGENCY GUIDE TO THE EVALUATION OF PSYCHOSOCIAL PROGRAMMING IN HUMANITARIAN CRISES

ACKNOWLEDGEMENTS

Writing Team
Alastair Ager, Wendy Ager, Vivi Stavrou and Neil Boothby.
Thanks
Thanks to the many individuals and organizations involved in the development and field-testing of the guide.

UNICEF acknowledges the financial support of ECHO in the piloting of the guide.

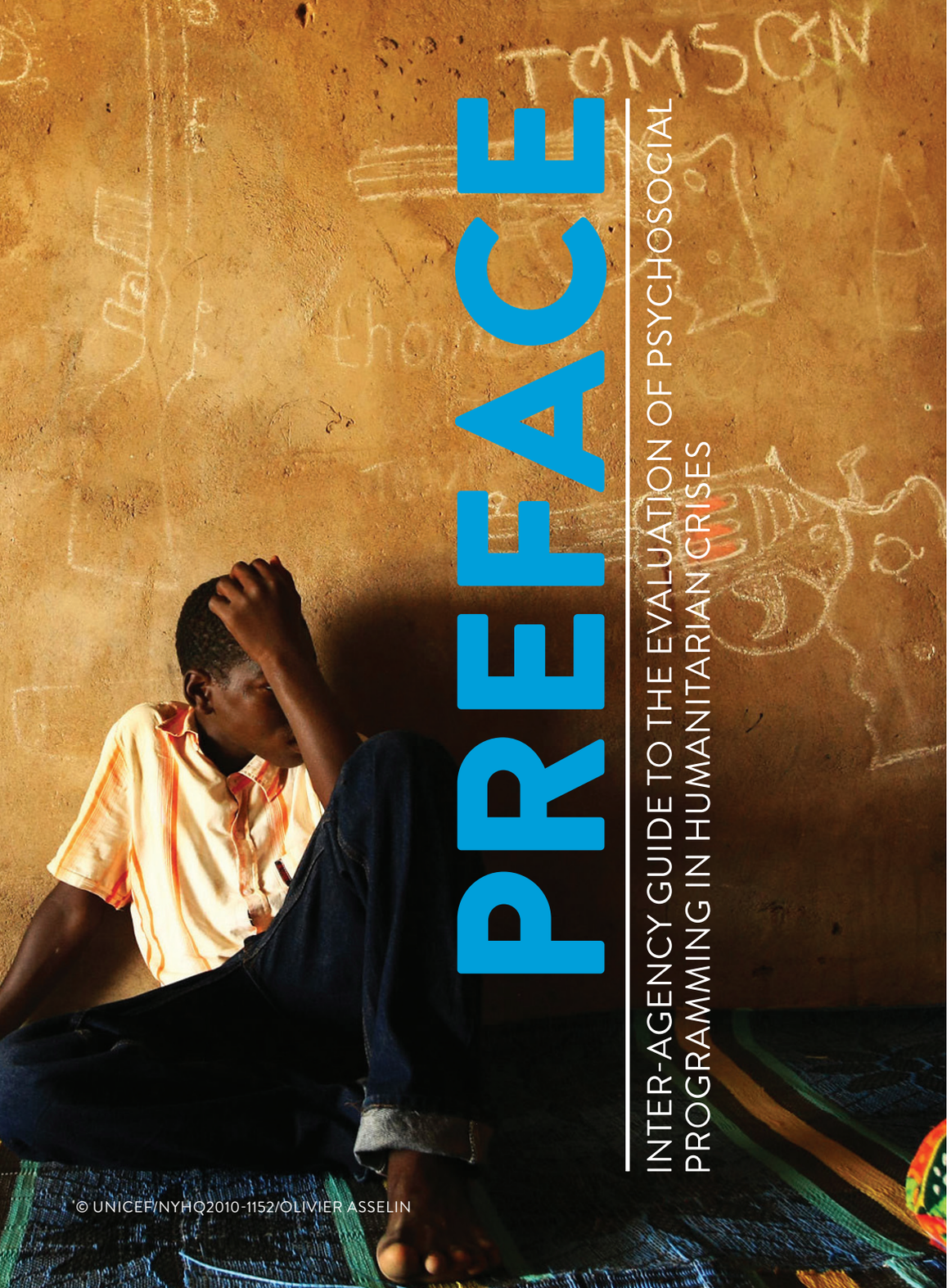
Recommended citation
UNICEF, (2011) Inter-Agency Guide to the Evaluation of Psychosocial Programming in Emergencies. New York: United Nations Childrens Fund.
or
Ager, A, Ager, W, Stavrou, V. & Boothby, N. (2011) Inter-Agency Guide to the Evaluation of Psychosocial Programming in Emergencies. New York: UNICEF.

Cover page photo
© UNICEF/NYHQ2009-1911/Giacomo Pirozzi
Graphic design
www.troispoints.ca

ENDORSEMENT

This guide has been endorsed by the following organizations

Action Contre la Faim
CARE Austria
Child Fund International
HealthNet TPO
International Federation of Red Cross and Red Crescent Societies
International Medical Corps
Medicos do Mundo
REPSSI
Save the Children
Terre des Hommes
UNICEF
War Child Holland
War Trauma Foundation
World Vision International



PREFACE

INTER-AGENCY GUIDE TO THE EVALUATION OF PSYCHOSOCIAL PROGRAMMING IN HUMANITARIAN CRISES

© UNICEF/NYHQ2010-1152/OLIVIER ASSELIN

PREFACE

UNICEF together with partners has developed this **Inter-Agency Guide to the Evaluation of Psychosocial Programming in Humanitarian Crises** to assist organizations working in the field of psychosocial support to think through key issues in planning and implementing an evaluation.

There are major challenges of conducting evaluations in humanitarian crises such as natural disasters and armed conflicts. 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, identify potential unintended negative consequences 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 effective psychosocial practice.

In recent years, psychosocial support has become an increasingly central part of development and humanitarian programming. The Inter-Agency Standing Committee (IASC) Guidelines on Mental Health and Psychosocial Support in Emergency Settings now present broad consensus on effective, ethical programming in the field. However, there is widespread recognition that there remains a need to build a stronger evidence for such work. More extensive and robust evaluations are required to develop a better understanding of what approaches to psychosocial support are the most effective, and in what situations.

A desk review[1] of existing psychosocial assessments and evaluations formed the foundation for design of this guide. This review indicated that, although some effective psychosocial evaluations have been conducted, recurrent problems frequently leading to questionable or inconclusive results include:

- a lack of clear and appropriate project objectives
- weaknesses in the design and methodology of evaluations, including a failure to collect baseline information, and
- a lack of appropriate quantitative tools for assessing psychosocial well-being.

The aim of this guide is thus to produce understandable, accessible guidance and tools for the field on how to conduct psychosocial evaluations. This effort is part of a wider goal to build capacity for evaluation for the sector through training and technical support. While the guide focuses primarily on children and emergencies, experience over the last two years of field testing has shown that the guidance can be useful with other crisis-affected populations and in other settings, such as those affected by HIV and AIDS.

This finalized version of the guide has been developed through wide consultation with organizations and experts working on psychosocial programming and evaluation since 2007. Strengthening the evidence-base for psychosocial intervention can only be achieved through interagency efforts, and I am thus particularly pleased to note the wide inter-agency endorsement of this guidance. I welcome the opportunity that this provides for improving the quality of our evaluations and, through that, the quality of our support for those affected by humanitarian crises.

Amanda Melville
October 2010

© UNICEF/NYHQ2009-0135/GIACOMO PIROZZI

In a collaborative approach to psychosocial programming, a number of Palestinian agencies agreed to specific indicators of aspects of psychosocial well-being, for example, reduction in troubling dreams (as one measure of emotional well-being) and increasing collaborative behavior with teachers and peers (as one measure of social well-being). Identifying such indicators has enabled psychosocial workers to gather clear results on their interventions. Additionally it has led to teachers increasing their awareness and focus on these aspects of behavior, and to an increase in parental involvement in children's activities at school.[2]

CONTENTS

INTER-AGENCY GUIDE TO THE EVALUATION OF PSYCHOSOCIAL PROGRAMMING IN HUMANITARIAN CRISES

CONTENTS

1 THE AIM OF THIS GUIDE

1 THE AIM OF THIS GUIDE

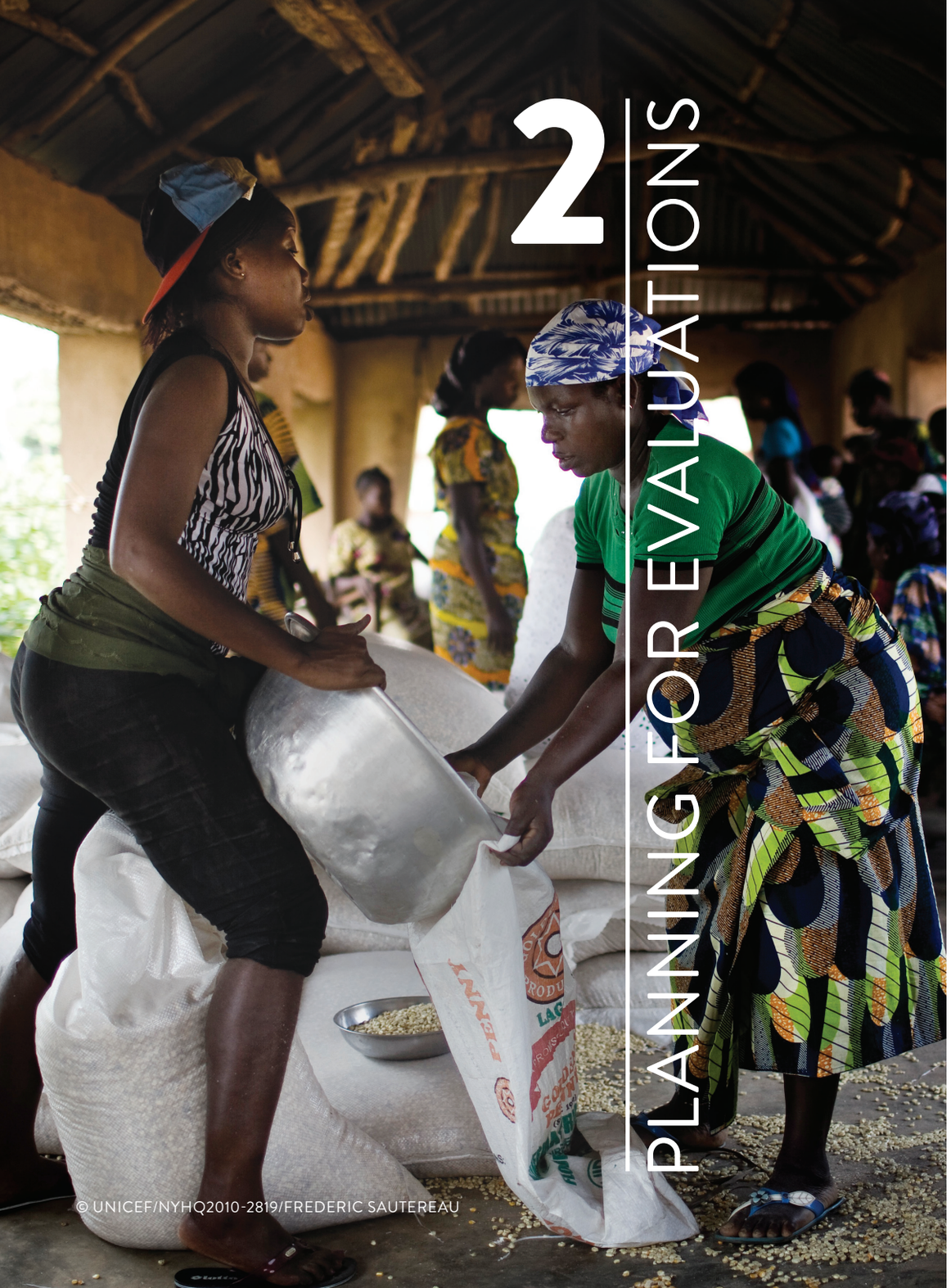
Exposure to the disruption, loss, and violence associated with humanitarian crises places significant psychological and social strain on individuals, families and communities. The way in which people experience and respond to conflicts and disasters varies greatly, yet with the right support the majority will be able to overcome these difficult experiences.

'Evaluation helps us to assess the activities we are involved with and to learn how we can improve our work.'

This guide provides practical guidance for the evaluation of psychosocial programs in crisis settings. In running a psychosocial program, it is important to see if we are actually making a difference to the individuals, families and communities we are working with. Evaluation helps us to assess the activities we are involved with and to learn how we can improve our work.

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.

The guide focuses on **psychosocial programs**[3], and not those that deal with mental disorders, as the tools and methodologies for assessing mental disorders may be very different from those measuring psychosocial distress and well-being. The emphasis is on psychosocial programming targeting the needs of children and their families but, with appropriate adaptation, the principles and approaches proposed should prove of relevance to those working with other populations.



2 PLANNING FOR EVALUATIONS

© UNICEF/NYHQ2010-2819/FREDERIC SAUTEREAU

Evaluations set out to understand whether or not programs have achieved their goals and what has been learned in the process. They should:

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

THE PLACE OF MONITORING & EVALUATION IN PROGRAM DESIGN AND MANAGEMENT

Evaluation is a central feature of project design and 'project cycle management'.[4] An evaluation strategy should be considered right from the start of the project cycle and integrated into each stage of the cycle.

Many organizations now seek to ensure that monitoring and evaluation activity is fully integrated with processes of designing and managing interventions. Figure 1 shows how World Vision International approaches this with their LEAP (Learning through Evaluation with Accountability & Planning) system.[5]

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.

Evaluation can also be understood as a process that continues throughout the life of a project: formative evaluations clarifying future strategy in the early stage of work; mid-term evaluations checking out progress and providing a basis for adjustment mid-way through a project; and end-of-project evaluations summarizing outcomes and impacts.

This guide encourages agencies to view evaluation as an integral part of the project cycle, but acknowledges that evaluation strategies are not always planned in advance, as time and resources are often limited, especially in emergencies. But even in pressured circumstances, evaluation can provide the opportunity to capture valuable lessons.[6]

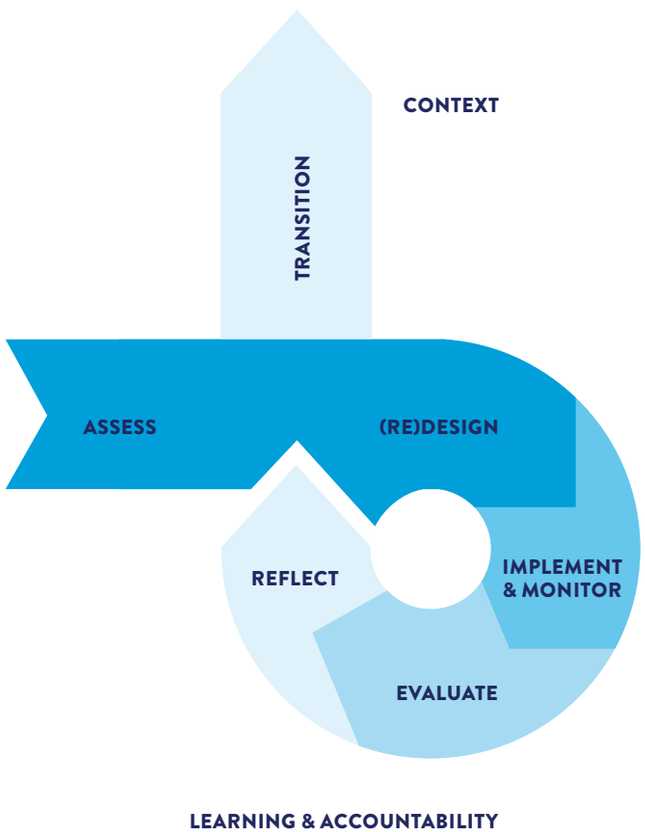


FIGURE 1
LEAP DIAGRAM FROM WVI

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.

Project outcomes are what 'comes about' during the course of a project as a result of the outputs achieved.

MONITORING

Evaluations are easier if they can build upon a strong foundation of routine **monitoring** of a program. Organizations can use the terms monitoring and evaluation in different ways. However, monitoring usually refers to a routine and continuous process of collecting relevant program information, analyzing this information at regular intervals (e.g. in quarterly and annual project reports) and comparing actual results to expected results in order to measure a program's performance. The major focus of monitoring tends to be at the level of project inputs, processes, activities and outputs. The sort of questions answered by routine monitoring are: is implementation proceeding as planned? Is there a need for revision or adjustment?

The continuous monitoring of key program information is an important tool for effective project management. For this reason, monitoring data are often integrated into the routine information systems of a program (or its management information system). Organizations usually specify these systems for use across all their program areas, and so we do not provide detailed guidance on them here. However, they can be a useful source of information on which to base more focused evaluation work.

EVALUATION

Evaluation describes work that tries to see if the changes that the program was hoping to bring about have happened. A good monitoring system can provide valuable information on this. But usually, additional actions to collect further information are required.

Exactly what actions are necessary depends upon the questions to be addressed by the evaluation. All psychosocial evaluations should seek to measure the change in the lives of individuals, families and communities that have come about during

the course of a project as a result of programming. This is the core task of an **outcome evaluation**. If done well, outcome evaluations provide valuable evidence that 'beneficiaries' have indeed benefitted from programming. As noted in the Preface, however, for the psychosocial field at this time we often need rather more. We need to show that programs result in lasting changes in people's lives. The rationale for programs often begins by identifying major needs regarding people's well-being and circumstances that must be addressed. We need strong evidence that psychosocial programming leads to impacts on the issues that were used to justify planning and funding this intervention. These need to be sustained changes that will continue beyond the end of the life of a project or program. Such questions are the focus of impact evaluations.

Impact evaluations identify the lasting change that has been brought about in individuals, families, communities and their broader environment as a result of an intervention. To build the evidence-base of psychosocial programming, impact evaluations are crucial, clarifying both the positive — and the negative — things that have resulted from a particular intervention.

This guide seeks to help plan such evaluations, although we acknowledge that some programs will only have the resources to conduct outcome evaluations. It is critical for the field, however, for some longer-term impact evaluations — with designs that provide strong evidence that change is clearly a result of programming and not other factors — to be conducted. These may be at country, regional or global level and may be particularly important in circumstances of large-scale humanitarian crises or in considering key interventions (such as the impact of child friendly spaces, or of counseling).

Project impact is the lasting change in individuals, families, communities and their broader environment that results from a project.

GETTING HELP

Performance monitoring focusing on outputs and immediate outcome is usually a routine process built into the running of the program, and is the responsibility of program staff.

Outcome evaluations 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 usually need the expertise of external consultants.[7] However, program managers will need to be able to manage and oversee the design and implementation of these evaluations, and therefore need a strong understanding of how they should be conducted.

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.

© UNICEF/NYHQ2007-2449/PIERRE HOLTZ

3 OVERVIEW OF PSYCHOSOCIAL PROGRAMMING

3 OVERVIEW OF PSYCHOSOCIAL PROGRAMMING

INTRODUCTION

Exposure to the disruption, loss, and violence associated with humanitarian crises places significant psychological and social strain on children and adults, their families and communities. The way in which people 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. A minority indeed may need professional mental health services. It is now widely accepted that early psychosocial interventions must be an integral part of humanitarian assistance.[8,9]

In work with children, 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 moving from trauma-based models of service delivery to those which recognize and strengthen resilience and local capacities. A resiliency-building approach to psychosocial well-being 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 programming
- Strengthening local networks that enable child protection, care, and well-being, such as women's groups or religious networks

A number of these objectives support the restoration of practices and resources disrupted by crisis. This is an increasingly common goal in psychosocial progamming for all ages and populations. However, rather than returning things to the way they were before, the situation of crisis may indeed create opportunities to address longstanding issues of social justice and empowerment for marginalized groups.

Humanitarian crises affect women, girls, boys and men in profoundly different ways. Members of each group face different risks and have different capacities, necessitating targeted interventions to address the various needs of all groups. To create truly inclusive and beneficial humanitarian interventions, all people — women, girls, boys and men — must be taken into account.

The IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings[10] summarize 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).[11]

There are many examples of psychosocial practitioners using different language and concepts to describe their work. Although this can be 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.

© UNICEF/NYHQ2009-0227/JOSH ESTEY

Acknowledging subtly different emphases across organizations, this guide suggests that the following **three domains** as the most helpful to evaluate how well psychosocial programs affect the lives and experiences of children:

1. Skills and knowledge

e.g. knowing how to communicate, knowing how to make decisions, using culturally appropriate coping mechanisms, vocational skills, conflict management, knowing who to go to for information.[12]

© UNICEF/NYHQ2009-2170/TOM PIETRASIK

© UNICEF/NYHQ2009-0234/JOSH ESTEY

2. Emotional well-being

e.g. feeling safe, trust in others, self-worth, hopeful for the future with realistic goals, not worrying about being hungry or sick.[13]

3. Social well-being

e.g. attachment with caregivers, relationships with peers, sense of belonging to a community, resuming cultural activities and traditions, willing and respectful participation in appropriate household responsibilities and livelihood support.[14]

Psychosocial programming is generally related to one or more of these domains — skills and knowledge, emotional well-being and social well-being. 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 Group[15] that sees psychosocial well-being 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

How these issues are understood across different settings — and how agencies feel best equipped to deal with them — vary widely. In situations where malnutrition and disease are major threats, the physical health of children may be seen to be an integral part of their psychosocial well-being, while in

other situations health issues may be seen as outside the core remit of psychosocial programming. For many cultures issues of belonging, meaning, and identity are integrally related to adherence to a religion or faith tradition. Spiritual well-being may be understood much more broadly in some contexts (and outside the remit of psychosocial work in others).

The reality is that psychosocial support is often a main source of addressing children's well-being in humanitarian crises, and inclusive, culturally appropriate understandings of priorities generally need to be adopted. In the case of spiritual well-being, useful guidance on integrating local beliefs and practices within psychosocial programming is available.[16] Psychosocial programming may be a key area to ensure the protection of people's rights under the Geneva Conventions for 'respect for their honor, family rights, religious convictions and practices' and being 'allowed to practice their religion with ministers of their own faith'.[17]

However, the complexity of the linkage between psychosocial issues and other aspects of well-being should not allow so broad a definition of psychosocial work that it is meaningless. In general, the domains of skills and knowledge, emotional well-being and social well-being capture the core of most psychosocial programming. In sections 5 and 6, therefore, these core domains are used to define objectives for evaluation.

CORE PSYCHOSOCIAL ACTIVITIES

Organizations get involved in many different types of activities in their quest to make a difference for individuals, families and communities. A helpful representation of the range of work undertaken in this field is provided in the IASC Guidelines[18]:

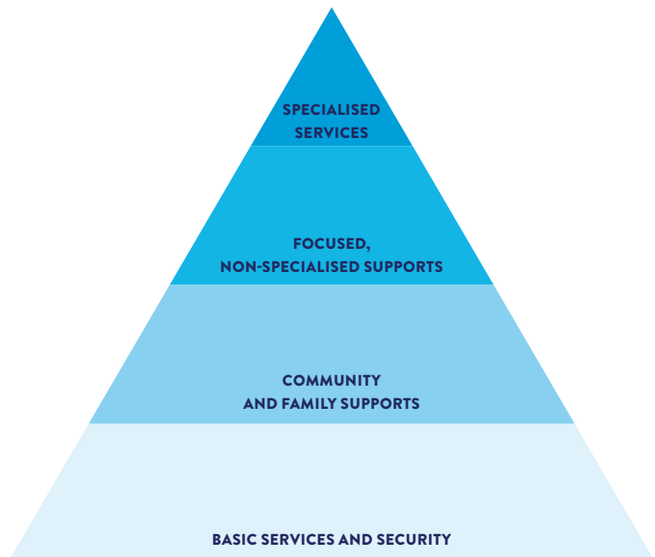


FIGURE 2

Core psychosocial work is typically at levels 1, 2 and 3, that is integration of social considerations in basic security and services, strengthening community and family supports and, providing focused supports. People with psychological disorders are referred to specialized mental health resources (where available). Briefly, services offered at these four levels are as follows[19]:

© UNICEF/NYHQ2010-0068/ROGER LEMOYNE

© UNICEF/NYHQ2010-2022/OLIVIER ASSELIN

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.

© UNICEF/NYHQ2009-0149/GIACOMO PIROZZI

© UNICEF/NYHQ2009-0157/GIACOMO PIROZZI

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. **Specialised 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 of 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 here are, firstly, core principles summarized from the IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings[20]; secondly principles grounded in the Convention on the Rights of the Child[21] and, thirdly, ethical principles for conducting psychosocial evaluations.[22]

CORE PRINCIPLES[23]

1. Human rights and equity

Humanitarian actors should promote the human rights of all affected persons and protect individuals and groups who are at heightened risk of human rights violations. Humanitarian actors should also promote equity and non-discrimination.

2. Participation

In most emergency situations, significant numbers of people exhibit sufficient resilience to participate in relief and reconstruction efforts. From the earliest phase of an emergency, the affected population should be involved to the greatest extent possible in the assessment, design, implementation, monitoring and evaluation of assistance.

3. Do no harm

Work on mental health and psychosocial support has the potential to cause harm because it deals with highly sensitive issues. The risk of harm is reduced by:

- Participating in coordination groups to learn from others and to minimize duplication and gaps in response
- Designing interventions on the basis of sufficient information[24]
- Committing to evaluation, openness to scrutiny and external review
- Developing cultural sensitivity and competence
- Staying updated on the evidence base regarding effective practices
- Developing an understanding of universal human rights, power relations and the value of participatory approaches.

4. Building on available resources and capacities

As described above, all affected groups have assets or resources that support mental health and psychosocial well-being. Key tasks are to identify, mobilize and strengthen the skills and capacities of individuals, families, communities and society.

5. Integrated support

Activities and programming should be integrated as far as possible. Activities that are integrated into wider systems (e.g. existing community support mechanisms, formal/non-formal school systems, general health services, general mental health services, social services, and child protection systems etc.) tend to reach more people, often are more sustainable, and tend to carry less stigma.

6. Multi-layered supports

See section above on ‘core psychosocial activities’.

PRINCIPLES GOVERNING WORK WITH CHILDREN[25]:

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. 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.

Structure and continuity in daily life. Programs 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. Grounding all psychosocial interventions in the culture, except where it is not in the best interests of the child[26], is both ethical and more likely to produce a sustained recovery. Aside from the basic principles of child development and local beliefs about children, those helping 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 EVALUATIONS[27]:

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 adults or children: informed consent should be documented for specific evaluation activities and limited to an agreed time period; 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 well-being 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 can be breached however to ensure the immediate safety of a child or vulnerable adult.

GETTING INFORMED CONSENT:

A CASE EXAMPLE FROM HUMULIZA, TANZANIA

"You will need to explain to the children that have been selected what you are doing. Even the youngest child must understand. During the Humuliza evaluation, the staff explained to the children what an evaluation was by getting them to make a small house outside out of natural objects. Once they had finished making their house, a friend was asked to give them feedback on the house and to make suggestions, if necessary, on how they could improve it. The group then discussed if the house was improved as a result of the feedback. Everybody agreed that it was. The children were then told that this was an example of an evaluation — it gave feedback on the Humuliza house so that it could be made better. You need to tell the children exactly how much time will be involved, where they need to be and what you will do together. Then you need to ask if they are happy to participate. If they are not, you must find other children to replace them." [28]

4 GENERAL GUIDANCE ON EVALUATION

4 GENERAL GUIDANCE ON EVALUATION

Nearly all organizations provide general guidance to their staff for the evaluation of projects and programs. This guide is not seeking to offer alternative evaluation criteria. Rather it aims to provide concrete guidance on how the criteria specified can be assessed in the specific case of psychosocial programs.

A number of agencies have adopted the framework of the standard Development Assistance Committee of the Organization for Economic Cooperation and Development (OECD-DAC) evaluation criteria.[29] These are relevance, efficiency, effectiveness, impact and sustainability.

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 a measure of the costs incurred to produce targeted outputs. This is very important in emergencies when one of the challenges is to take psychosocial programming to scale – for example if comparing two types of programs, which deliver similar benefits for children and the affected population, it is very important to know the relative cost of each.

Effectiveness is measured in terms of the outcomes of a program – what changes have come about for children, 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, lasting change. 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 provide robust evidence that changes — which can be negative as well as positive — 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. This can also reflect changes in political will, economic factors and other developments which increase the likelihood of change being durable.

In humanitarian crises, four additional criteria commonly used are coverage, coordination, coherence and protection:

In psychosocial work, **coverage** will mean the proportion of affected people 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). One of the difficulties when we have good psychosocial programs in emergencies is taking them to scale.

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 — and is thus complimentary to — the work of another, establishing common programming guidelines and strategies, coordinating geographical distribution of programs, establishing referral mechanisms and sharing of resources and information.

TABLE 1 **KEY EVALUATION QUESTIONS RELATED TO PSYCHOSOCIAL PROGRAMMING BASED ON OECD/DAC EVALUATION CRITERIA**

GENERAL EVALUATION CRITERIA	KEY QUESTIONS IN CONTEXT OF PSYCHOSOCIAL PROGRAMMING
RELEVANCE	• Did the program articulate objectives related to changes in children's well-being 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	• Have inputs resulted in the outputs targeted? • How did costs compare to other programs targeting similar outputs?
EFFECTIVENESS	• Have stated program outcomes been achieved? • What difference has come about for beneficiaries in terms of skills and knowledge, emotional well-being, and social well-being? • What factors contributed to success or failure with regard to targeted changes?
IMPACT	• Has the central goal of the project – the needs that provided the rationale for intervention – been met? • What lasting changes – attributable to programming — can be identified in the lives of individuals, families, communities and the broader environment? • Did any negative changes result from programming?
SUSTAINABILITY	• 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? • Have root causes (such as attitudes to children) been impacted?
COVERAGE	• 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 appropriate addressed?
COORDINATION	• Have agencies worked well together towards the common goal of improved psychosocial well-being amongst children?
COHERENCE	• Has work been consistent with the IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings?
PROTECTION	• Does the project contribute to protecting children by strengthening the child protection mechanisms such as legislation, services, community norms etc?

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

Finally, in terms of **protection**, the issue for psychosocial programming is whether activities have worked to strengthen child protection mechanisms (such as laws, services and community norms) supporting children and others at risk. Psychosocial programs need to advocate for affected populations, and making sure they do not inadvertently contribute to protection issues.

As noted above, many organizations have developed general evaluation frameworks reflecting their specific mandates (see, for example, World Vision International and Save the Children).[30] Many of the issues listed above are reflected in these frameworks. As with the guidance that follows, the aim is to encourage overall good practice in evaluation design and implementation within the policy and processes of specific agencies.

© UNICEF/NYHQ2010-0807/KATE HOLT

5 OBJECTIVES OF PSYCHOSOCIAL PROGRAMMING

5 OBJECTIVES OF PSYCHOSOCIAL PROGRAMMING

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

Evaluations examine how successful programming has been in achieving what it set out to do and what the consequences are for beneficiaries' psychosocial well-being. 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 psychosocial well-being.

Project documents usually feature different types of objectives, often in the form of a 'results framework' or 'logical framework'. These will not only include the key 'inputs' and 'activities' planned, but also the project 'outputs' that should result, the expected project 'outcomes', and a clear statement of the lasting changes in the lives of individuals, families and communities that the project seeks to encourage: project 'impact'.

The precise language used by organizations to refer to these 'levels' of the logic of a program differs widely (see Table 2). Some organizations talk of 'results' and 'objectives'; others talk of 'purposes' and 'goals'. Some distinguish between 'intermediate outcomes' and 'end outcomes' rather than between 'outcomes' and 'impacts'; still others distinguish two types of 'outputs'.

© UNICEF/NYHQ2004-0707/GIACOMO PIROZZI

Although the exact terminology used varies, it is crucial to distinguish between these levels. The framework we are using provides a key tool for both monitoring and evaluation specifying:

- the key steps required to effectively implementing the work, and
- the benefits that it is anticipated will result from the project.

To strengthen psychosocial programming, evaluations need to always look at what has been achieved at all levels of the program. However, as noted earlier, to strengthen the evidence-base of what works in the field, it is crucial that particular attention is given to the 'higher levels' of outcomes and impacts.

TABLE 2 **COMPARISONS BETWEEN TERMINOLOGIES OF DIFFERENT DONOR AGENCIES FOR RESULTS / LOGICAL FRAMEWORKS**

	ULTIMATE IMPACT	END OUTCOMES	INTERMEDIATE OUTCOMES	OUTPUTS	INTERVENTIONS	
NEEDS-BASED	Higher consequence	Specific Problem	Causes	Solution	Process	Inputs
CARE TERMINOLOGY	Program Impact	Project Impact	Effects	Outputs	Activities	Inputs
CARE LOGFRAME	Program Goal	Project Final Goal	Intermediate Objectives	Outputs	Activities	Inputs
PC/LOGFRAME		Goal	Purpose	Outputs	Activities	
USAID RESULT FRAMEWORK	Strategic Objective	Intermediate Results		Outputs	Activities	Inputs
USAID LOGFRAME		Final Goal	Strategic Goal/ Objectives	Intermediate Results	Activities	
DANIDA + DFID	Goal		Purpose	Outputs	Activities	
CIDA + GTZ	Overall Goal		Project purpose	Results/ outputs	Activities	Inputs
EUROPEAN UNION	Overall objective	Project purpose	Results	Activities		
FAO+UNDP+NORAD	Development Objective		Immediate Objectives	Outputs	Activities	Inputs
UNHCR	Sector Objective	Goal	Project Objective	Outputs	Activities	Inputs/ Ressources
WORLD BANK	Long-term Objectives		Short-term Objectives	Outputs		Inputs
AUSAID	Scheme Goal		Major Development Objectives	Outputs	Activities	Inputs
RESULTS FRAMEWORK	**IMPACT**		**OUTCOME**	**OUTPUTS**	**ACTIVITIES**	**INPUTS**

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.

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.

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)

K=Knowledge and skills; E=Emotional well-being; S=Social well-being.

Note that at the **Output** level, objectives are typically relevant to more than one domain.

LEVEL TWO: OUTCOMES

Project outcomes are what 'comes about' during the course of a project as a result of the outputs achieved.

For psychosocial programs, projects usually seek changes relevant to one or more of the 'domains' of children's lives, described 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 support the psychosocial well-being of children in the longer-term.

An evaluation at this level would look at what differences have occurred as a result of individuals, 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.

To identify project outcomes ask: What changes have come about for individuals, families and their communities through programming?

© UNICEF/NYHQ2010-1222/ROGER LEMOYNE

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
- Decrease in frustration, anger and aggression
- 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 lasting change in the lives of individuals, families, communities and the broader environment that results from a project.

It is extremely important to see if project outcomes lead to lasting benefits. For this to happen, changes need to be able to be sustained over time (that is, they continue beyond the end of the project). They may also represent changes in the broader environment, shaping the lives of communities, families and individuals.

To identify project impacts ask:
What is the lasting change in people's lives that was hoped for when proposing the project?

Project documentation usually begins with some statement about current circumstances 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 example about Sierra Leone given in this section, although achieving outcomes such as acquiring vocational skills, gaining a greater sense of hope and establishing mentorship relationships would all have been welcomed 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.

Impact evaluations should not just examine whether targeted impacts have been achieved. As well as these intended consequences, they need to consider the unintended consequences of psychosocial programming that have arisen. These can be negative, as well as positive. For instance, while children associated with armed forces might have gained skills from a program 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 program.

© UNICEF/NYHQ2008-1779/GIACOMO PIROZZI

PSYCHOSOCIAL SUPPORT THROUGH SCHOOLS FOR WAR-AFFECTED YOUTH IN SIERRA LEONE (3)

The Impacts targeted in 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-affected youth enhanced

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

Organizations face many challenges in trying to demonstrate such impacts. 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 objectives.[31] If objectives remain vague or unstated 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 (see Annex A for a tool to assist in this).

DIFFERENT APPROACHES TO MAPPING OUTCOMES AND IMPACTS

The previous section of the guide about impacts — and indeed the whole guide — is built around the idea of a 'results framework'. This structure of agreed objectives — at a range of levels — specifies not only what a program aims to achieve but how it will achieve it. It is thus a key tool for managing a program.

Some people argue that 'management by objectives' is difficult to achieve in humanitarian or development contexts, where many complex factors are at work — many outside of the control of an agency implementing a program.

An approach called **outcome mapping** (OM)[32,33] has emerged as an alternative way to structure planning, monitoring and evaluation. It is designed for situations when attributing a change (in a community or children's well-being) to a specific intervention would be difficult (for example, when there are a number

© UNICEF/NYHQ2010-1414/MARC HOFER

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.[a]

of organizations working in the community on different, but related, projects). Although OM is a very different approach to evaluation, there have been some interesting examples of combining it with the sort of approach taken in this guide.[34]

SETTING OBJECTIVES RELATING TO INDIVIDUALS, FAMILIES AND COMMUNITIES

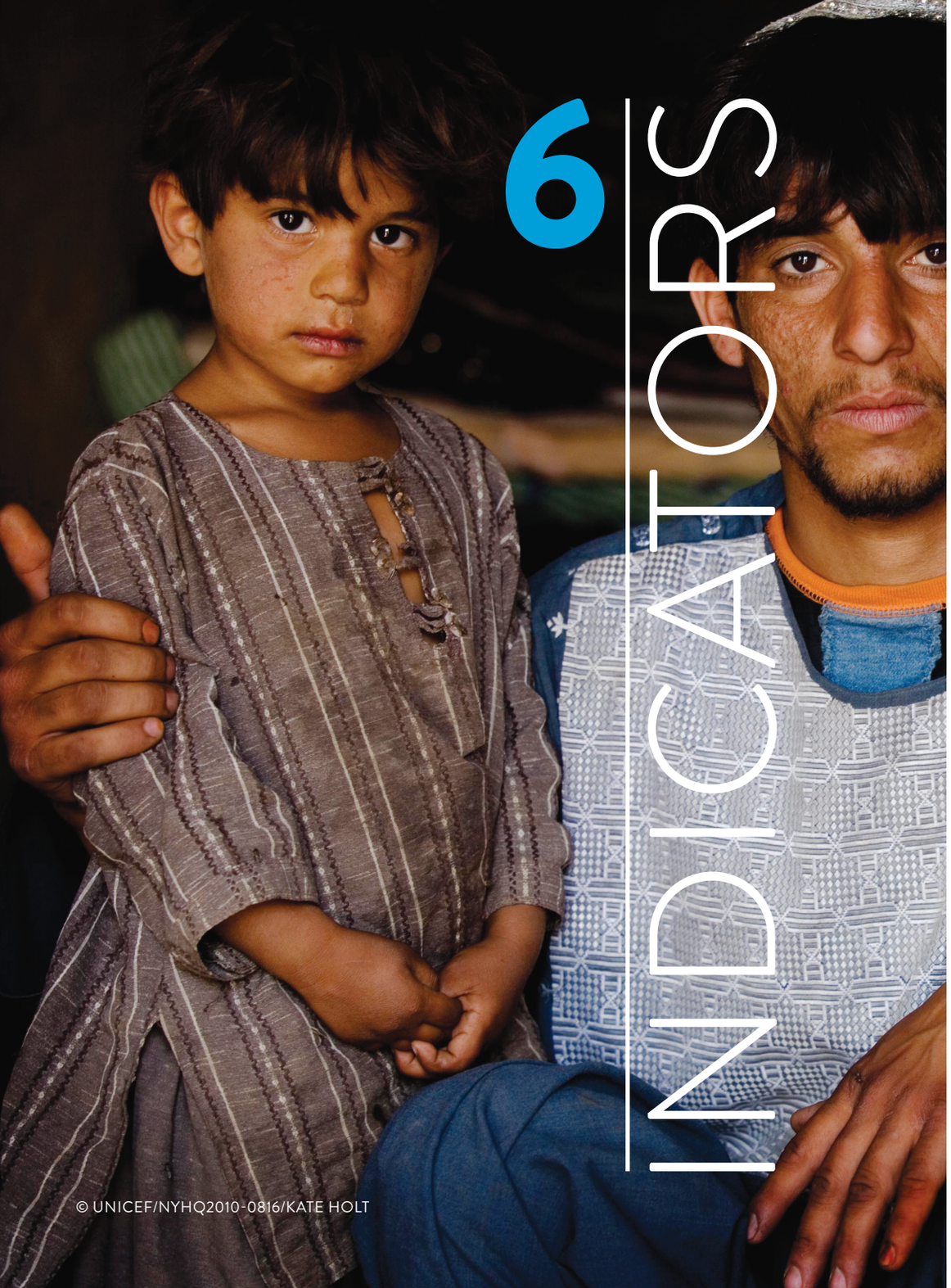
It is very important that the objectives of psychosocial programming – whether these are outputs, outcomes or impacts – are not decided upon by project staff alone, but with active engagement with beneficiaries and other relevant stakeholders.[35]

When objectives are discussed in this way, usually this includes not just aspects of the lives of individuals, but also issues related to their families, and the wider community. 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 3 is an example of the kind of objectives identified to strengthen psychosocial support to children and youth in Palestine, developed through practitioner consultation and later refined with children, parents and teachers.

TABLE 3 **SAMPLE OBJECTIVES ACROSS THE CORE DOMAINS RESULTING FROM A PALESTINIAN CONSULTATION**

FOR CHILDREN	OBJECTIVES
SKILLS AND KNOWLEDGE	• Higher Increase in the use of locally defined life skills by children (e.g. adolecents 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 WELL-BEING	• Increase in the self-confidence, playfulness and sense of security of children attending activities • Increased sense of locally defined purpose and meaning
SOCIAL WELL-BEING	• 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	**OBJECTIVES**
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 WELL-BEING	• 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 WELL-BEING	• Increase in engagement of parents in activities that support children's development
FOR COMMUNITIES	**OBJECTIVES**
SKILLS AND KNOWLEDGE	• Increased open debate on psychosocial issues in public forums (media, community meetings etc.)
EMOTIONAL WELL-BEING	• Increase in access to psychological and social services by vulnerable/ marginalized groups • Reduction in levels of distress of front-line workers
SOCIAL WELL-BEING	• Community involvement in social and cultural activities • Decrease of interfamily violence and decrease of causalities 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 used. 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 a change 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 well-being or social well-being 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 psychosocial programs will help in attempts to establish a more robust evidence-base for the sector. Over time it should then be possible to compare results across programs and countries.

However, given the different ways that psychosocial well-being 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 domains of **skills and knowledge, emotional well-being** and **social well-being** which were highlighted in the introduction to psychosocial programming. For most psychosocial programs it should be possible to specify outcome and impact indicators for each of these domains:

Domain	Core indicators
Skills and knowledge	Some measure of acquisition of skills
Emotional well-being	Some measure of improved emotional adjustment
Social well-being	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 using both quantitative and qualitative methods (see Section 9). We can also use existing information to help define relevant and practicable measures of change.

© UNICEF/NYHQ2010-1755/LEONIE MARINOVITCH

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 illustrate how existing information can be used in this way. 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 addressing girl empowerment might specify as an indicator the proportion of girls enrolled in school (which could be obtained from school records)
- a program focused on parent health 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 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 — for child-focused interventions — children, their caregivers and 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 4 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.

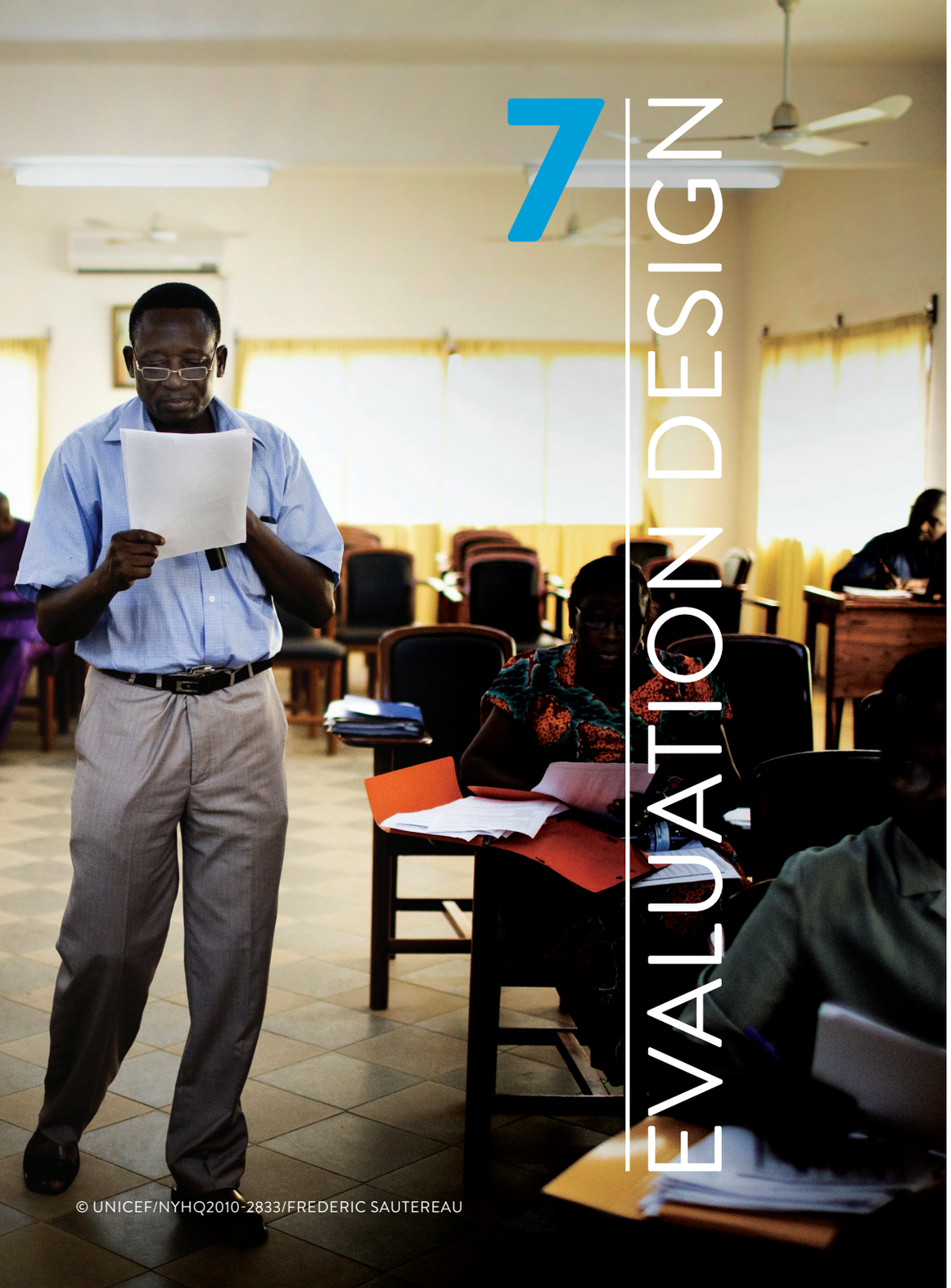
TABLE 4 **DEVELOPING INDICATORS FOR PSYCHOSOCIAL PROGRAMMING IN EASTERN SRI LANKA**

DOMAIN ▸ INDICATORS ▾	SKILLS AND KNOWLEDGE	EMOTIONAL WELL-BEING	SOCIAL WELL-BEING
IMPACT	Quality of care provided to children by families improved *[measured by interview with children and parents, and by reduced incidence of reports of child abuse].*	Stronger referral networks between communities service providers *[measured by increased case loads of community referrals].*	Reduction tensions between IDPs and host communities *[measured by community interview, and by reduction in police reports of social disturbances involving Internally Displaced Persons (IDPs)].*
OUTCOME	Increased awareness of parenting skills and children's rights *[measured by increase of knowledge in parents and children].*	Appropriate referral of children requiring specific support *[measured by increase of referrals through District Child Protection Committee].*	Reduction tensions between IDPs and host communities *[measured by community interview, and by reduction in police reports of social disturbances involving Internally Displaced Persons (IDPs)].*
OUTPUT	Psycho-educational program delivered through drama and theatre activities for children and families in all IDP camps and host communities in Batticoloa *[measured by number of program sessions delivered, number of sites for delivery].*		

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 parenting skills or children's rights? Or 'how much reduction in IDP-host community conflict would represent a good impact, and how will we measure this?' There are a number of sources of guidance available to help project evaluation teams develop unambiguous, measurable indicators of this sort.[36] However, as with many other issues, these are best addressed at the planning stage of a program. Indicators can then be used to define key milestones and targets for programming, drawing upon baseline data collected.

Annex A provides a template for developing and refining objectives and indicators for a psychosocial program, using the ideas in this section. Annex E provides some case examples that involved the development of appropriate indicators for particular programs.



7 EVALUATION DESIGN

© UNICEF/NYHQ2010-2833/FREDERIC SAUTEREAU

7 EVALUATION DESIGN

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

This section looks specifically at how to design an evaluation to measure project outcomes and impacts. Evaluation design should be considered early in the life of a project, as soon as clear objectives and a set of corresponding indicators have been developed. Ideally, ways of collecting information for evaluation will be set in place throughout the whole project management cycle. In fact these design principles can also be used to plan formative and mid-term evaluations, as well as overall project outcome and impact evaluations. However, the guide recognizes that evaluations often have to be conducted when these things have not been put in place.

This section looks at how comparisons can be made for evaluation, beginning with **baselines** (planned ideally before programming begins), then describing different ways that **comparison groups** can be set up and then looking at the merits of designing evaluations that '**look forwards**' and '**look backwards**'.

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.

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 the well-being of individuals, their families and communities and so there is no reliable way to measure changes over time.[37]

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 program are much stronger than those that don't have such measures. Needs assessments/situation analyses may have been conducted before beginning programming (as is good practice). If done correctly, then this information 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 humanitarian crises. And conducting a baseline potentially provides a hugely useful base for program planning and development, as well as evaluation.

COMPARISON GROUPS

A comparison group that receives no intervention (sometimes referred to as a 'control group') is a group of people that in every way — except their not receiving an intervention — is as similar as possible to those receiving the intervention. People and communities change and adapt, in both negative and

In Sierra Leone, IRC followed up a baseline assessment in 2002 with an interim evaluation in 2004. Measures of confidence, prosocial behavior, depression, anxiety and hostility were used.[c]

Prior baseline assessment allowed War Child Holland to conclude from a 2006 assessment in Sierra Leone that awareness on child rights had increased from 40% to 95% after one year implementation of their community based psychosocial program.[d]

positive ways, after crises whether there are interventions in place or not. It is important in judging the effects of an intervention to consider these changes. Change may have occurred with the group receiving an intervention, but is this any different from change in evidence with other members of the population who didn't?

© UNICEF/NYHQ2009-1236/GIACOMO PIROZZI

© UNICEF/NYHQ2010-1221/ROGER LEMOYNE

To evaluate the impact of its PSSA intervention in schools, Save the Children in Uganda compared children's experience across schools that received the intervention in 2007 and schools where it was planned to be 'rolled out' in 2008.[64]

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

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 3 on principles of psychosocial support). Providing a comparison group is established sensitively, potential harm should be minimized and this will be better than having no comparison group at all.

An evaluation in northern Uganda directly compared an innovative group counseling intervention with a play therapy intervention offered by another organization.[39]

One way of establishing a comparison group arises in the common practice of rolling out programs over time. 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.

Establishing a comparison group should be given high 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.[38]

Another way of making some form of comparison between those who receive an intervention and those that don't is by using information about existing patterns or 'norms' for beneficiaries. 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. If fewer than 50% of girls of a particular age are known to generally be in school and an intervention results in over 80% now being enrolled, that is an 'indicator' of success.

A final strategy for having some basis of comparison is to compare two interventions with each other. If it is not clear which of a number of approaches to a particular issue is more effective — and that is often the case with psychosocial programming — 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 people fare without any intervention, however. Identifying a non-intervention group would also be preferable in this circumstance.[39]

Whatever strategy is used for a comparison, the comparison group really does need to be 'as similar as possible' as those receiving the intervention. This usually means finding people who are of a similar range of ages; similar gender balance; similar cultural or religious background; and have had similar experience related to the crisis, as those who are receiving the intervention.

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.[b]

LOOKING FORWARDS

One of the most powerful designs is one where we combine the features of a baseline and comparison groups. We take measurements at the start of the project (i.e. a baseline) and at various points during the project, assessing those who receive access to an 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 that can be linked to the project intervention. This design is particularly suited to impact evaluations. It can be used for outcome evaluations, but the costs can be significant. Also, if assessments points are too close together in time, little change on indicators is likely to be found.

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 program 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.

LOOKING BACKWARDS

Sometimes, indeed quite often, we have to try to evaluate a project for which no clear objectives or indicators were originally identified and no good baseline measures completed. 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 this[40]:

USING EXISTING DOCUMENTS

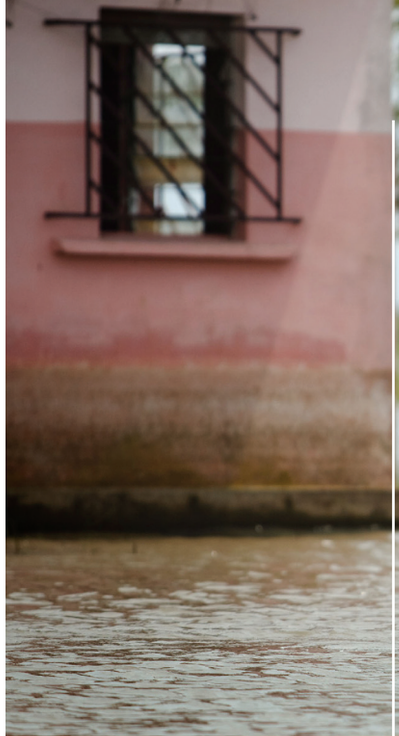
'Secondary data', such as data from health and education agencies, government surveys, school enrolment and attendance records, project records, can be used. It is important to check that these records have been completed accurately and regularly. If so, they can be very helpful.

USING RECALL

It is generally not a good idea to expect children, families, communities or workers to remember in detail how they felt or behaved a long time in the past. This information may not — with the passage of time — be very accurate, and recalling past events can also be experienced as intrusive. There may be a risk of reviving stressful memories.

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. Key informants, such as community leaders, teachers, nurses and doctors, may be able to pinpoint key events — such as the opening of a clinic or new bridge, or the appointment of a new chief – in the life of the community. Beneficiaries can then be asked to recall information relevant to the evaluation with respect to these events (e.g. were you married before or after the bridge opened? Were you in school when the new chief was appointed?).

In conclusion, evaluations should normally plan to use a design with a baseline and, wherever possible, an ethically sound comparison group. Where conducting effective baselines is really not possible — for example, after rapid onset emergencies like the Indian Ocean tsunami of 2004 — then establishing an adequate basis for comparison of those receiving support with those who (currently) are not is an even greater priority if the effectiveness of intervention is to be examined.



In evaluating its programs 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.[e]

© UNICEF/NYHQ2009-1898/JOSH ESTEY

8 PREPARING FOR AN EVALUATION

ENSURING PARTICIPATION

The participation of children is vital:

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

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.

There are a range of diversity issues that should be considered in terms of participation.[41] Age and gender are, in most circumstances, especially crucial. Women, men, girls and boys should all be involved to ensure that no group is marginalized or overlooked. Women, girls, boys and men offer different perspectives on the same problem, and it's important to capture these experiences.

Dialogue with women, girls, men, boys and older people should occur in settings in which they are comfortable. Cultural differences may inhibit women from talking in front of men and vice versa. Adolescent boys and girls, and older people have valuable ideas and should not be overlooked. Some vulnerable groups will require special support to participate, e.g. transportation, childcare, a male relative to accompany them. Women typically have less access, due to factors such as language skills, community leadership (being predominantly male), mobility and time (women and girls typically undertake childcare or household duties).[42]

© UNICEF/NYHQ2007-2043/SHEHZAD NOORANI

ENGAGING CHILDREN IN PREPARATION FOR AN EVALUATION

Humuliza is an organization in Tanzania that works with orphaned children. The program is centered around a children's organization run by the children themselves. They did an impact evaluation and used the children's quotes to help formulate indicators[45], for example:

SKILLS AND KNOWLEDGE

Survival knowledge (HIV/AIDS) "I know how I can take care of myself."

Income generation "My capital for selling small fish has expanded."

SOCIAL WELL-BEING

Social support network "I felt like I am the only orphan child. Now I know there are others.", "I could not play before. Now I have many friends."

EMOTIONAL WELL-BEING

Stress "My head was full of feelings before."

Confidence "I can talk with confidence.","How to live even if life is difficult"

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

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.[43] Children should be involved in the planning, implementation and evaluation of programs. Their involvement adds authenticity to all these processes.

Involving children in program evaluations may itself have significant benefits for children's well-being. Participation may empower children by valuing their opinions and enhancing their knowledge and skills. Opportunities for social interaction also contribute to children's psychosocial well-being being linked with social and emotional development.[44] 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 the context of crisis. It also requires an appreciation of the political, social and economic factors that favour (or limit) the participation of particular groups. Analysis can reveal how the context limits the participation of different members of society, as well as how participation can be increased.

There are different levels of participation. Arnstein's influential 'ladder of participation'[46] provides a useful prompt to understanding the way we might be 'involving' beneficiaries. 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. Based upon Arnstein's thinking, some organizations working with children refer to the different approaches of 'consultative participation', 'collaborative participation' and 'child-led participation'.

For the most part, children and other members of the affected community 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 who review a program, for example, or working in 'partnership' with adults to complete an evaluation. Such roles should be encouraged, if it is in the best interest of the child.

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 well-being
- providing their views on the project
- providing feedback on draft results and the implications for their community/future similar projects

MOVING THE GOALPOSTS, KILIFI [F]

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 strategy.[47] 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 respectful way is essential. You also need to decide how to ensure that ethical considerations (e.g. ensuring appropriate confidentiality and consent, and minimizing any risk from participation) are addressed (see Section 3).

All activities should be gender sensitive, culturally appropriate, and, where children are involved, child focused.

There are particular challenges in participative work with children and young people. The Population Council's guide 'Ethical Approaches to gathering information from children and adolescents in international settings – guidelines and resources'[48] provides extensive detailed guidance in this area. We also highlight here some general points about working with children based upon fieldwork in Afghanistan and Sri Lanka.[49] 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 topics, where appropriate.[50]

EQUIPPING STAFF

We also need to consider the capacity of project staff to engage effectively in evaluations. Training may be needed to prepare all those involved, and supervision and support will need to be in place.

A training package, *The Psychosocial Support Monitoring and Evaluation Technical Training Kit*[51], has been developed to accompany this guide. There are a number of additional resources that may be helpful in training and development of staff in evaluation work.[52,53] Here is a list of core competencies identified after a review of frontline staff's successes and difficulties in implementing evaluations.[54] The list describes the skills needed to work 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 participants, encourage participation, explore/elicit participants' views, manage group dynamics);
- Conceptual and practical knowledge related to aspects of psychosocial well-being, to be able to understand the meanings/implications of people'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 participants.

© UNICEF/NYHQ2005-0515/RACHEL DONNAN

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.

The choice of methods is much easier if indicators have already been clearly defined for the program. This clarifies the core key information that is required. There is often a strong temptation to collect too much information. Indicators should be a good guide to the information that is really important.

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 complements 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 builds a fuller picture.

The most important factor, clearly, is choosing methods that will enable you to collect information that will allow you to judge whether an intervention has been successful. Seek technical advice on choice of methods from M&E colleagues or researchers at local academic institutions if it is available. If not, some of the resources listed at the end of this guide may be of assistance.

© UNICEF/NYHQ2011-0091/SHEHZAD NOORANI

In an assessment of *Palestinian children*, focus groups discussions explored children's concerns and aspirations. 95 focus groups were held. At the end of each focus group, facilitators recorded the three most prevalent responses to each prompt question. Using quantitative analysis to combine information across all the groups, it emerged 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.[8]

QUALITATIVE METHODS

Using qualitative methods can provide vivid insights into the experience, perceptions and beliefs of program beneficiaries and other relevant stakeholders. Qualitative methods allow us to gather detailed information of people's experiences and descriptions of the places where they live and work and play. If done well, qualitative methods allow the voices of the affected people to be heard and can bring to life the reality

of the changes that the humanitarian crisis and the psychosocial intervention have brought about. After all, we do evaluations not only to find out if staff have done the work that they said they would do, and not only to show that some change has taken place, but to understand how and why these changes have come about. Used appropriately such methods can provide robust, insightful information to support an evaluation.

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.

As discussed later, qualitative methods can produce large amounts of unstructured data that can be challenging to collate and analyze. It's important to only collect information that will be analyzed, which again means being clear on the information that is really important. If time for analysis is a major constraint, consider using qualitative methods that involve some sort of group ranking or mapping of issues, which significantly reduces time for analysis.

Although the activities described may appear unobtrusive and unthreatening, any work of this kind may trigger strong emotions. This needs to be anticipated and adequate support put in place to support children or adults in distress. This should reflect the principle of 'do no harm' and use of the ethical guidelines discussed in section 3.

Here is a range of different qualitative methods:

FREE LISTING

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

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

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'. This involved children developing a long list of the qualities they associated with resilient children. The listed qualities were then discussed and grouped under 6 major headings: playful & sociable; intelligent; happy; respectful; responsible; and healthy. The evaluation was then focused around these local 'indicators' of resilience.[55] A similar exercise was conducted with parents and teachers which revealed — despite some overlap — subtly different perspectives on what signaled children's well-being (all of which were used in the evaluation).

In another example[56], a 'Well-being 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 well-being.

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,

© UNICEF/NYHQ2009-1029/SHEHZAD NOORANI

ASSESSING AFGHAN CHILDREN'S PSYCHOSOCIAL WELL-BEING: A MULTI-MODAL STUDY OF INTERVENTION OUTCOMES[57]

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?

adding detail or correcting one another, and in this way the information gathered is likely to be more accurate.

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, suggested that this was among their greatest 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.

PARTICIPATORY RANKING METHODOLOGY (PRM)

PRM is an approach to convening a focus group type discussion that draws on the ideas of 'free-listing' and other participatory methodologies.[58] Participants — typically a group of between eight and ten (but this can be larger) — identify key issues, challenges or resources related to a framing question posed by a facilitator.

A physical object is identified by the group for each of the issues raised. These are placed in a pile in front of the group, who then proceed to rank these issues in order of importance. This produces a rank ordering of issues that can readily be collated and/or compared across groups (e.g. men's groups vs. women's groups; boys' groups vs. girls' groups). Also, the account given by participants justifying the relative ranking of issues is recorded, and serves as a rich source of narrative quotations from participants. The method is generally experienced as an open, even entertaining, means of discussing the relative importance of a group's concerns.

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 — reflecting the diversity of the community — 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 or adults in drawing a map of some kind which is then used to generate discussion. There are many different kinds of mapping tools. Participants 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 well-being.

© UNICEF/INDA2009-00013/ANITA KHEMKA

For children in Kabul[h], 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

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 participants drawing a map of their immediate surroundings and community and other areas they frequently visit. This method identifies the things that group members 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 Lanka.[59] This enabled comparison of the circumstances of children 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.

CREATIVE SELF EXPRESSION

Drawing, drama and other forms of self expression can be used to gain insight into needs, concerns, ideas and opinions. Children's drawings, for example, provide 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.

One of the challenges in evaluation is to talk with children and young people about their feelings. 'Feelings are an abstract idea that children sometimes find hard to articulate, and in many languages there are very few words to describe feelings'.[60] **Dino pictures** are a tool described in the REPSSI manual, 'Are We Making a Difference?' (which outlines a wide range of participative methods for use with children). Children are taught the emotions associated with each card, and then encouraged to indicate the card (emotion) that they associate with a particular situation. This provides a potentially reliable means of assessing children's emotional well-being, and its change over time. The images to the side depict two of the cards, representing 'happy/glad/joyful' and 'angry/cross' respectively.

In evaluating psychosocial programs addressing the needs of children affected by conflict in the Balkans, Kostorova-Unkovskai 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.

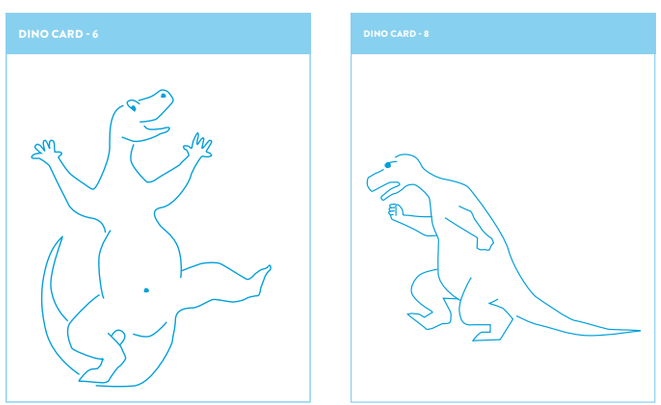


FIGURE 3
DINO PICTURES
FROM REPSSI MANUAL

Particular forms of drawing can be used to address key issues in the lives of adults and children. For example, with a **timeline** participants 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, participants present their timelines to one another and discuss the different events and when they happened.

In a **lifeline**, participants 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 (figure 4).

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.'[61] 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.[62]

All of these methods need to be used in a manner that fits in with the capacities and culture of the communities in which information is being collected. The example below demonstrates what this may mean in practice: adjusting planned methods in the light of realities in the field.

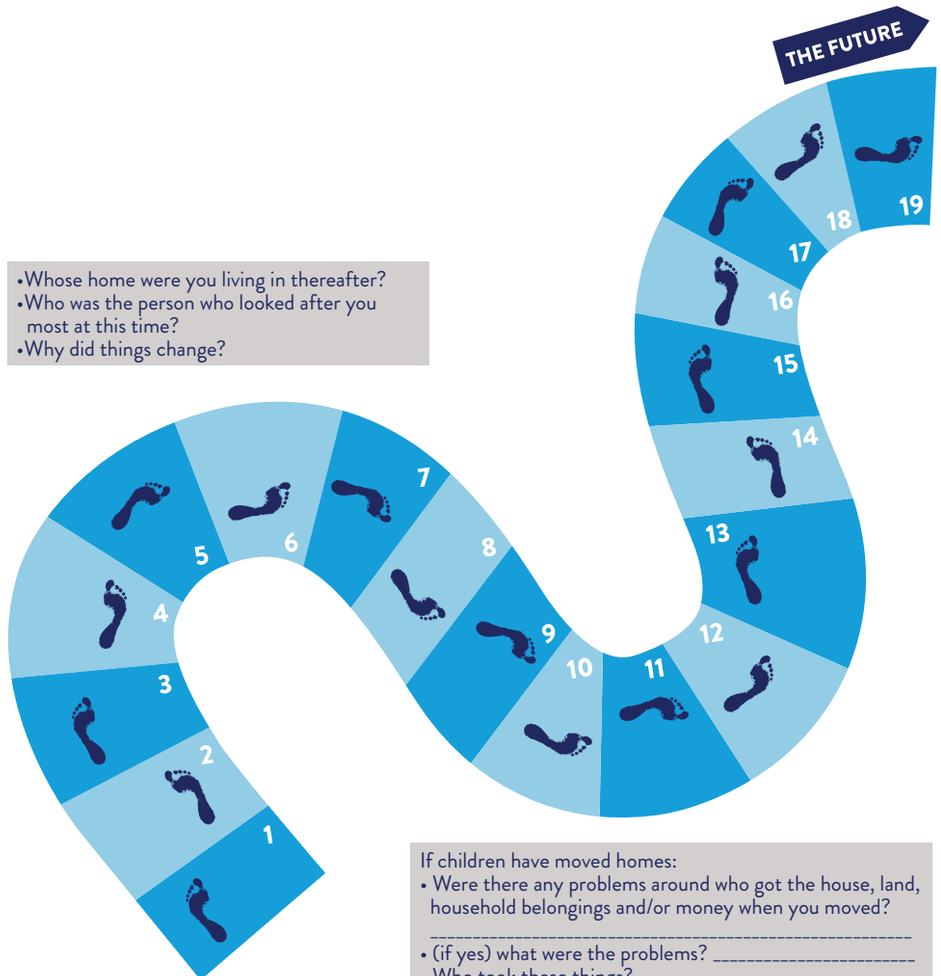
© UNICEF/NYHQ2010-0664/OLIVIER ASSELIN

REVIEW OF UNICEF SUPPORTED RIGHT TO PLAY INTERVENTIONS IN RESPONDING TO THE PSYCHOSOCIAL NEEDS OF CHILDREN AFFECTED BY CONFLICT IN NORTHERN AND EASTERN UGANDA[J].

'Three, four-hour evaluation workshops were held with 100 children, whose aged ranged from 9 to 16 years old. The children's workshops were held in schools on a Saturday or a Sunday within the different program's implementing areas. We found out during the first focus group discussion with the youth that adolescents in the last few years of primary school struggled to read and write in Acholi. In response to this, the facilitator-administered questionnaire was abandoned and four general questions were asked during the group discussions instead. The first three exercises were adapted to be administered verbally in Acholi and to include games, music and drama in their execution. So, for example, the Coat of Arms with the incomplete sentences were drawn on a flipchart paper and discussed with the group. Then, using movement, song and drama re-enactments, the children completed the sentences and the co-facilitator wrote down all the children's responses.'

FIGURE 4

MY JOURNEY OF LIFE



UNPUBLISHED QUESTIONNAIRE DEVELOPED FOR THE: RESEARCH ON THE IMPACT OF PROGRAMMATIC INTERVENTIONS ON THE WELL-BEING OF ORPHANS AND VULNERABLE CHILDREN IN KAFUE DISTRICT, LUSAKA, ZAMBIA. A RESSEARCH COLLABORATION BETWEEN REPSSI AND THE SWISS ACADEMY FOR DEVELOPMENT, 2008.

FIGURE 5

MY LEISURE ACTIVITIES

Please circle the activities that you usually do in an ordinary week.

Did you have an opportunity to play and have fun last week?	☐ YES	☐ NO
Do you have a place to play?	☐ YES	☐ NO

UNPUBLISHED QUESTIONNAIRE DEVELOPED FOR THE: RESEARCH ON THE IMPACT OF PROGRAMMATIC INTERVENTIONS ON THE WELL-BEING OF ORPHANS AND VULNERABLE CHILDREN IN KAFUE DISTRICT, LUSAKA, ZAMBIA. A RESSEARCH COLLABORATION BETWEEN REPSSI AND THE SWISS ACADEMY FOR DEVELOPMENT, 2008.

QUANTITATIVE METHODS

There are a number of types of quantitative measure. These include: surveys and questionnaires (e.g. community-based surveys, standardized interviews, etc.); project records, registers and other service statistics; and observation (with respect to a checklist of behaviors).

Quantitative methods are focused on telling you 'how much', 'how many', and 'to what extent'. Information is generally presented in the form of numbers and percentages. If a sufficient number of people are properly sampled, conclusions can be generalized to the larger population from which the sample has been drawn. Measures may give insight into levels of skills and knowledge, emotional well-being and social well-being, and many other characteristics. Concerns about the use of quantitative methods in the psychosocial field usually focus on the issue of how well such measures actually capture the experience of participants with respect to these domains. Quantitative methods can also be limited in their ability to shed light on why things are the way they are.

DEVELOPING LOCAL MEASURES

With quantitative methods there are two main approaches. One is to develop a survey, questionnaire, 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. However, this is a complex process and can be very time-consuming. It is wise to seek expert advice so that the findings will be sound, and potentially 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

© UNICEF/CBDA2004-00378/JOHN VINK

In 2006 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.

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.

TABLE 5

PERCEPTIONS OF CHILD WELL-BEING

'HOW WOULD YOU DESCRIBE A CHILD THAT IS DOING WELL, A CHILD THAT THE COMMUNITY ADMIRES?'

DOMAIN	INDICATORS
SOCIAL WELL-BEING: POSITIVE SOCIAL FUNCTIONING BEHAVIOR.	The top ranked indicator of a child that is doing well, a child that the community admires, is a child with 'good values' that 'feels useful in the community'. This is a child that is respectful and obedient, primarily indicated by respect towards elders, visitors, parents and other adults; carrying out instructions; completing household and agricultural tasks. With girls, this is also indicated by the girl 'respecting herself' in terms of how she relates to boys and men.
SKILLS AND KNOWLEDGE AND EMOTIONAL WELL-BEING: POSITIVE APPROACH TO LEARNING	A child that is doing well is also one that can 'think well'/is intelligent – is at school or has gone back to school; has good or is improving their school performance; is good at sports or music, dance and drama; has learnt traditional stories, music and dance, and can 'discuss useful matters' like giving good advice to friends and useful ideas on academic issues.
SOCIAL AND EMOTIONAL WELL-BEING: PRO-SOCIAL BEHAVIOR	A social, friendly and caring child that likes to associate with other children her age.

Table 5 provides is an example from Uganda[63], where responses were used to develop a survey form to assess the extent to which children were 'doing well' in the terms specified by the local communities. A similar approach was used Save the Children in Uganda's evaluation of its Psychosocial Structured Activities Program in schools noted earlier.[64]

The survey questionnaire aimed at children and adolescents below was developed 'from scratch' in Zambia. Many of the questions, examples and pictures that were used come from the children themselves. It was piloted extensively to ensure that it was child-friendly and culturally appropriate (figure 5).

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 well-being. Tools imported from one culture to another without careful consideration can result in misleading comparisons.

The Child Behavior Checklist (CBCL) has been used in many cultural contexts, and there are 'norms' for many settings.

Pre-existing psychosocial assessments should only be used when there is clear evidence of their validity in a given cultural context. This will usually mean having the measure translated into the local language and then 'back-translated' into the original language to check that concepts have been properly understood. Best of all is using a measure that has been used before in that cultural setting and for which 'norms' (typical scores for people in that context) have been established. In practice this leaves very few options for appropriate use of such measures. The Child Behavior Checklist[65] is one of the few measures that meet these criteria. Some widely used assessments — such as the Hopkins Symptom Checklist and the Harvard Trauma Questionnaire — are set up to measure mental disorder[66] and are generally not appropriate means of assessing broader psychosocial needs.[67]

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. Although it takes some time to think through, sampling potentially reduces the time and resources required for an evaluation by allowing you to select information from a representative subset of participants in a programme, rather than having to use data from all or most of the participants. 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.

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.

For quantitative research methods — discussed below — there are mathematical formulae to work out the size of the sample required.[68] 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 programs[69]

© UNICEF/NYHQ2009-0178/GIACOMO PIROZZI

In 2008 a survey was administered to a sample of approximately 1,900 Palestinian children and adolescents living the West Bank and Gaza by several agencies. Each agency used their program rosters as their sampling frame. In all but one case, random sampling strategies were used, and questionnaires were administered to at least 80-100 children in program groups and in comparison groups.

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

© UNICEF/NYHQ2011-0059/OLIVIER ASSELIN

In an evaluation of a classroom-based intervention in schools in Gulu District of N Uganda in 2008[64], 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.

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 both camps. 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 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.

THE THREE APPROACHES TO SAMPLING WITH QUANTITATIVE EVALUATION METHODS TABLE 6

RANDOM SAMPLING	getting a list of everyone, and choosing a way of selecting at random 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.

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.[70]

Quota Sampling. Random sampling or cluster sampling 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 collecting information from (such as people with disabilities, or children from certain minority ethnic groups).

However, as quota sampling is more likely to lead to a biased selection of participants, it is a statistical 'rule of thumb' that the sample size should be **doubled** with this selection method. For evaluation of any distinct population of beneficiaries, you are therefore looking at a sample of between 160 and 200. With this method you select your sample to fit in with a 'quota' of beneficiaries defined by certain criteria. For example, in a sample of 200 children in a school, you could decide to sample 50 young girls, 50 older girls, 50 young boys and 50 older boys.

If 5% of children in the school come from a particular ethnic minority with random or cluster sampling you would have been likely to recruit only 4 or 5 such children in a sample of 80-100.

With a quota sampling approach you can choose how many you wish to include: 10 or so would be a 'representative' number within a sample of 200, but if you are really interested in information about these children you could select perhaps 20 (an approach that is known as 'over-sampling').

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 a program. 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 collecting information 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 information 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 participated in a program. 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 experience will not be representative of others staying in more inaccessible camps.

© UNICEF/NYHQ2010-2782/OLIVIER ASSELIN

In Northern Uganda in 2006, CCF arranged focus groups to discuss issues around sexual violence against women and girls. To ensure a representative range of participants, the groups were convened at locations that had been selected at random from a listing of all 'blocks' within a refugee camp.

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 to identify others that fulfill certain criteria for inclusion (a particularly useful method when dealing with sensitive issues such as recruitment by armed groups or sexual violence).

10 ANALYSIS, REPORT WRITING AND SHARING FINDINGS

© UNICEF/NYHQ2010-1369/MARTA RAMONEDA

This section looks at analyzing information you have found in doing an evaluation, and provides a brief guide in report writing and sharing findings.

ANALYZING INFORMATION

There are generally three major steps in analyzing the information collected in an evaluation: collating or organizing the information; data analysis; and interpretation.

COLLATING OR ORGANIZING INFORMATION

Evaluations often generate a lot of data and it is important to organize material so that analysis can be conducted in a clear and structured way. The stated objectives of the program (and the indicators developed with respect to them) will be key in guiding your organization of material.

For quantitative information — collected through survey forms, questionnaires, existing records etc — it is usually best to collate all information onto a single spreadsheet using simple software such as Excel. Typically, this would involve a row of the spreadsheet representing each child (or family or community) featuring in the evaluation, with columns for each of the 'fields' (or variables) for which data was collected. It is important to check over the spreadsheet carefully once data has been entered, checking back with the original documents to verify entries as necessary (a process call 'data cleaning').

With qualitative data it is generally harder to reduce information so concisely before analysis. But there is still value in clearly organizing the material — for example, grouping materials from different communities or marking the age and gender of participants clearly on the material — so that analysis can proceed in a logical way and any missing information readily identified.

© UNICEF/NYHQ2009-1005/SHEHZAD NOORANI

'With quantitative data, at its most basic, we are looking for difference in the mean score of participants who haven't received an intervention with those who have'.

DATA ANALYSIS

Data analysis involves identifying trends or patterns in information collected through the evaluation. A main analytical tool is to examine **differences** – either differences between before the program and after it, or differences between those who received an intervention and those in some comparison group (and preferably both). With quantitative data, at its most basic, we are looking for differences in the mean (that is, average) score of participants who haven't received an intervention with those who have.

Analysis can get much more complicated than this, but we shouldn't lose sight of this as the core question. So a very first step has to be calculating the average scores for participants, and seeing if there appear to be important differences linked to whether people have received some psychosocial support.

It can be helpful to use Excel to plot out a basic bar graph showing this information. If the bars suggest very different levels of scoring, then it may be that the intervention has had an impact and the trend needs to be looked at more closely. If the bars are very much the same height, although we might pursue some sort of statistical analysis, we should not expect this to reveal something different from what is apparent – there is not much difference (on this particular issue, at least) between those who received psychosocial support and those who had not.

Amongst the most basic forms of statistical analysis, **bivariate** (that is, two variable) analysis provides a means of judging whether one variable (e.g. attending or not attending a youth empowerment program) seems to have influence on another (e.g. self-reported well-being). The concept of statistical significance reflects the chances that the observed trend could have arisen purely by chance (and thus not providing a sound basis for concluding that there is some real substance in the connection that would justify extending coverage of the youth empowerment work to confidently improve self-reported well-being).

Simple forms of bivariate analysis, such as the Mann-Whitney U-test[71] and the t-test[72], are on the syllabus of most social science degree programs, so if you don't have the confidence to do these yourself, you should readily be able to find a colleague that can assist.

© UNICEF/NYHQ2008-1405/IYAD EL BABA

Qualitative analysis can provide rich insight into people's experience. A survey may indicate that 70% of children agreed with the statement 'I feel sad every morning'. We understand how to interpret this figure much better when we observe through qualitative analysis a pattern of many statements like this: "When I wake in the morning I am happy to see the sun, but then I remember that my mother has died and I feel very sad. I feel like putting my head under the blanket and never coming out. I want to stay there all day."

More complex **multivariate** (that is, multiple variable) analysis provides a means of considering whether other factors (such as whether trainees were girls or boys, or older or younger, or were from male – of female – headed households) influenced the relationship between receiving the training and its Impact on well-being. These questions are often very interesting, but to look at them rigorously will usually require a larger sample size than discussed earlier (calculations for which are based on a fairly simple bivariate analysis) and, more generally, more advanced statistical knowledge. If such technical support is available to an evaluation team, that is very helpful. Indeed, if this is the case, such technical expertise should ideally be engaged in the design of the evaluation from the outset.

The analysis of **qualitative data** builds on the same general principles of the identification of patterns and trends. Qualitative data can suggest why a trend noted in quantitative analysis occurred, or give insight into what a pattern of responses means. For instance, survey analysis might indicate that girls who were associated with armed forces are more likely to drop out of school than boys who were also recruited. Qualitative analysis may then suggest why this occurred, e.g. it is because they have babies, their families prioritize boys' education, and they are seen as 'dirty' and are ashamed to go to school.

Analysis of qualitative information provides rather different challenges to analyzing quantitative information. Usually, the key focus is identifying issues that recurred across a number of participants and in a range of situations. Such **thematic analysis** means some structure has to be imposed on the data, using some form of **coding.**

This usually involves grouping statements together that share a similar theme. Decisions need to be made about whether to 'clump' many items together or to 'split' information across a larger number of themes. For most practical purposes in

planning and evaluating psychosocial provision, aiming for between three and five major organizing themes will usually be most helpful. 'Sub-themes' can be identified within these if necessary.

There are a number of practical manuals offering assistance in analysis of qualitative data.[73] Some encourage use of specialized computer software to help organize data and its analysis. Many others suggest that traditional methods of grouping statements (e.g. marking with colored pens, or cutting out statements from transcripts and placing in piles — or into plastic pots or trays — for sorting) is the best way to become familiar with your data.

Once data is grouped and coded in this way, you can make an analysis table, and compare your "before" and "after", or programme and control group notes, for each tool that you used. Look for patterns in the responses that suggest how things have changed, or how the circumstances of those participating in programming differ from those who did not.

Because dealing with large amounts of qualitative data can be very time-consuming, superficial analysis of such data (i.e. 'cherry-picking' or 'choosing a couple of good quotes') is quite common. Sufficient time needs to be given to the analysis of qualitative data if one is to respect the time that participants contributed in its collection.

Transcription and translation of interviews (whether digitally recorded or from written notes) consumes significant resources. Generally, allow two to three times more time to analyzing qualitative data than collecting it. That is, if it took two days to run some focus groups, allow between four and six days for analyzing the transcripts from such groups.

If it is not possible to find this amount of time for analysis, then you should consider using qualitative methods that lend themselves to more rapid identification of themes such as participative ranking methodology and other mapping methods. These methods involve identifying key priorities for participants in the course of data collection itself, and analysis involves consolidating responses across groups, rather than beginning 'with a blank sheet'.

Whether quantitative or qualitative data is being considered, seeing the different patterns of response across different groups will often be an important consideration during analysis, as this has potentially major implications for programming.

One of the most significant differences to track is that shaped by gender. The box below provides some guidance on this issue of gender analysis, which involves thinking through the way evaluation is conducted as well as how data itself is then analyzed.

TABLE 7

GUIDANCE ON GENDER ANALYSIS	Gender analysis examines the different experiences of males and females. It examines their roles, their access to and control of resources, and the constraints they face relative to each other. Assessments need to focus on gaining an understanding of the possible differences in experience for women and girls, boys and men.
GENDER FOCUS	Focus on women, girls, boys, and men as the centre of your assessment: start with the smallest unit – the household and its dynamics. Understand how each family member participates, what role they play, and what they need in order to improve their well-being, security and dignity. Ask how services are accessed. Who obtains resources and decides how they are used?
UNDERSTAND THE CULTURAL CONTEXT	Who is head of the household? The ability of a family or household to access services will differ depending on the situation of the head of household. Adolescent boys, widowed women, single mothers and the elderly may all have problems accessing services and support during emergencies. Who asserts control? Analysis of relations and roles can help identify vulnerabilities and solutions.
COORDINATE AND COOPERATE	Field workers in all sectors and areas of activity need to assess whether and how the situation affects women and men differently. Field workers must ascertain how their programs will address the immediate practical and longer-term strategic needs of women and men. This information needs to be shared with other humanitarian actors to ensure that programs are well coordinated. Local networks and organizations – such as local NGO's, informal youth or women's networks, male working groups and community groups for the elderly – are a useful resource base for assessments.
DON'T MAKE ASSUMPTIONS	Gender analysis helps explain the different ways women, girls, boys, men and the elderly are affected by, or participate in, the political, economic, social and cultural decision-making in a society. Awareness of who is making the decisions helps to ensure a more accurate understanding of the situation and the varying needs of different groups affected by the crisis.

Reference: IASC Guidelines: Women, Girls, Boys and Men: Different Needs, Equal Opportunities, Gender Handbook in Humanitarian Action, December (2006).

INTERPRETATION

The main task of interpretation in evaluations is deciding whether analysis indicates that programming has achieved targeted outcomes, whether these have resulted in meaningful impacts for beneficiary communities and what factors have helped (or hindered) these achievements. If indicators suggesting targeted change have been clearly specified, and if a strong evaluation design has been used, then this task will be very straightforward. Are things better for people than before, and have those that have received support done better than those who haven't?

In many situations, though, things may not be as straightforward as this. Data may suggest that things have improved for some groups but not others. Or perhaps there is no evidence of the intervention having made a difference overall, but there is a suggestion it may have been helpful for some (e.g. younger girls). Or families may be very positive about their experience of the intervention, but other measures suggest they are doing no better than those who received no support.

For most agencies there is, in such circumstances, a natural inclination (given the need to demonstrate outcomes and impact to funders) to focus on positive findings (i.e. evidence of success) and to 'play down' negative ones (i.e. those suggesting no change, or even potential harm). This whole guide is based on the argument, however, that we need a stronger evidence base to learn about what works and what doesn't. That means being honest about lack of progress and disappointment.

Program evaluations can contribute a lot by indicating why progress was not as good as planned. If progress was good for some groups of participants but not others, what may explain this? The answer to this question could lead to significant improvements in programming in the future. If most indica-

tors suggest poor program impact, but beneficiaries report very positively on the program, were appropriate indicators selected? Again, acknowledging disappointing findings can lead to more informed thinking about programming and its aims.

REPORT WRITING AND SHARING FINDINGS

The conclusions of an evaluation have implications for those involved with delivering the intervention, the beneficiaries that the program has aimed to serve and potentially many other stakeholders. It is obviously good practice to share findings with all of these groups. However, it is better still that — before final reports are circulated — preliminary analysis and conclusions are shared with such stakeholders such that they can assist in shaping recommendations.

Recommendations should flow logically from analysis. If an Intervention has been successful, the focus should be on potential means of replicating or scaling-up the program. If it has been unsuccessful, or only partially successful, focus on potential strategies to strengthen the Intervention or adopt another approach. Recommendations can also focus on the need for further analysis or data. For instance, an inter-agency evaluation of psychosocial programs in oPt found programming generally effective in supporting children's psychosocial well-being, but indicated that further research will be required to understand which programming approaches are the most effective.

Recommendations are usefully grouped in terms of the people or organizations that are expected to act upon them (e.g. separating recommendations for the program intervention team, from those addressing the agency as a whole and, potentially, from government ministries or Inter-agency working groups working in that area).

REPORT WRITING

Standard guidance on writing evaluation reports is usually provided by organizations[74], and so the advice here should be considered as complementary to such information.

Because of the different people that will read your report it is helpful to think of three 'products' that you are generating, each of which should make sense on their own:

- a three page Executive Summary;
- the full report;
- a one page summary in plain, clear language.

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 suitable for sharing with both community and professional networks. There are a number of sources of guidance on making reports more accessible to readers, and thus more likely to impact practice and policy.[75]

It is important to include a comprehensive account of the evaluation methodology used in the report, including a full description of the tools used. This information is helpful for other evaluators and will contribute over time towards better understanding of psychosocial programming evaluation.

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 issues.[76] 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.' One of the most common problems of evaluations is that they are not widely disseminated. Resources, time and energy need to be dedicated to the dissemination of the findings and a clear plan developed for this. It is usually the case that findings are shared and discussed with donors; but it is important to remember accountability to other groups too, especially affected populations and program staff.

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?
- What program developments should be adopted given the findings?

When you look at your findings, you will need to consider[77]:

- **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. Remember that the passive distribution of written materials or presentations about findings with no opportunity for discussion is not likely to have much impact.
- **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. Using workshops to discuss findings or arranging a launch party or showing a DVD featuring main findings is more likely to engage attention.
- **the user**
  target audiences may have different needs and — as noted above — 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.'

In a working paper about child and older carer participation, the writers note that 'children and adolescents are calling for action and to be involved. (However) child participation in public fora and ongoing and structural national, community and district child participation still remains a gap in many countries'.[78] There are remarkably few examples of children being effectively engaged in the sharing of findings about psychosocial programming targeting their needs. Evaluators could usefully share their experience in seeking to more effectively engage children and youth — and other groups commonly marginalized from such activities — in the discussion of evaluation findings.

Longer term benefits can be gained by developing interaction and ongoing links with practitioners, NGOs and policy makers through informal and formal partnerships. Networks such as the Mental Health & Psychosocial Support Network[79] and the CPC Learning Network[80] provide useful mechanisms to share findings and approaches.

Crucially, reviewing findings should be linked with planning a response. One way to do this is to develop an action plan with program management that outlines steps to be taken in response to the recommendations of an evaluation and then monitors implementation of that action plan.[81]

11 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 psychosocial support influences the lives and experiences of individuals, families and communities:

- **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 (often captured in a project 'results framework'). Evaluation should, whenever possible, be fully integrated into the design, monitoring and management of programs with respect to such objectives.

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 seek to feature baseline measures. A valid and ethical approach to comparison of those receiving programming support with those not receiving such support helps establish whether observed changes can be attributed to programming.

8. As a minimum, all evaluations should feature the participation of individuals, groups or communities in:

- determining objectives and indicators for the project
- determining local definitions of well-being
- 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.

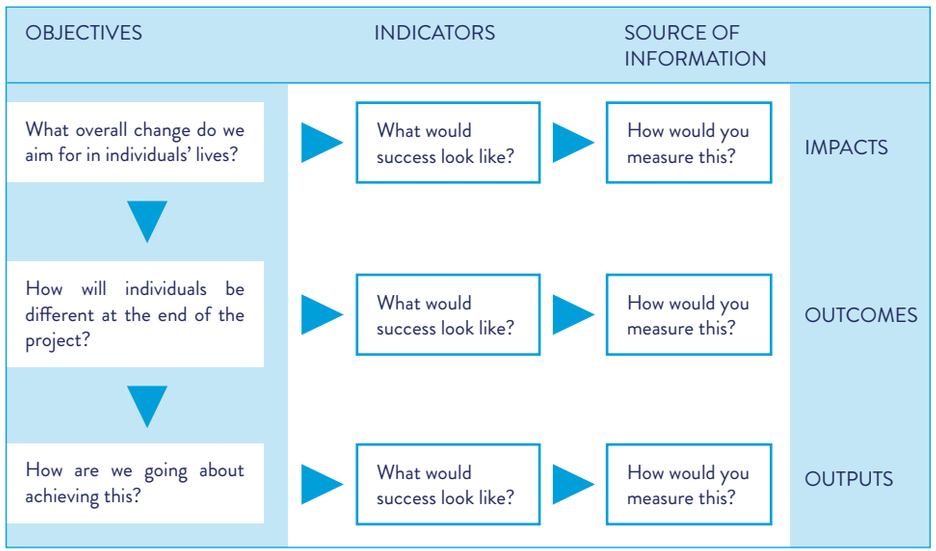
10. Rigorous evaluations can generally be conducted with a sample of between 80 and 100 participants (with a comparison group of a similar number).

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 circumstances.

12. Evaluation reports should be accessible and suitable for a variety of audiences. We encourage wide dissemination, including the preparation of one page summaries of major findings, written in plain language. Reports should provide detailed descriptions of methodology used, including tools and measures adopted.

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

ANNEXE A
A GUIDE TO THE DEVELOPMENT OF INDICATORS



Consider desired changes in families and communities also

At minimum, specify culturally relevant indicators of
(1) skills & knowledge,
(2) emotional well-being and
(3) social well-being

ANNEXE B
A STEP-BY-STEP GUIDE TO CONDUCTING AN EVALUATION

STEP	TASK	RELEVANT SECTION	COMMENTS
1	Engage Beneficiaries and Staff in Evaluation Planning	2, 3, 4, 5, 6 & 8	Review scope of evaluation and resources available/ required for its completion.
2	Confirm/Develop Program Objectives and Indicators	3, 5, 6 & 8	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	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	Use indicators to identify required methods of quantitative and qualitative data collection.
6	Plan Baseline Assessment (or 'Reconstruct' Baseline, if required)	7 & 9	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	Arrange focused, supervised program of collecting information from beneficiary and comparison settings;
10	Analyze and Share Findings	10	Ensure prompt analysis of information. Ensure multiple versions of findings are prepared for range of stakeholders.
11	Plan Action	10	Support action planning based on findings.

In*Steps are indicated in sequence, but earlier steps will often need to be 'revisited' subsequent to later decisions. Evaluation should be integrated within the 'design, monitoring & evaluation' project cycle – and linked to agreed results frameworks – as much as possible.

ANNEXE C
KEY RESPONSIBILITIES FOR PROGRAM EVALUATION

The responsibilities for protection and ethical working relationships lie with program staff and the evaluators. Managers commissioning evaluation activities are responsible for ensuring that protection issues are identified and resolved in methodology design, staff training and supervision.

KEY RESPONSIBILITIES AT THE TIME OF PROGRAM PLANNING (PRIOR TO IMPLEMENTATION):

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

KEY RESPONSIBILITIES WITHIN THE YEARLY CYCLE OF PROGRAM PLANNING:

- Organizations complete regular monitoring and evaluation including process and mid-term monitoring and evaluation;
- 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 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;
- Where children are involved, have a 'child-friendly lens' in the process of drafting the evaluation terms of reference, selecting the evaluation team and developing the proposed methodology;
- Provide training to ensure staff have skills & knowledge to conduct specified means of data collection (as well as having the skills to cope with distress if it occurs);
- Continue — through 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 individuals, families and communities.

ANNEXE 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).

ANNEXE E
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.[82] The second example is an evaluation of a program supporting the reintegration of children formally associated with fighting forces in Mozambique.[83]

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 Children. 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 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 here.

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	• Improved emotional well-being • Improved social well-being	• Child Behaviour Checklist (Palestinian Version)
	• Stronger parent-child relationships	• Children report closer relationship with parents	• Parental Checklist
	• Enhanced capacity of local community organizations for work with children	• Organizations develop wider range of activities utilizing skills acquired	• Interviews with organizations • Organizational records
OUTCOMES	• Children engaged in play in a safe environment	• High proportion of local children regularly attended safe play area • Children active In play	• Program records • Site visits
	• Children participate actively in cultural activities	• Parents report positive engagement by children • Children report active engagement with activities	• Site visits • Focus groups with parents • Participative activities with children
OUTPUTS	• Safe play areas established and equipped	• Number of safe play areas built or refurbished • Equipment delivered to play areas	• Program records • Site inspections
	• Training conducted in delivery of cultural activities	• Number of trainings offered in facilitating cultural activities • Number of persons trained	• 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. 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. 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 psycho-social 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.

GENERAL GUIDANCE ON PROGRAMMING AND EVALUATION:

ACT Alliance website on psychosocial issues addressing issues of spirituality.
Retrieved from http://psychosocial.actalliance.org

Active Learning Network for Accountability and Performance in Humanitarian Action (ALNAP): Chapter 6 'Monitoring' and Chapter 7 'Evaluation'. (2003). In Participation by Crisis-Affected Populations in Humanitarian Action: A Handbook for Practitioners (pp. 193-227). Retrieved from http://www.globalstudyparticipation.org/index.htm

Action for the Rights of Children Resource Pack CD-Rom (ARC). (2009). Retrieved from http://www.arc-online.org.

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

Horizons Program/Population Council and Family Health. (2005). Ethical Approaches to gathering information from children and adolescents in international settings – guidelines and resources. Retrieved from http://www.popcouncil.org/horizons/childrenethics.html

Inter-Agency Standing Committee (IASC). (2007). IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Geneva: IASC. Retrieved from http://www.humanitarianinfo.org/iasc/content/products/docs/IASC%20MHPSS%20guidelines%20Feb%2025%202007.pdf

IASC resources in relation to action sheets.
Retrieved from http://www.who.int/mental_health/emergencies/mh_key_res/en/index.html

International Institute for Environment and Development (IIED). (2005). Participatory Learning and Action (PLA Notes). See specifically notes 31 and 42, which cover PRA monitoring and evaluation. Retrieved from http://www.iied.org/NR/agbioliv/pla_notes/backissues.html

REPSSI. (2009). Mainstreaming Psychosocial Care and Support through Child Participation. Copyright © REPSSI. Retrieved from http://www.repssi.net/index

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

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

Perez-Sales, P. (2006). Repensar Experiencas. Evaluación de programas psicosociales y de salud mental. Metodologías y técnicas. (Ed Popular). (Rethinking experiences. Assessment of mental health and psychosocial programs: Methods and Techniques). Retrieved from http://www.psicosocial.net

Sphere Project. (2004). Humanitarian Charter and Minimum Standards in Disaster Response. Retrieved from 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. Retrieved from http://www.crin.org/docs/Handbook%20new%20update.pdf

UNICEF. (2005). Understanding Results-Based Programme Planning and Management. UNICEF: New York.

UNICEF. (2006). Child Protection Information Sheets. Retrieved from
http://www.unicef.org/protection/files/Child_Protection_Information_Sheets_(Booklet).pdf

GUIDES ON SPECIFIC METHODOLOGIES:

Action Aid International. Participatory Vulnerability Analysis: A step-by-step guide for field staff.
Retrieved from http://www.actionaid.org.uk/wps/content/documents/PVA%20final.pdf

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

Silverman, D. (2010). Doing Qualitative Research: A Practical Handbook (3rd ed). Sage: London.

ASSESSMENT TOOLS:

Achenbach, T. M. (2001). Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, Vermont.
Retrieved from http://www.aseba.org/aboutus/laterdevelopments.html

Bragin, M. (2005). The community participatory evaluation tool for psychosocial
programmes: A guide to implementation. Intervention: International Journal of Mental Health, Psychosocial Work & Counseling in Areas of Armed Conflict. 3, 3-24.
Retrieved from http://www.interventionjournal.com/downloads/31pdf/03_24%20Bragin%20.pdf

Hopkins Symptom Checklist. Retrieved from http://www.hprt-cambridge.org/Layer3.asp?page_id=10

Impact of Events Scale: Ioanna Giannopoulou et al. (2006) Factor structure of the Children's Revised Impact of Event Scale (CRIES) with children exposed to earthquake. Personality & Individual Differences. 40 (5), 1027-1037.
Retrieved from http://www.childtrauma.com/chmies8.html

O'Donnell, K., Nyangara, F., Murphy, R. & Nyberg, B. (2007). Child Status Index: A Tool for Monitoring and Evaluation of Orphans and Vulnerable Children programs: User's Guide. Measure Evaluation and Duke University.
Retrieved from http://www.hciproject.org/node/1237

REPSSI's Information and Action Tool. Retrieved from www.repssi.org

REPSSI (2010) Are we Making a Difference? Participatory Evaluation Tools For Monitoring and Measuring the Impact of Psychosocial Support Programmes For Children Aged 6 To 18. (2nd ed).
Retrieved from http://www.repssi.org

Senefeld, S., Strasser, S. & Campbell, J. (2009). Orphans and Vulnerable Children Wellbeing Tool: User's Guide. Catholic Relief Services: Baltimore. Retrieved from www.crs.org/publications/ovc-wellbeing-tool

OUTCOME MAPPING:

Retrieved from http://www.outcomemapping.ca/ and http://www.idrc.ca/en/ev-26586-201-1-DO_TOPIC.html

Jones, H. (2006). Making Outcome Mapping Work: Evolving Experiences from Around the World. Ottawa: IDRC. Retrieved from http://bit.ly/19zzsH

KEY PUBLISHED STUDIES:

Ager, A., Stark, L., Olsen, J., Wessells, M. & Boothby, N. (2010). Sealing the Past, Facing the Future: An Evaluation of a Program to Support the Reintegration of Girls and Young Women Formerly Associated with Armed Groups and Forces in Sierra Leone. Girlhood Studies, 3 (1, Summer), 70-93.

Bolton, P. & Tang, A.M. (2002). An alternative approach to cross-cultural function assessment. Social Psychiatry and Psychiatric Epidemiology, 37, 537-543.
Retrieved from http://www.who.int/mental_health/emergencies/2.2_key_resource_4_bolton_article.pdf

Loughry, M., Ager, A., Flouri, E., Khamis, V., Afana, A. H. & Quota, S. (2006). The impact of structured activities among Palestinian children in a time of conflict, Journal of Child Psychology & Psychiatry, 47 (12), 1211-18.

Jordans, M.J.D., Komproe, I., Tol, W., Kort, B., Luitel, N., Macy, R. & de Jong, T. (2010). Evaluation of a classroom-based psychosocial intervention in conflict-affected Nepal: a cluster randomized controlled trial. Journal of Child Psychology and Psychiatry. 51 (7), 818–826

Tol, W., Komproe, Dessy Susanty, D. et al. (2008) Randomized Trial Affected by Political Violence in Indonesia: A Cluster School-Based Mental Health Intervention for Children. JAMA. 300(6), 655-662.

REFERENCES

[1] 1 UNICEF. (2006). Assessment and Evaluation of Psychosocial Programming for Crisis-affected Children. Prepared by the Program on Forced Migration and Health, Mailman School of Public Health. Columbia University, USA.

[2] IFRC Reference Centre for Psychosocial Support. (2009). Psychosocial Interventions: A Handbook. (pp.175).

[3] For an introduction to the mental health and psychosocial impact of emergencies, please see pages 1 – 13 of the Inter-Agency Standing Committee (IASC). (2007). IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Geneva: IASC.

[4] UNICEF. 2005. Understanding Results-Based Programme Planning and Management:
Tools to Reinforce Good Programming Practice.

[5] World Vision International (2007) LEAP: Learning through Evaluation with Accountability & Planning. World Vision's Approach to Design, Monitoring and Evaluation (2nd ed). World Vision International.

[6] Section7 of this guide also looks at these issues in more detail. Bamberger, Michael et al. (2006). Real World Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage, Thousand Islands, California.

[7] Impact evaluation consultants require expertise in research design and quantitative and qualitative methods, as well as experience in humanitarian crises.

[8] Introduction to Child Protection in Emergencies: an Interagency Modular Training Package - Psychosocial Care and Support Module. (2007). Christian Children's Fund, ICRC, IRC, Save the Children, Terre des Hommes; UNICEF and UNHCR.

[9] Action for the Rights of Children Resource Pack CD-Rom (ARC). (2009).
Retrieved from http://www.arc-online.org.

[10] Inter-Agency Standing Committee (IASC). (2007). IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Geneva: IASC.

[11] Psychosocial Working Group. (2005). Psychosocial Intervention in Complex Emergencies: A Framework for Practice. Retrieved from http://www.forcedmigration.org/psychosocial/papers/A%20Framework%20for%20Practice.pdf

[12] REPSSI (2010) Are we Making a Difference? Participatory Evaluation Tools For Monitoring and Measuring the Impact of Psychosocial Support Programmes For Children Aged 6 To 18 (2nd ed).
Retrieved from http://www.repssi.org

[13] Stavrou, Vivi (2007). Review of UNICEF Supported Right to Play Interventions in Responding to the Psychosocial Needs of Children Affected by Conflict in Northern and Eastern Uganda. UNICEF Uganda.

[14] Onyango, G.R. (2005) A Randomized Controlled Trial of Group Interpersonal Psychotherapy and Creative Play for Psychosocial Problems in War-Affected Acholi Adolescents in N.Uganda (PPT presentation). World Vision, Uganda.

[15] Psychosocial Working Group. (2005). Psychosocial Intervention in Complex Emergencies: A Framework for Practice. Retrieved from http://www.forcedmigration.org/psychosocial/papers/A%20Framework%20for%20Practice.pdf

[16] http://psychosocial.actalliance.org/default.aspx?di=65871&ptid=66401

[17] Articles 27 and 38 of the 4th Geneva Convention, In, The Geneva Conventions of 1949 and their Additional Protocols. Retrieved from http://www.icrc.org/web/eng/siteeng0.nsf/html/genevaconventions

[18] Inter-Agency Standing Committee (IASC). (2007). IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Geneva: IASC.

[19] Inter-Agency Standing Committee (IASC). (2008). Inter-Agency Child Protection in Emergencies Training and Resource CD Rom, UNICEF, New York.

[20] Inter-Agency Standing Committee (IASC). (2007). IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Geneva: IASC. (pp. 9- 11).

[21] Office of the United Nations High Commissioner for Human Rights. (1989). Convention on the Rights of the Child. Geneva, Switzerland, 1989

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

[23] Inter-Agency Standing Committee (IASC). (2007). IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Geneva: IASC. (pp. 9 – 11).

[24] See action sheet 2.1 for details on what kinds of information should be Collected. In, Inter-Agency Standing Committee (IASC). (2007). IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Geneva: IASC. (pp. 46 – 49).

[25] Inter-Agency Standing Committee (IASC). (2008). Inter-Agency Child Protection in Emergencies Training and Resource CD Rom, UNICEF, New York. (pp. 21 – 22).

[26] Interventions may sometimes be part of transformational processes, as explained on page 11. Not all beliefs and practices are free of harm or discrimination.

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

[28] REPSSI (2010) Are we Making a Difference? Participatory Evaluation Tools For Monitoring and Measuring the Impact of Psychosocial Support Programmes For Children Aged 6 To 18. (2nd ed). (pp. 14).
Retrieved from http://www.repssi.org

[29] Overseas Development Institute (ODI). (2006). Evaluating Humanitarian Action Using the OECD-DAC Criteria: A Guide for Humanitarian Agencies. ODI, London. Retrieved from http://www.alnap.org/pool/files/eha_2006.pdf

[30] Guidance on refining monitoring and evaluation (M&E) indicators for Child Protection, Education, Health & HIV, and Hunger Reduction programmes. In, World Vision International (2007) LEAP: Learning through Evaluation with Accountability & Planning. World Vision's Approach to Design, Monitoring and Evaluation (2nd ed). World Vision International.

[31] UNICEF. (2006). Assessment and Evaluation of Psychosocial Programming for Crisis-affected Children. Prepared by the Program on Forced Migration and Health, Mailman School of Public Health. Columbia University, USA.

[32] Jones, H. & Hearn, S. (2009). Outcome Mapping: a realistic alternative for planning, monitoring and evaluation. ODI Background Note, ODI, London.

[33] Jones, H. (2006). Making Outcome Mapping Work: Evolving Experiences from Around the World. IDRC, Ottawa. Retrieved from http://bit.ly/19zzsH

[34] Ambrose, K & Roduner, D. (2009). A conceptual fusion of the logical framework approach and outcome mapping. OM Ideas Paper No. 1 2009. Retrieved from http://bit.ly/WfOtn

[35] See Section 3 for a discussion of principles that should inform such participative working.

[36] Bamberger, Michael et al. (2006). Real World Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage, Thousand Islands, California.; Duncan, J. & Arntson,L. (2004). Children in Crisis: Good Practices in Evaluating Psychosocial Programming. Save the Children Federation.

[37] UNICEF. (2006). Assessment and Evaluation of Psychosocial Programming for Crisis-affected Children. Prepared by the Program on Forced Migration and Health, Mailman School of Public Health. Columbia University, USA.

[38] Columbia University & UNICEF/oPt. (2010). OPT Survey: Interagency Psychosocial Evaluation Project, Final Report.

[39] Bolton P, Bass J, Betancourt T, Speelman L, Neugebauer R, Murray L, Onyango G, Clougherty K, Verdeli H. (2007). Interventions for Depressions Symptoms Among Adolescent Survivors of War and Displacement in Northern Uganda: A Randomized Controlled Trial. JAMA, 298 (5), 519 – 527.

[40] Bamberger, Michael et al. (2006). Real World Evaluation: Working Under Budget, Time, Data and Political Constraints, Sage, Thousand Islands, California. (pp. 92 – 93).

[41] See consideration of broader diversity issues in assessment section of the Inter-Agency Standing Committee (IASC). (2007). IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Geneva: IASC.

[42] IASC Guidelines: Women, Girls, Boys and Men: Different Needs, Equal Opportunities, Gender Handbook in Humanitarian Action. (2006, December).

[43] UNICEF. (2003). Technical Notes: Special Considerations for Programming in Unstable Situations, UNICEF. (pp. 412).

[44] Armstrong, M., Boyden, J., Galappatti, A. &Hart, J.(2004).Piloting Methodsfor the Evaluation of Psychosocial Programme Impact in Eastern Sri Lanka, Final Report for USAID. (pp. 22).
Retrieved from http://www.rsc.ox.ac.uk/PDFs/rrpilotingmethods04.pdfOxford

[45] REPSSI (2010) Are we Making a Difference? Participatory Evaluation Tools For Monitoring and Measuring the Impact of Psychosocial Support Programmes For Children Aged 6 To 18. (2nd ed).
Retrieved from http://www.repssi.org

[46] Arnstein, S. (1969, July) A Ladder of Citizen Participation. JAIP, 35 (4), 216 – 224.

[47] Holte-McKenzie, M., Forde, S. &Theobold, S. (2006) Development of a participatory monitoring and evaluation strategy. Evaluation and Program Planning, 29, 365 - 376

[48] Available at: www.popcouncil.org/horizons/childrenethics.html

[49] De Berry, J. et al. (2003) The Children of Kabul: Discussions with Afghan Families. Save the Children USA and UNICEF, (pp 3-4); and Armstrong, M., Boyden, J., Galappatti, A. &Hart, J.(2004). Piloting Methodsfor the Evaluation of Psychosocial Programme Impact in Eastern Sri Lanka, Final Report for USAID. (pp. 17).

[50] 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.

[51] REPSSI & UNICEF (2011).The Psychosocial Support Monitoring and Evaluation Technical Training Kit. (in publication).

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

[53] Oxfam. (2007). Impact Measurement and Accountability in Emergencies: The Good Enough Guide. Oxfam, Oxfam.

[54] Armstrong, M., Boyden, J., Galappatti, A. &Hart, J.(2004).Piloting Methodsfor the Evaluation of Psychosocial Programme Impact in Eastern Sri Lanka, Final Report for USAID. (pp. 50).
Retrieved from http://www.rsc.ox.ac.uk/PDFs/rrpilotingmethods04.pdfOxford

[55] Ager, A, Akesson, B, Stark, L, Okot, B, McCollister, F, Flouri, E. & Boothby, N. (in review) The impact of the school-based Psychosocial Structured Activities (PSSA) program on conflict-affected children in northern Uganda.

[56] Armstrong, M., Boyden, J., Galappatti, A. &Hart, J.(2004).Piloting Methodsfor the Evaluation of Psychosocial Programme Impact in Eastern Sri Lanka, Final Report for USAID.; Hart, J., Galappatti, A., Boyden, J., & Armstrong, M. (2007). Participatory tools for evaluating psychosocial work with children in areas of armed conflict: a pilot in eastern Sri Lanka. Intervention, 5(1), 41-60.

[57] Christian Children's Fund. (2005). Assessing Afghan Children's Psychosocial Well-Being: A Multi-Modal Study of Intervention Outcomes. CCF, Oxford University and Queen Margaret University College, Edinburgh.

[58] Stavrou, Vivi (2007). Review of UNICEF Supported Right to Play Interventions in Responding to the Psychosocial Needs of Children Affected by Conflict in Northern and Eastern Uganda. UNICEF Uganda.

[59] Available: www.rsc.ox.ac.uk/PDFs/rrpilotingmethods04.pdf

[60] REPSSI (2010) Are we Making a Difference? Participatory Evaluation Tools For Monitoring and Measuring the Impact of Psychosocial Support Programmes For Children Aged 6 To 18. (2nd ed). (pp 19).

[61] Hart, J., Galappatti, A., Boyden, J., and Armstrong, M. (2007). Participatory tools for evaluating psychosocial work with children in areas of armed conflict: a pilot in eastern Sri Lanka. Intervention, 5(1), 41-60.

[62] Stavrou, Vivi (2007). Review of UNICEF Supported Right to Play Interventions in Responding to the Psychosocial Needs of Children Affected by Conflict in Northern and Eastern Uganda. UNICEF Uganda.; Hart, J., Galappatti, A., Boyden, J., and Armstrong, M. (2007). Participatory tools for evaluating psychosocial work with children in areas of armed conflict: a pilot in eastern Sri Lanka. Intervention, 5(1), 41-60.

[64] Ager, A, Akesson, B, Stark, L, Okot, B, McCollister, F, Flouri, E. & Boothby, N. (in review) The impact of the school-based Psychosocial Structured Activities (PSSA) program on conflict-affected children in northern Uganda.

[65] Achenbach, T. M. (2001). Manual for the Child Behavior Checklist/4-18 and 1991 Profile. Burlington, Vermont.

[66] Bolton,P., Wilk,C. & Ndogoni,L. (2004). Assessment of depression prevalence in rural Uganda using symptom and function criteria. Social Psychiatry Psychiatr Epidemiol. 39 (6), 442-7

[67] Additionally, in humanitarian settings, without extensive local validation such measures may be unable to distinguish disorder from severe non-disordered distress.

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

[69] 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).

[70] The Food and Nutrition Technical Assistance (FANTA) Sampling Guide (1997). Retrieved from www.fantaproject.org/publications/sampling/shtml

[71] See http://en.wikipedia.org/wiki/Mann%E2%80%93Whitney_U or http://statpages.org/ or http://faculty.vassar.edu/lowry/utest.html

[72] See http://studentsttest.com/

[73] Silverman, D. (2010). Doing Qualitative Research: A Practical Handbook. (3rd ed).Sage: London.

[74] Monitoring and Evaluation Quick Reference: Extracts from the Programme Policy and Procedure Manual, (2005, May). UNICEF Evaluation Report Standards. Retrieved from www.unicef.org

[75] See, for example, the 'Research Matters' website and associated links at:
http://www.idrc.ca/en/ev-54056-201-1-DO_TOPIC.html

[76] Inter-Agency Standing Committee (IASC). (2007). IASC Guidelines on Mental Health and Psychosocial Support in Emergency Settings. Geneva: IASC. (pp. 43, 48).

[77] Drawn from notes from Research Unit for Research Utilization, St Andrews and University of Edinburgh (2005).

[78] Ward,N. & Stavrou,V. (2009, May) Child and Older Carer Participation in Events: Are We Really Learning? Regional Inter-agency Task Team (RIATT) on Children and AIDS Working Paper, RIATT, Johannesburg.

[79] www.psychosocialnetwork.net

[80] www.cpclearningnetwork.org

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

[82] Evaluation of the psychosocial intervention for Palestinian Children undertaken by the CRS and the Save the Children Federation by Loughry, M & Ager, A (2004). On behalf of the Gates Foundation. Refugee Studies Centre, University of Oxford.

[83] Boothby, N, Crawford, J, and Halprin, J. (2006). The Life Outcomes of Former Mozambican Child Soldiers, Global Public Health. Routledge Publishing, 1 (January), 87-104.

Boxed text references

a. Jareg, E. (2005) Community-based Foster Homes in Ethiopia: An Account of a Follow-up Experience Ten Years After Phase Out. Save the Children Norway.

b. Annan, J., Blattman, C. & Horton, R. (2006). The State of Youth and Youth Protection in Northern Uganda. SWAY: AVSI & UNICEF.
Retrieved from http://chrisblattman.com/documents/policy/sway/SWAY.Phase1.FinalReport.pdf

c. MacMullin, C. & Loughry, M. (2004). Investigating Psychosocial Adjustment of Former Child Soldiers in Sierra Leone and Uganda. Journal of Refugee Studies, 17 (4), 460-472.

d. http://www.warchild.nl/ons-werk/publicaties/

e. Ager, A., Stark, L., Olsen, J., Wessells, M. & Boothby, N. (2010). Sealing the Past, Facing the Future: An Evaluation of a Program to Support the Reintegration of Girls and Young Women Formerly Associated with Armed Groups and Forces in Sierra Leone. Girlhood Studies, 3 (1, Summer), 70-93.

f. Moving the Goalposts, Kilifi. www.mtgk.org

g. Arafat, C. and Boothby, N. (2003). A Psychosocial Assessment of Palestinian Children. USAID

h. Arntson, L. (2002). Review of the Former Soviet Embassy Compound IDP Camp Psychosocial Support Activities Kabul, Afghanistan. Save the Children US.

i. Kostorova-Unkovska, L. & Pankovska, V. (1992). Children Hurt By War. Skopje: General Children's Consulate of the Republic of Macedonia.

j. Stavrou, Vivi (2007). Review of UNICEF Supported Right to Play Interventions in Responding to the Psychosocial Needs of Children Affected by Conflict in Northern and Eastern Uganda. UNICEF Uganda.


ACF
INTERNATIONAL


care


ChildFund
International


HealthNet TPO


International Federation
of Red Cross and Red Crescent Societies


International
Medical Corps


MÉDICOS
DO MUNDO


The International Federation
Reference Centre for
Psychosocial Support


Save the Children


Terre des hommes
aide à l'enfance Kinderhilfe
aiuto all'infanzia child relief
www.tdh.ch


unicef


WAR
child


WARTRAUMA
FOUNDATION


World Vision