Guidelines for evaluation of education projects in emergency situations

by

Dr. Eva Marion Johannessen

for

The Norwegian Refugee Council

March 2001
# Table of contents

1 INTRODUCTION .......................................................................................................................... 4

2 EDUCATION IN EMERGENCY SITUATIONS ............................................................................... 5

3 WHAT IS EVALUATION? .................................................................................................................. 7

4 WHY EVALUATE? .......................................................................................................................... 9

5 WHAT CAN BE EVALUATED? ......................................................................................................... 10

6 TERMS OF REFERENCE (TOR) .................................................................................................... 12

7 QUANTITATIVE AND QUALITATIVE ISSUES IN EDUCATION .................................................. 13

8 WHEN TO EVALUATE .................................................................................................................. 15

9 FACTORS THAT INFLUENCE EVALUATION .............................................................................. 16

9.1 LACK OF QUALITY AND OBJECTIVITY .................................................................................. 16

9.2 EXTERNAL OR INTERNAL EVALUATORS? ............................................................................. 16

9.3 THE TEAM OF EVALUATORS .................................................................................................. 17

10 METHODOLOGICAL ISSUES ..................................................................................................... 18

10.1 SAMPLING .................................................................................................................................. 18

10.2 METHODS .................................................................................................................................. 18

10.2.1 Reviewing existing information ............................................................................................ 19

10.2.2 Asking questions: Interviews and questionnaires ................................................................. 19

10.2.3 Group interviews .................................................................................................................. 20

10.2.4 Interviewing children ......................................................................................................... 21

10.2.5 Observation ........................................................................................................................ 22

10.2.6 Other methods .................................................................................................................... 23

10.2.7 Which research methods are good enough in evaluation of education projects in emergencies? 25

11 EVALUATION OF EDUCATION PROJECTS IN EMERGENCIES ............................................. 27

11.1 EDUCATION IN A CHRONICALLY UNSTABLE SITUATION: TEP (Teacher Emergency Package) IN ANGOLA ...................................................................................................... 27

11.2 EDUCATION IN ACUTE EMERGENCIES: EDUCATION FOR REFUGEE CHILDREN FROM KOSOVO IN MACEDONIA .................................................................................................. 28

11.3 THE INTERNATIONAL RESCUE COMMITTEE EDUCATION PROGRAM FOR REFUGEES IN GUINEA 1991-98: A best practice study in a "care and maintenance" phase ........................................................................ 29

11.4 IN SEARCH OF INDICATORS ................................................................................................... 29

11.4.1 Checklist ............................................................................................................................ 32

11.4.2 Guidelines for classroom observations .............................................................................. 35

12 WRITING THE REPORT .............................................................................................................. 37

13 EVALUATING THE REPORT ....................................................................................................... 40

14 EVALUATION STEP - BY - STEP ............................................................................................. 41
1 INTRODUCTION

These guidelines were commissioned by the Norwegian Refugee Council who wanted a model for evaluation of education projects in emergency situations. The idea was to elaborate a tool for project planning and development, assessments, internal and external evaluations.

When we did the evaluation of the Teacher Emergency program in Angola (Johannessen and Romeu 2000a), we reviewed the reports from the provinces involved in TEP searching for commonalities which could make it possible to draw comparisons between them. But they were all different and it was not possible to find common indicators. We also found serious flaws in many provinces regarding the statistics at provincial and district level, which again were fed into national level and gave a biased or incomplete picture of the project's status. In the end we decided to present a set of indicators from the four provinces visited. In some provinces simple units were missing, and the numbers varied from one report to another.

It is any organization's dream to make an unambiguous standard format for information gathering and evaluation reports. It is however not possible to find a format that suits all purposes. The indicators vary from one project to another depending on the objectives, and they have to be specified when the project is being planned not after it has been finished.

A problem in emergencies is that the situation by definition is unstable. This explains why it is difficult to present correct information. But that is not the only reason. Lessons learned from our evaluation in Angola are that the key to secure correct statistics is to train those who are engaged in the project on a daily basis, in our case the teachers, headmasters and supervisors, to collect the information needed. The training has to focus on the understanding of the task not only on techniques. Ambiguous cases and flaws have to be reported and explained.

The presentation is divided into chapters. Some of the chapters apply to evaluation and planning of education projects in general while some are specifically geared to emergencies. A checklist and guidelines for classroom observation have been developed and a list of questions related to the checklist is enclosed in annex I (the checklist and questions may also be relevant in other education projects).

The paper is based on experience with shorter evaluations, which is quite common in NGOs and other organizations (cf. Cracknell 2000).
2 EDUCATION IN EMERGENCY SITUATIONS

An emergency arises as a result of natural disasters or human-made crises like war and civil strife (UNICEF).

The term "complex emergencies" has become widespread. It denotes an emergency produced by armed conflicts and political instability and not by natural disasters (Hallam 1998).

An educational emergency means that a disaster or conflict whether due to war or natural disasters, "have destabilized, disorganised or destroyed the education system, and which require an integrated process of crisis and post-crisis-response" (UNESCO 1999 in Sinclair 2000 p. 5).

NRC's objective is to contribute to the protection of refugees and internally displaced people (IDP) in close cooperation with UNHCR. "Protection" is defined according to Human Rights conventions and standards, and education is a human right.

Apart from assisting the refugees and IDPs in their basic needs in an emergency, education is considered important to restore normal life routines and activities. Refugees and IDPs who are placed in a camp tend to become passive recipients of services and not active participants in establishing a new life. Taking part in restoring schools and engaging in their children's education are ways to get out of the passivity.

We tend to think that first of all people's urgent needs must be met in an emergency. Education is luxury in such situations and comes last. However, education has gradually become what is called the "fourth pillar" in humanitarian assistance seen as equally important as food, shelter and health. "Education provides opportunities for students, their families and communities to begin the trauma healing process, and to learn the skills and values needed for a more peaceful future and better governance at local and national levels" (Sinclair 2000 p. 3). NRC has contributed actively to this process and proposed already in 1998 to the Ministry of Foreign affairs that education should be the fourth component in humanitarian assistance (Midttun 2000a p. 4). NRC has a strategy for how to meet education emergencies rapidly based on the four areas: personnel, material, principles/procedures/tools and networking. Teachers and educators have been recruited in the NRC Emergency Standby Force and are offered training to assist in acute emergencies.

There are three different emergency situations: acute emergencies, chronic emergencies and return/rehabilitation.

An acute emergency is defined as "any situation in which the life or well-being of refugees and internally displaced persons will be threatened unless immediate and appropriate action is taken, and which demands an extraordinary response and exceptional measures". Examples of such situations are Kosovo, East Timor, Eritrea and Chechnya.

A chronic emergency is "A general situation where
• an acute emergency has become a coping or maintenance phase in exile or within IDP camps
• refugees or IDPs have in part been absorbed by the local population, but there is no solution to the basic conflict
• recent or dormant conflicts may erupt again, and where strategic support may ease tensions and encourage peaceful development and reconciliation”. Angola, Sierra Leone and Burundi are examples of a chronic situation.

The return/rehabilitation phase is characterized by the following:

• "The security situation has been improving or has been stable and acceptable for a while
• Spontaneous return has started
• UNHCR is planning organised return
• Host and/or home country authorities have decided that refugees and IDPs should return to their home areas and are closing camps and settlements” (Midttun 2000b, draft).

The situations call for different actions. They are not static and may change rapidly, from stable to unstable. Rapid action and flexibility are key words in emergencies. We shall take a closer look at core elements in education programs in such situations in chapter 11.
3 WHAT IS EVALUATION?

To evaluate means to judge the value of a piece of work or objects. It covers determining "the worth of" in quantitative terms as well as judging the same qualitatively. When we look up the term "evaluate" and synonyms (review, assess, appreciate, audit) in English dictionaries it is associated with: "critical examination", "consideration", "survey", "judgement."

There are basically two types of evaluations, **formative** which are directed to on-going projects and **summative**, which take place when the project is finished. For our purpose, monitoring is a type of formative evaluation which is useful. **Monitoring** is a continuous summing up, analysis and control of a program/process to find out whether it develops according to its objectives. Its purpose is to give hints as to corrective actions. (Lewin 1993, Dale 1998). Monitoring may be initiated by donor organizations or by the organization who runs the project/program.

A **summative evaluation** is a more thorough examination of the project/program. It normally takes place when the project/program is completed. Formal evaluations may also be commissioned when a project is still in progress.

Sida distinguishes in its manual on evaluation between
- **feasibility studies** that take place before a project has started to test an idea
- **appraisals**, that is a critical examination of a program or project
- **monitoring** which is defined as a continuous control and analysis to find out whether the project/program develops according to its objectives.
- **Evaluation**, which takes place when the project/program is terminated (Lewin 1993 p. 9)

OECD's expert group has presented the following definition of evaluation:

"An evaluation is an assessment, as systematic and objective as possible, of an on-going or completed project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. An evaluation should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process of both recipients and donors" (Lewin 1993).

The term "evaluation" is used to describe different activities that vary in scope and methodology. It would be more appropriate in most cases to call shorter evaluations for "appraisals", "reviews" or "assessments" and reserve the term "evaluation" for examinations that are more thorough and of longer duration. To provide "information that is credible and useful" one needs to use tools and procedures that are reliable and transparent. As we shall see later, this is not always the case in evaluations.

Evaluation is also a term that is used in research. Long-term evaluations have characteristics in common with scientific research. Research however, in our modern times, is no longer carried out by trained scientists only, but is used to describe the work of journalists, authors and artists, among others.
A main characteristic of scientific research is that it is **systematic**. The design is systematic, in the sense that it follows certain procedures to find an answer to a research question. The research questions guide the collection of data. Depending on the research methods used, the findings may be presented in figures or words (descriptive). Another main feature is that research is **analytic**. The findings must be analyzed in terms of how they relate to one another and/or by drawing upon relevant research and literature. Research is presented in a final written report following agreed-upon procedures. The methodological part in a research report is critical as the findings are judged according to the their methodological validity and reliability\(^1\). A research report may or may not have recommendations as to future actions.

Although evaluation is different from scientific research in many respects, we need to meet some of the same criteria of rigor doing evaluation. Some of these criteria will be considered in the following.

---

\(^1\) "Validity" refers to whether the results measure what they claim to measure and "reliability" to whether the procedure of measurement would yield the same results whenever it is carried out. Although the term "measurement" is not applicable in qualitative research, the concepts are still relevant (Kirk and Miller 1986).
4 WHY EVALUATE?

Evaluation may be initiated for various reasons. The official reasons for doing evaluations are to improve strategies and programs and to investigate how the money has been spent. As Cracknell points out, an evaluation may focus mainly on the accountability in order to find out whether aid works or how effective it is. There may be less interest in why it works or why it does not. Another motive for evaluation is that it is a learning process for the organization. It seems as if these two purposes are difficult to reconcile. A study of evaluation reports where the lesson-learnt purpose was focused, shows that they could not be used as evidence for accountability purposes, "although they certainly yielded useful lessons about many other aspects of development aid" (Cracknell 2000 p. 56).

Other unofficial reasons for initiating evaluations are to justify political decisions that have already been made. The politicians may have another agenda than the evaluators and those who order the evaluations. An evaluation may also be decided in order to postpone a difficult political decision, to justify an on-going project that has been criticized, to cover up for mistakes by asking for evaluations which only cover some aspects of a project, or in order to stop a project (cf.. Lewin 1993).

It is important to ask why one wants evaluation and then find the most appropriate way to do it. An evaluation is costly and time-consuming. Breakwell and Millward (1995) recommends analyzing the financial and psychological costs and benefits involved before deciding to undertake an evaluation. They give the following advice:
"do not start an evaluation until you have specified at least one question which stands a good chance of being answered"
"do not begin an evaluation until you are satisfied that its benefits will outweigh its costs" (p. 16-17).

In most cases evaluations are imposed by politicians or donor agencies. They need proof that it is or has been worth while spending money on the project and advice as to its continuation.
5 WHAT CAN BE EVALUATED?

It is common in many evaluations to focus on the OECD criteria: relevance and fulfillment of objectives, development efficiency, effectiveness, impact and sustainability. **Relevance** refers to whether the project addresses the needs of the beneficiaries, **efficiency** to how productively the resources have been used while **effectiveness** relates to whether the objectives of the project were achieved. **Impact** is the long-term effect while **sustainability** relates to the maintenance of the changes after the project or program has been terminated (Dale 1998, Lewin 1993, Samset 1993). Sida adds "**lessons learned**" to the criteria which refer to important general experience the evaluation has yielded. All or some of the criteria are commonly addressed in evaluations of education projects although the terms may be used differently.

Although the concepts seem relevant and reasonable, there are several problems associated with them. First of all, they need to be defined and specified in terms of identifiable activities when the project is being planned. This is often not the case. Secondly, the concepts "relevance", "impact" and "sustainability" may be difficult to investigate. The project may have been going on for a too short period and/or there is not enough time available to find answers to this type of questions. Another objection is that the OECD concepts are related to final products and not to description of the **processes** that have taken place. No projects/programs, whether in research or evaluations, develop neatly in a linear way. The description of the project and the steps we plan in the beginning are always transformed as they are being implemented. Major decisions regarding adjustments and turning points should be described and justified and be part of the evaluation.

Others have suggested to expand the OECD criteria by introducing concepts that are more useful in evaluation of emergencies. These criteria are:

- **Connectedness**, which means "to assure that activities of a short-term emergency nature are carried out in a context which takes longer-term and interconnected problems into account". This term is considered more appropriate than "sustainability" as many humanitarian actions are not designed to be sustainable.
- **Coverage**, i.e. the need "to reach major population groups facing life-threatening suffering wherever they are, providing them with assistance and protection proportionate to their need and devoid of extraneous political agendas". One has to assess how a program affects different groups who are included.
- **Coherence**, refers to the need "to assess political, diplomatic, economic and military policies as well as humanitarian policies, to ensure that there is consistency and, in particular, that all policies take into account humanitarian considerations" (Hallam 1998).

Two more concepts are introduced: timeliness and appropriateness. The timing is of course particularly relevant in emergencies and should be explicitly stated as an important criterion. ** Appropriateness** describes the need "to tailor humanitarian activities to local needs, increasing ownership, accountability and cost-effectiveness accordingly. The concept supplements the criteria of "relevance" (ibid.).

**Gender issues** must be taken into account. The OECD/DAC WID Expert Group and Evaluation group has elaborated a plan for how gender aspects should be included in evaluations (cf. Lewin 1993 appendix 6). Gender is particularly important in education since we know that in many developing countries there are far less girls than boys who start and
finish primary school and continue their schooling. Girls are subject to early marriage, sexual harassment and unwanted pregnancies which stop them from further studies. We also know that many teachers treat boys and girls differently in class and have higher expectations for the boys. They do not know how to encourage the participation of girls. Some teaching methods favor boys more than girls. In some countries and programs there are far less women than men in the teaching profession although we know that it may be easier for the girl to identify with a female than a male teacher. These aspects have to be taken into consideration when a project is planned and evaluated and have to be reflected in the indicators.

Breakwell and Millward (1995) says the following broad categories of things can be evaluated: activities, personnel and provision of resources, organizational structure, objectives. When we evaluated the TEP program in Angola (Johannessen & Romeu 2000a), we looked at how the pedagogical activities in the classroom were carried out. The TEP methodology was thought to promote a more child-centered teaching method, which was rarely the case. We found that some of the personnel involved did not always have the necessary skills to carry out their tasks and that lack of physical resources in some places delayed the construction of schools and the supervision of the teachers. There were also weaknesses in the management of the project (organizational structure) and the cooperation between the parties involved. Furthermore we studied the relationship between the objectives and the projects' activities. Were the activities in line with the objectives?

Although the evaluation questions need to be thoroughly considered before the evaluation starts, one always has to look out for unexpected/unplanned effects. An example of an unplanned effect is that a project which is intended to reach the poorest becomes attractive to families who are better off and willing to find illegal ways of access. In unstable situations it may be important to identify security restrictions and sudden events. An evaluator must be open and flexible and willing to add new questions to her list as the investigation proceeds.
6 TERMS OF REFERENCE (TOR)

The Terms of reference describe the purpose of the evaluation, main objectives and evaluation questions. It is a document which directs the work of the evaluation team and the methodology, and it is therefore important to spend some time preparing a clear and complete terms of reference.

First of all the organization who commissions the evaluation has to find out what they want to focus on. Second, the team of evaluators and other involved parties or external expertise may be invited to contribute to the final version of the ToR.

Generally speaking it is better to focus on broader aspects of the project/program that are in line with the objectives and not on details (Samset 1993). It is the evaluation team who on basis of the ToR prepares a more detailed work plan that describes the methodology and design they find appropriate.

The following elements in a ToR are recommended by NORAD (Samset 1993) and Sida (Lewin 1993):
- Background (description of the project and its objectives)
- The purpose of the evaluation
- Scope and focus (which elements that will be evaluated, what is the main focus, evaluation questions and analysis, limitations)
- Methodology
- The evaluation team
- Time schedule
- Report (date for handing in a draft and final report)

Often the terms of reference are too vast and ambitious taken the time and resources into consideration. Sometimes questions and topics are suggested that are not possible to answer, at least not within the limited time available. As a result the evaluators know already before they start working that it is not feasible to follow the ToR. It is therefore better to make a realistic and simple document and listen to the evaluators' comments.
An evaluation of education project should include both quantitative and qualitative aspects. The quantity relates to factors that can be counted. How many children, teachers, members of the community are reached by the project, is often the question asked, or how many textbooks, notebooks, pencils have been distributed. How many teachers have been trained, how many children pass the exams, how many drop out and how many have to repeat, are other quantitative questions.

Although these factors can be counted, we also now that statistics lie. They do not tell the whole truth. Although a certain amount of books have been distributed, we cannot draw the conclusions that they reach the target children. An episode from a situation in Luanda comes to my mind. We gave a sandwich and a coca cola to a hungry little boy in the street. The minute after his elder brother came and took it from him. What does it mean that so and so many people have been "reached"? How do we know that a project which includes some of the children in a community has effects on the rest of the population?

The relatively simple question to find out how many children attend class, is not so simple after all. It depends on the rigor with which the attendance has been recorded. It is not uncommon to find a high attendance rate in the beginning and a considerable reduction by the end of the school year. If this is the case, it does not give much sense to record the average attendance rate over a year.

It is of interest to find out how many children pass the exam and with what results. However, even if the number is low compared with the number of children who entered the class in the beginning of the school year, this quantitative result gives us a piece of information that needs to be analyzed. Is the reason that the teaching was boring, inadequate, irrelevant? Or that the children had to leave school because of poverty in the family? That the students did not get enough individual attention and support? Or that it was the weakest children who did not pass the exam? Or may be the timing of the exam was bad? (because of harvest, rites of initiation).

Let's take a closer look at the results of exams as another example of the complexity of quantities. A class with a high pass rate and good results gives credit to the teachers. Our immediate thought is that the teaching is of high quality. But there may be several factors which contribute to the results, like for instance interested parents, a better standard of living and stability in the area. The parents' attitude to schooling contribute to the children's motivation and results. This is a consistent research finding from many countries. Even poor and uneducated parents may value schooling and encourage their children to continue despite of their difficult situation.

Another thing is how the teacher uses the criteria. Even if there are objective standards there are always differences between the teachers in how they apply them. The teachers' qualifications influence the way they evaluate the students' performance. The results also depend on the general level in the class. In a class with many high performers the level is different than in a class with low performers. Bribing and cheating are common in several countries where thousands of students sit in for national exams and only a small percentage passes. Girls may be sexually harassed and exploited by their teachers in promise of better
results. The tests and final exams (often the multiple choice format) may be of low quality and make cheating easy. By and large, it is not possible to judge what certain marks actually mean qualitatively. It is important to pass, but it is not a guarantee that the pupils really master the requirements.

Qualitative aspects cannot be counted, they have to be described. They relate to the way the children have been taught and how they have studied. The teaching in many countries is teacher-centered and the children spend most of the time copying from the blackboard - because they have no books, - and repeating simple sentences in unison. They are not allowed to work on their own solving relevant problems and therefore not able to use their knowledge independently in new settings. The quality of teaching also refers to how the teachers motivate the children, encourages them, uses positive and negative sanctions, supports weak learners and so on. It relates not only to the number of the questions the teacher asks, but how they are asked and what results they produce. Quality may also refer to the standard of curricula and text-books.

As we see a lot of the activities going on in education projects are of qualitative nature and difficult to assess rapidly. On the other hand, we need to take the counting of quantities seriously when it is possible and meaningful to count. The number of children in class and material distributed give us important indications that have qualitative implications. If the number in a class is 50 in stead of 25 (which it is supposed to be in TEP), it means that it is difficult for the teacher to assist weak learners individually. Unfortunately, there is often a lack of reliable statistics in many countries, i.e. on one hand it is an unrealistic belief in figures, on the other hand they may not have been collected correctly. There is a tendency to present figures without reservations and critical comments as to incompleteness and weaknesses.
8 WHEN TO EVALUATE

Ideally the planning of an evaluation should start parallel with the activities/project. But more often the opposite is the case; the coordinators start too think about evaluation when the project is concluded or about to finish.

The ideal is to arrange formative as well as a summative evaluations. The two are interlinked in the sense that results from formative evaluations feed the summative and have impact on the quality of the latter. Regular self-monitoring procedures can be developed and incorporated in the plan of operation. One may decide to review the project/program once each semester for example. It is important to invite all main stakeholders and beneficiaries to participate in the self-monitoring procedures. This will allow modifications and contribute to lessons learned in the organization. It is also valuable to arrange evaluation/monitoring across provinces involved in the project allowing them to learn from each other.

Questions to be asked in self-monitoring may be the following:
• What have we accomplished? (describe even small steps although they seem insignificant)
• Are the activities in line with the objectives?
• How far have we come in reaching the objectives?
• Do we need to modify the objectives?
• Is it necessary to modify other parts of the project?
• Management and coordination of the project - positive and negative aspect
• A review of the training of teachers
• The teaching in the classroom
• Material distributed, provision of equipment, school buildings etc.
• Partner cooperation

In a project in a developing country it is easy to focus on negative aspects, and to find things that are not going as planned, that are delayed or slow. It is important to concentrate equally on positive aspects, on what has been accomplished, although they may seem insignificant. During our evaluation of TEP in Angola, we found positive and negative aspects related to the project in all the four provinces visited. Some had succeeded where others had failed.
9 FACTORS THAT INFLUENCE EVALUATION

9.1 Lack of quality and objectivity

There are high expectations regarding the results of an evaluation which may have considerable impact on the future of a project/program. The expectations may be too high and sometimes the quality of the evaluation report does not justify such an attitude. The restrictions of an external evaluation, for instance that most of them are based on 2-3 weeks' fieldwork, are not communicated or under-valued. Vast and unrealistic recommendations may be presented that do not take the costs involved into consideration (Hallam 1998).

If time, money and resources are scarce, the evaluators have to draw upon their experience from other, similar projects to compensate for the lack of time to do the fieldwork properly (i.e. if fieldwork is part of the evaluation). Often the organization who engages consultants does not have qualified people who can look critically at the methods used and if the conclusions are grounded in the findings. Sometimes the results are presented without reservations, and even if they are, the organization chooses to overlook the reservations and emphasize the conclusions they expect or like.

Evaluations are like scientific research; the results depend on the evaluators' qualifications and on their preconceived ideas and preferences. It is an illusion to believe that research and evaluation are absolutely objective. The history of research is full of examples of reports focusing on the same phenomena which point in opposite directions, even though they pretend to apply similar criteria of scientific rigor. They may have used different methods which are more or less relevant, their scope and focus vary as well as their research questions. They highlight different results and they analyze the same results differently. This tells us that we need several types of evaluation that may contribute to the validity of the results.

9.2 External or internal evaluators?

The Swedish Aid development agency, Sida, emphasizes that the evaluator should be external and independent of the donor agencies as well as of the project management. It secures more objectivity and credibility (Lewin 1993). It is true that an independent, external, experienced evaluator's perspectives add to the evaluation's credibility. However, an external, experienced evaluator is not objective or value-free. Distance guarantees the ability to see a broader perspective because a person who is not familiar with the project observes details and raises questions which the local staff take for granted. However, external evaluators do not present everything they hear and see but have to consider the effects of the evaluation on those who have engaged them and other recipients and stakeholders.

Evaluations made by those who take part in the project on a daily basis also adds to the validity. The participants inside-perspective supplements the outsider's, even though they may point in opposite directions. The internal evaluators are restricted by home-blindness, loyalty to the management or the beneficiaries, they are too familiar with the project, but the external evaluator is restricted by snap-shot visits, limited time, lack of familiarity with the culture and may base their findings on what they believe they see, because it corresponds with other similar findings, not on what is actually going on (because the time in the field is limited). It is also easy to see how a group of evaluators on field work may not be able to find out what is
really going on when they are not there. Their massive presence influences the contacts they get and how the people contacted are willing to talk to them.

Many evaluations of projects in developing countries are restricted by time, money and resources. The purpose of the evaluation is practical: to give recommendations regarding improvements and modifications. To evaluate a project's impact on the development of a community more time is needed and different approaches (participatory observations).

Another restriction on evaluation is the attitude of the partners, Government and other organizations in the country where the evaluation takes place. They may not see the need for evaluation which may represent too much work and unpleasant results.

9.3 The team of evaluators

The evaluators may be internal or external. It is however, reason to define what is meant by "internal". It may mean persons from the project management or other people from the organization who are not directly involved in the project. External evaluators are contracted by the organization and have no obligations or loyalty to the employer.

It is an advantage to form a team that consists of local and foreign evaluators. It is never recommendable to engage only one category of consultants. The way the consultants should cooperate has to be defined as well as their role, positions and responsibilities in the evaluation. It has to be a leader or a coordinator of the team. It is also possible to include other locals and trainees in the team, but in such cases training is needed. The advantage of including local staff in the team is that it enhances the skills and the level of competence in the organization.

The evaluators' relationship with their informants has to be considered. Breakwell and Millward (1995) suggest to consider whether the evaluation is imposed or invited, participatory or non-participatory which describe the relationship between evaluators and those being exposed to evaluation. An evaluator who is known is liked by some, disliked by other. The evaluator's status and position influence the relationship. If the evaluators come from the Ministry or local education authorities, they may be associated with control and inspection, which influence what the informants are willing to share. But it may also mean that they want to talk about aspects of the project that they know the evaluator is familiar with. On the other hand, a foreigner with a different cultural background, language and appearance may also restrict the communication with the informants, or they may talk more easily with her/him since she/he is not affiliated with the project or the organization. However, his/her interpretation of the information is limited because she/he is not familiar with the context.
10 METHODOLOGICAL ISSUES

10.1 Sampling

Sampling regards the selection of sites and people who are going to be included in the evaluation. In research methodology there are rules as to the selection of samples, depending on whether it is a qualitative or a quantitative research design. In quantitative designs we talk about random sampling, which means that the objects of the study have to be picked at random allowing the researcher to generalize her/his findings. Purposive sampling is a procedure where the units under study are chosen for a specific purpose, for instance people who are reflective and willing to talk. Strategic sampling is another option which means dividing the population into homogenous groups, for instance by separating female and male teachers.

In education projects, the sampling is often a practical issue, which has a built-in bias. Which schools and districts to visit depends on the time available for the field visit, means of transportation and the time of the year. Many remote places are not accessible during the rainy season, at other times the community is busy with the harvest and many children do not show up in school. More often than not the geographically accessible sites are visited most frequently, and as a consequence remote schools and communities who are in most need of support and therefore should be included in an evaluation are not selected. In our evaluation of the TEP program in Angola we had more time than usual devoted to field work, and in some of the districts we insisted on visiting all the schools who were involved in the program. The schools who were farthest had the least trained teachers and the most difficult working and living conditions but at the same time they were closely attached and devoted to the community.

The evaluator may plan his/her sampling, but the final choice has to be negotiated with the local education authorities. The evaluators present their concern to visit schools that are different in size and location and sites where the project seems to be successful/not successful. The best possible choice has to be made taking time and other factors into consideration.

Bias in the sampling means that we cannot draw the conclusion that our findings can be generalized to all the schools in the province. The final results might differ if the sample were more balanced.

10.2 Methods

The choice of methods depends on the questions that need to be answered. If we want to know people's opinions and beliefs, interviewing is an appropriate tool. But if we want information on their actual behavior we must observe them. As a rule of thumb it is advisable to combine several methods, called "triangulation" in the research litterateur, for instance to start by interviewing and to observe afterwards if the interviewees practice what they preach. Observation feeds the interviewing and vice versa. Combining various methods allow comparisons, checking the data gathered and thus strengthens the validity.
Taking the time and resources into consideration there are basically three different methods that are most common in evaluation: observation, interviews and reviewing existing information. We shall look more closely into these methods.

10.2.1 Reviewing existing information.

This is often called "secondary sources", but this does not mean that it is of less importance than first-hand information. It is not uncommon to start an evaluation without first checking out what information already exists. Depending on the objective, relevant existing information in education evaluation may be the Government's policy documents and white papers, Government statistics, other statistical information, documents/reports from other development agencies, evaluation reports, research, literature, correspondence, local records, archives, minutes, newspapers and journals, and people who are familiar with the topic and the historical background (Pratt and Loizos 1992).

The advantage of analyzing existing written information is that it can be done unobtrusively, i.e. the evaluator does not need to interact with the informants. The disadvantage is that the information may be incomplete and therefore need to be double-checked. Talking to resource persons with knowledge on the subject may allow you to check the written material and lead you to other existing sources of information. Different types of written material have to be compared and analyzed.

Reviewing existing information may reveal that there already exist evaluation reports done by other organizations on the same topic, or that there are already answers to some of the questions raised in the Terms of Reference.

A historical background is always needed in an evaluation as well as an introduction to the topic and the context, based on existing information. It is worth while to spend time reviewing existing information in all evaluations. When the available information has been studied and the various pieces of information compared, the next step may be to check and supplement it by talking to resource persons who are well acquainted with the subject.

10.2.2 Asking questions: Interviews and questionnaires

An interview is a planned conversation which the interviewer has prepared according to questions that need to be answered. It is advisable to interview different beneficiaries, like education authorities, female and male teachers, headmasters, children (girls and boys), parents (mothers and fathers) and people from the local community, and tailor the questions to their position but also ask them some of the same questions in order to make comparisons. We cannot draw the conclusion that a particular educational method is a success just by talking to the education authorities and the teachers.

The interview may be highly or loosely structured. A written questionnaire is an example of a highly structured set of questions where the questions are sequenced and the optional answers fixed or open. It may be conducted by an interviewer or answered individually with or without the presence of the interviewer. It is difficult to make a good and understandable questionnaire that the informant wants to answer. The questions may not be specific or clear enough, the answers you want to give do not fit into the prepared categories, or the
questionnaire is too long and complicated. It also requires that the informants master the language well.

The advantage of a questionnaire is that it is a way to collect relevant facts in an easy, cost-efficient and non-obtrusive way. Questionnaires are also used to find out the informants opinions and attitudes. However, the use of questionnaires has to be considered in the cultural context within which it takes place and the motive for using this type of method explained. In many cultural settings it is not appropriate. Written statements are associated with the police or political militants (Pratt and Loizos 1992) and people may be reluctant to answer or give answers they believe the interviewers like to get.

On one occasion I distributed a simple questionnaire to collect background information about the interviewees who were taking part in a group interview. By the end of the session it turned out that one of the participants thought that the reason I had collected the information about education and professional experience was because of the remuneration he expected to receive when he participated in the interview.

In many cultures, the Western way of conducting a structured or half-structured interview is alien to the way people talk and get information from one another. It may be considered impolite and inappropriate to sit the way we normally do when we conduct interviews, to look at the interviewer/interviewee and to write down their answers. It is difficult and even impossible for a male interviewer to interviewee females in many cultures and also to engage children in an interview where they talk at ease with a stranger. Adults are authority persons who may report to the teachers, directors or parents what they say.

In emergencies one has to take into account the vulnerability of the interviewees who have witnessed traumatic situations. They may not be willing to talk about their experience and the interviewer must be sensitive to their wishes and able to handle unexpected reactions.

It is therefore recommendable to engage people in an open, non-formal conversation based on the questions prepared which the interviewers have in the back on their mind and prepare how to take down notes.

It is sometimes better to gather the interviewees for a group discussion than for individual interviews. The group discussion is focused around some main issues.

10.2.3 Group interviews

Interviewing groups of people is less time-consuming than individual interviewing. In groups interviews the focus is more on the topic discussed, than on the individual sentiments and opinions. If for instance, the evaluators observe in a school, they may want to gather the individual teachers in a group for an interview after the observations in stead of talking to each teacher individually.

In focus group discussions, a certain number of people are selected who are known to have special knowledge and interest in the topic. In such cases the interviewer does not need to prepare many specific questions. Her/his task is to observe how the group discusses the topic (see more about focus group discussions later).
If the group has many participants, it will affect the individual participation. In a group of 10-12 people, only a few will dominate the discussion, while the general participation will invariably grow if the group consists of 5-6 persons. Another issue in group discussions, is to avoid false consensus. There may be different opinions among the members which are equally interesting to the evaluators.

10.2.4 Interviewing children

When children are among the beneficiaries, their voice should also be heard as part of the evaluation. The children's opinions are different from the teachers' and the parents'. It is however difficult to engage children in a general, philosophical discussion about teaching and teachers, if they are young (grade 1-3/4). Specific questions related to the children's experiences have to be asked. In a project where the impact of a particular educational method was evaluated in Mozambique (Johannessen 2000b), we asked the teachers how they handled discipline situations in the classroom, like when the children come late to school, make noise or fight in the classroom, do not listen to the teacher ans. Afterwards we asked the children the same questions, and the picture which emerged was that there was little correspondence between what the teachers said they did and what the children told us about the same situations.

Often there is an attitude among adults not to put too much emphasis on what children say. Their view is considered biased and unreliable because of their young age while the adults' views and opinions are considered more reliable. In the above-mentioned evaluation we asked the children the same questions in slightly modified ways during the interviews and thus managed to separate teachers who tried to use the educational method from those who did not, according to the children.

The difference in age and status between an adult interviewer and a child creates challenges, particularly in hierarchical societies. The children look upon the adult as a powerful person who they have to obey. They are not used to say their opinion in the presence of adults. This may be counteracted by the attitude of the interviewer and by interviewing children in groups. When gender sensitive issues are discussed, the group should be divided by gender. The interviewer must pay special attention to the girls' participation who may be shy and talk less than the boys. The groups should not have more than 4-6 children. If all the children in a class is interviewed, only a few will participate.

Because of the traditional way children are treated in many societies, they are used to a type of communication were the adults ask the questions, which they have to answer. The answers may sometimes be brief, or the children may be shy to answer (particularly girls in the countryside). Because of tradition and habit, many adult interviewers tend to ask too many questions. If the children don't answer readily or are more eager to find out which answers the interviewer wants than expressing their own opinion, the interviewer may put the "right" answer in their mouth by asking leading questions. Interviewing children presupposes enough time to make them feel comfortable and relaxed and to encourage them to talk freely. Even though the children may say things that are strange and contrary to what the adults believe in - the interviewers must respect and appreciate their answers. In a group interview with children during the same evaluation, we asked the children what a good teacher is or does, and four of the five children in a group said it is a teacher who does not beat. When we started to discuss beating and if they believed it helped, the same four said it did not, but the fifth boy said ha-
believed it did, and one of the adult interviewers tried to suppress a laughter. I congratulated the boy because he said what he believed and told him that it was ok to have a different opinion from the others and that we should all respect each others right to express our personal opinions.

Adult interviewers may feel obliged to teach or educate the children in an interview, particularly if they are teachers or inexperienced interviewers. It may also be a problem with two adult interviewers and a few children if the interviewers talk more than the children. It is therefore advisable that only one adult interviews the children and of course, without the presence of teachers or other adults.

10.2.5 Observation

If we want to know more about people's behavior, we need to observe them. It is far more common in evaluations to use checklists and interviews than observation. Interviews give us access to how the interviewees interpret their reality, while observation gives us direct access to that reality. On the other hand, observation is difficult because what the observer sees is filtered through his/her own belief system and preconceived ideas. What we observe is colored by our own experience as students and teachers. It is therefore excellent to combine interviews and observations.

However, there are several pitfalls. In order to draw any valid conclusions, one has to make several observations. The quality of the observations depends to a large degree on the training and experience of the observer. In practice, it is not possible for an evaluator on a short fieldtrip to make a sufficient amount of observations; he/she has to be selective. One can either choose to make several observations of the same teacher or to visit several classrooms. It is important to ask the question: Is this a special day, a special time of the year, a special week? I have for instance experienced that teachers extend the lesson when I arrive because they want to fulfill our expectations. But when the pupils are tired and ready for a break, it is not the right time for observations. Many times lessons have been arranged for the visitors to show us their teaching and several classes are mixed, but this does not give us an impression of the relationship between the regular teacher and their students. As a rule of thumb the observer should stay for the whole lesson and not interrupt the teaching. The observer needs to check the information she/he gets with other sources of information in order to control biases.

The observer's focus of attention determines the results. The observations should be guided by the principal questions in the ToR. Often it is necessary to transform the questions into recognizable activities that take place in the classroom. We had to do this when we observed in the classrooms in Mozambique in order to find out the impact of the method on the teaching in primary schools. The same has to be done when we are looking for so-called child-centered teaching methods or "quality of teaching". Such methods have to be concretized and separated from methods that are not child-centered or of poor quality. This is a difficult process and it is related to the issue of validity (see footnote 1). Do we measure what we pretend to measure, are we sure that our "translation" of child-centered methods into teaching methods really cover the term child-centered?

Some observers use checklists as the basis for their observations. The checklists have to be elaborated in line with the questions in the ToR. The reason why checklists are used is that they are less time consuming than other methods. The problem with checklists is that several
observations are required to give accurate answers. Another problem is that the points in the prepared checklist do not always correspond with what goes on in the classroom; thus they do not catch unplanned events. Therefore the checklists have to be supplemented by comments and elaborations and in the end they may prove not to be so time-efficient after all.

Another option is to make guidelines for the observations (see 11.4.2). Guidelines may be similar to checklists or ask for descriptive observations that are more elaborate than checklists. It is generally not sufficient to focus narrowly on the aspects defined as relevant according to the ToR when doing observations. Notes on "irrelevant" aspects may prove valuable in order to understand the context. And the use of strictly structured checklists and guidelines do not save the observers from the laborious interpretations of the observations. As a rule of thumb extensive, descriptive notes based on direct observations are an excellent source of understanding the teacher-student interaction in the classroom. It is difficult for an inexperienced observer to separate observation of what actually goes not from his/her own interpretations. The inexperienced observer should try to put his/her interpretations in brackets and train her/himself in describing what she/he sees as detailed as possible. The subjectivity involved in observation is a common objection to using this methods. This is why the observers should put more emphasis on substantiating their interpretations by referring to the observations.

The observer must also reflect critically upon how his/her presence influences the observations. The teacher and the students may act differently because of the intruder. This is also the reason why there should not be more than one observer in a class and that the observer stays for the whole lesson and tries to use a non-participatory approach. My experience is that if the observer stays for 45 minutes or an hour without interrupting, the teacher and students will after a while fall back into their ordinary classroom behavior, at least partially.

Participant observation means that the evaluators spend more time in the field and observe while they are participating in the activities. It helps the evaluator to understand the project from the perspective of the beneficiaries. If the period of fieldwork is short and the evaluators many, it is not possible apply this method correctly.

10.2.6 Other methods

Logical Framework Analysis (LFA) was first introduced by USAID, and has later been adopted by many other governmental agencies and NGOs (among others NORAD and Save the Children Norway). LFA presents a matrix to ensure a logical relationship between objectives, goals, purpose, inputs and outputs. It emphasizes elaboration of indicators to find out how objectives are being met. It is used in the design, planning, implementation, monitoring and evaluation of a project (for more information on LFA, see Cracknell 2000, Dale 1998). As previously mentioned, it is not uncommon to find that the concepts have not been satisfactorily specified when the project was planned which makes it difficult to know precisely what to focus on in the evaluation. In stead, objectives, goals, purpose, inputs and outputs are presented in general and vague terms and when the time comes for a summative evaluation and the terms of reference require more specific descriptions, the weaknesses in the planning are discovered. However, it is not always easy to define the difference between the LFA concepts and one may risk spending a lot of time on this exercise. It has been argued that LFA is too rigid, particularly for humanitarian assistance program where situations may change rapidly and require new and swift actions. It forces people to fill in boxes without...
thinking. It may look neat on the paper to see the specification of inputs and outputs but the relationship between the two may be difficult to prove. (Hallam 1998). Another objection is that it is a typical top-down Western way of thinking which comes from the donor agencies and which does not take the recipients' perspectives into consideration (Cracknell 2000, Pratt and Loizos 1992).

Rapid Appraisal Methods. Rapid appraisal methods have been developed and used as an alternative to traditional and costly research methods. Rapid rural appraisal (RRA) is used by many NGOs (for instance Oxfam) as a way "to identify the problems, goals and strategies of households, groups and communities" (Pratt and Loizos 1992 p. 66). It is carried out by a team of professionals and project staff. The most common techniques are:

- investigation of existing secondary material
- direct field observation
- semi-structured interviews
- group interviews
- informal workshops with local people
- ranking exercises
- role-playing exercises
- sketch-maps, diagrams, and other visual aids

Participatory rural appraisal grew out of RRA with more emphasis on local participation. It is the participation which is the most important not the time aspect. Agency staff and locals do the data-gathering and analysis together (ibid.)

Other rapid appraisal methods used in research and evaluation are

- key informant interviews, i.e. selecting a group of individuals who are in a position to provide the needed information, ideas and insights (Kumar 1993).
- Focus group interviews. This method has long been used by marketing researchers to find out the consumers' reactions to new products. The focus group interviews are organized to discuss a specific topic in group sessions. The participants are encouraged to share their ideas and opinions freely, and to comment and criticize each other. The group is limited to 8-12 carefully selected participants. The group should be homogenous in composition, whenever possible. It is conducted by a mediator whose job is to use probing techniques to keep the discussion going. The discussion lasts for about 2 hours.
- Community interviews. It takes form of a public meeting open to all community members. Unlike the focus group interview the interviewer has prepared a set of questions and the interaction takes place between the interviewer and the community members.
- Structured direct observation. The observation is structured according to the focus of the study. It can deal with physical objects as well as social processes, and it tries to take the inside perspective into consideration. It is conducted by a team of observers.
- Informal surveys. Is mostly used in studies of agricultural and rural development interventions. An open-ended questionnaire which allows the participants to answer in their own words, is applied. The participants are selected according to a non-probability sampling procedure, for instance, people in the market or in a public meeting may be approached.

An objection that has been raised against participatory methods are that the evaluators lose their independence and personal objectivity since they become too involved with the participants (Cracknell 2000). Creating a genuine equal relationship between evaluators and local participants is not easy among poor and illiterate people. Limitations of rapid appraisal
methods are that although they claim to be rapid, it does not mean that they are easy to use. If the evaluators are not properly trained, the information they get may be biased and confirm their own preconceived ideas. The gathering of data may be too rapid focusing mainly on what is quantifiable and not on qualitative data which are more difficult to record and code (Kumar 1993).

Participatory methods are never rapid, as it takes time to build a relationship between the evaluators and the community members. True participatory methods are similar to the requirements in participant observation where the time aspect is crucial.

Hallam (1998) suggests that the use of narratives is useful in evaluation of emergency programs. That is to collect information from various actors who have been involved in a situation or an event in order to reconstruct the various details and important turning points.

10.2.7 Which research methods are good enough in evaluation of education projects in emergencies?

When it comes to evaluation methods, we have seen that there are several approaches called "rapid". But in order to use such methods rapidly and with ease, the evaluators need training and experience.

There are no exact answers as to the number of interviews and observations needed for example, in order to qualify the methodology. Evaluations we are talking about here never allow enough time to use scientific research methods as prescribed. The choice regarding which methods to use may often be decided on the spot. It depends on the availability of educational authorities, headmasters, children, teachers and parents and schools. The evaluators have to be flexible. Moving around freely is not always possible and therefore certain places are not possible to visit. However, in acute emergencies when the children and their parents are placed in a camp, common obstacles related to long distances, transportation and availability no longer apply. The informants are available due to their life situation in the camp. In such situations it might be possible to organize interviews with many refugees/IDPs as well as to observe the education activities.

As a rule, it is advisable to use different methods even if the number of interviews and observations is reduced. Methods and results are not good enough as such. Only interviewing is not good enough. Interviewing, observation and document analysis give a better chance of getting a complete and valid picture of the status of the project. Interviewing one group is not good enough, but interviewing several and different stakeholders on the same matter is better thus getting different perspectives. It is possible to ask them the same questions, to see if the answers are different. Observing may include several or one teacher only. If possible and relevant, one may choose teachers who have been trained and those who have not been trained to make comparisons.

Here follows an example from the evaluation of Mediated Learning Experience (MLE) in Mozambique (Johannessen 2000b) which illustrates the use of different methods and how one may design this type of evaluation.

• Preparation. Before I arrived we had asked the local staff to provide lists of all the teachers who had been trained in MLE and select participants to one group in each province. The participants came from different districts in the province.
• **Evaluation teams.** Two evaluation teams were formed, one in each of the provinces visited. Each consisted of two educators from the local staff and myself.

• **Kick-off seminar.** The design, timetable and methods were discussed and planned in a one-day seminar with all members of the evaluation teams present.

• **Reviewing existing document.** Documents that had been compiled by the local staff (Save the Children Norway) allowed us to trace the historical background of MLE and how it was introduced.

• **Group discussions with teachers.** The participants had been chosen by the local education authorities and local staff and consisted of 10-15 people in each group. We spent one day in each province on this type of group discussions. A list of questions for this purpose had been prepared in the kick-off seminar.

• **Classroom observation.** The group participants were observed (one observer in each classroom) in action in their classrooms. Guidelines for classroom observation were prepared in the kick-off seminar.

• **Interviews with children.** Small groups of children (4-5 in each) from the same classrooms as we had been observing met with one interviewer. Questions had been prepared in the kick-off seminar.

• **Conversations with parents** took place in some communities.

Some of the questions we asked teachers and children were the same, and we also focused on the same areas in the classroom observation. The design allowed us to make systematic comparisons between what was said in the group discussions, observed in the classroom and told us by the children (Johannessen 2000b).

No methods and design are flawless. It is not uncommon to find that the choice of methods and how they were used is lacking in evaluation reports. The most important criterion which defines the quality of the results is how the evaluators are able to present and reflect upon the strengths and the weaknesses of their methodology and the findings.
11 EVALUATION OF EDUCATION PROJECTS IN EMERGENCIES

In the following we shall give a few examples of main results and lessons learned in three different education projects in emergencies, one present main findings of an evaluation of a program in a chronic emergency, the second focuses on lessons learned in an acute emergency and the third example is from an education program for refugees in a pre-return phase. They will give us ideas regarding which indicators that are relevant in the different situations. This is followed by a checklista and guidelines for classroom observations.

11.1 Education in a chronically unstable situation: TEP (Teacher emergency package) in Angola

Angola is an example of a country in a prolonged chronically unstable situation. A civil war has been going on since the liberation from Portugal in 75. There have been periods of hope of reconciliation, followed by new tensions. Provinces and districts that have been considered secure for some time, may suddenly be attacked forcing the population to flee to other places. Many IDPs have fled to the cities. The influx of people to the cities have resulted in overcrowded classrooms and lack of education material in some places, while schools are left empty in rural areas headed by the least qualified teachers.

The introduction of the TEP program in Angola is meant to reach children in school age (6-14) who have never been to school or have had to leave school due to the war. The main challenges noted in the TEP program are the following:

- The target group is not always reached. Around 1/5 of the children in the program in some provinces were below school age. This is not a result of the "baby-sitter" problem of older children taking care of their younger siblings, but may be related to the fact that it is easier to enroll smaller children than older who have never been to school.
- A high drop out rate among the children in some provinces. One reason is the unstable situation for the IDPs who are on the move.
- Trained teacher who flee to other provinces. This may be due to the unstable situation, difficult working and living conditions and/or to the meager remuneration they get.
- Incomplete statistical information at all levels. This may also be due to the instability as well as to lack of training of teachers and officials in how to record and process statistical information.
- The recruitment and training of unqualified teachers. When there are not enough trained teachers who are IDPs, persons without teacher training have been recruited. Their remuneration is a main problem.
- Formal cooperation with the national and provincial education authorities is a decisive factor for the outcome of the project. It regards among other things: the transfer of TEP students to ordinary school, criteria for the evaluation of TEP students' performance, the transfer of TEP teachers to the government's payroll, the planning of the TEP, supervision and training, salaries etc.
• That the emergency project does not become more attractive than the ordinary public schools. The reason is that education material is provided for free in TEP and the size of the class is small.
• Conflicts between cooperating parties, NRC, UNICEF and MEC (Ministry of education and culture) due to how they are able to work in emergencies. NRC is a small and efficient NGO tailored to act rapidly in emergencies, while UNICEF is a bureaucratic UN organization, and MEC is also burdened by bureaucratic, lack of financial means and inefficient procedures. This creates tensions and slows down the efficiency of the implementation of the program.
• "The emergency character of the TEP can be over-emphasized and result in unplanned and spontaneous actions. Although the situation in Angola at times calls for rapid actions this does not mean that the organizations need to respond to emergency needs immediately, particularly not with respect to education which cannot be compared with food distribution" (Johannessen 2000a).
• Lack of coordination of the projects' various parts: training and recruitment of teachers must go parallel with the construction of schools, the enrolment of children, the provision of complete TEP kits, the teacher's salaries, and the practical arrangements in running the schools
• Remuneration and incentives to teachers turn out to be a constant problem. The teachers who work without any type of incentives may leave the program and the teachers who get incentives demand more. Food distribution (food for work) is apparently not a good solution because the deliveries are not regular.
• The selection and training of supervisors, and the organization of a supervisory system. The supervisors are particularly important in an emergency project. They visit the teachers regularly, keep track of them and encourage them to continue. It is reason to believe that supervision may function as an incentive to the teachers.
• Gender specific issues. How to encourage the girls' active participation in the classroom and recruit more female teachers
• The quality of the teaching. Lack of proper use of the educational kit, use of local didactic material, low academic level among the teachers, little psychological and pedagogical knowledge and training, no/little use of child-centered teaching methods.

11.2 Education in acute emergencies: Education for refugee children from Kosovo in Macedonia

These are some of the questions raised in another education project that took place in an acute emergency situation for children from Kosovo who fled to Macedonia (Midttun 2001):
• Some difficulties in the cooperation between UNICEF and NRC due to the first-mentioned organization's lack of capacity to act rapidly and provide funds in order to buy the necessary school material (pens and notebooks).
• Problems with curriculum and text-books. The Kosovo-Albanians were not allowed to use their own curriculum and text-books for political reasons. Only Macedonian curriculum and books were allowed
• Not enough textbooks.
• Important to attend to the children's need for recreation and play taking into account their traumatic background, and not only attend to their academic needs. Subjects like arts, music, drama and sports were included.
• An educational kit should be available
• Provision of temporary certificates were done by NRC and the headmasters of the local schools while waiting for official ones
• The running of the schools. Important to leave the running of the schools and the programs to the refugee themselves, and only facilitate and support when necessary
• The use of local schools. Many advantages of camp children using local schools
• After approx. 2 months, the decision was taken to stop the education project. This had negative effect on the camp life, the teachers and the children who felt let down. The decision was taken by people high up in the system and it was felt that it was not coordinated with the situation in the camps
• Terms of reference changed according to the new situation when the return started. Necessary to renegotiate.
• Necessary to develop better communication lines with partners and other actors in the camp
• Even in emergency situations it is essential to take the time needed to secure good communication and routines of information, to clarify expectations, tasks and duties.
• Frequent and effective meetings between UNICEF education officers and the NRC team are extremely important.
• Cooperation with local education authorities and other parties involved in and around the education program is necessary
• It is necessary to make a decision regarding whether to provide incentives to teachers or not (although it is even more relevant in chronic situations).

11.3 The International Rescue Committee Education program for refugees in Guinea 1991-98: A best practice study in a "care and maintenance" phase

In a review of the International Rescue Committee's education program for refugees from Liberia and Sierra Leone in Guinea (Lange 1998), many of the above-mentioned points are mentioned as contributing to the program's success. The following are highlighted:
• establishment of active Parent-Teacher Associations (PTAs), particularly responsible for construction of school buildings
• education coordinators. Continuous teacher training and guidance is an outstanding feature of the program. A group of field-based, mobile advisers, called education coordinators, is considered a key factor in this respect. Their job is to monitor teacher performance, provide in-service training and assistance, arrange mini-workshops for teachers, assess needs and communicate IRC policies. They are the link between the teachers in the field and the headquarters.
• Each subject has a curriculum manual for the teachers.
• The program has a strong and active administration who has developed good routines regarding the collection of information from the field, a functional report system and monthly statistical reports.

11.4 In search of indicators

To indicate means to "direct attention to; to point out". We need to direct our attention to the aspects of a project that are most important and in line with the objectives and questions we want to find an answer to. Indicators are associated with measurement and items that can be quantified (called "objectively verifiable indicators", Cracknell 2000). Many donor agencies
have developed indicator banks or lists of indicators that might be used in different projects. The World Bank is currently developing sets of indicators that are suitable for its various sectors.

The development of indicators is part of the planning of a project when using Logical Framework as a tool (see page 24). They should cover the objectives of the project. Cracknell (2000) holds that indicators "need to be verifiable, reliable, relevant, sensitive, specific, cost-effective, and timely" (p. 110). But he also warns that it is not always meaningful and possible to use quantified indicators. The irony of measurement is that "all too often we measure the relatively important things because the really important ones cannot be measured" (ibid.). He recommends the use of quantitative and qualitative indicators "rather than straining for quantification, it is often better to use qualitative indicators which may be mainly descriptive, or which may depend upon an opinion given by someone well qualified to judge" (ibid.). As we have mentioned earlier (chapter 7), there are many qualitative issues in education which have to be described. This means that we must put emphasis on clear, specific and detailed descriptions.

Developing specific and useful indicators is time-consuming and laborious. The more specific and detailed they are, the better chance we have of getting a complete picture of the project. Even if we manage to produce a set of indicators, we cannot be sure that it will give us the key to explain the success or failure of a project. There may be other indicators than those we have focused on that turn out to be more important. If we for instance train teachers in the best possible way and they still do not seem to deliver better work, one possible explanation is that they are not motivated for the teaching profession. But "motivation" may not be on our list of indicators because they belong to what Cracknell calls "really important indicators" that are difficult to measure. It is important to elaborate indicators together with people who have first-hand field knowledge of the project.

Some of the key elements in education program are:
- the teachers and their qualifications
- in-service training, monitoring and follow-up
- provision and development/adaptation of education material, teaching aids, educational kits
- mobilization and involvement of the parents and the community

The relevance and importance of each of these elements in emergencies vary according to whether it is an acute, chronic or rehabilitation emergency.

In acute emergencies - assuming there is an established interest and need for the activities - the main elements are to
- identify and mobilise teachers
- procure tents/classrooms and desks
- provide equipment and teaching aids
- provide text-books
- provide specific material related to the situation (Health, sanitation etc). (Midttun 2000b)

It is also important to mobilize the parents and the local community and to cooperate with the local education authorities when possible. Experience has shown that monitoring and supervising the teachers are of particular importance in emergencies (cf. Lange 1998).
may not be possible in acute emergencies because of lack of qualified supervisors. However, it is necessary to consider the teachers' need of support and follow-up.

Even though books and material are not ready, the teaching can start when the teachers have been identified. The most important is to gather the children rapidly and organize play or other activities with the material that is available. If a regular routine of coming to school every day is established, it may ease the children's and their parents' difficult situation.

If the situation of emergency is prolonged, it is possible to focus more on the quality and development of the program, i.e. on the development of teaching aids, textbooks, curricula, training of trainers and teachers, supervision and monitoring and on the follow-up of the program. In addition comes organizing extra-curricular activities, training of mothers and youth for care of pre-school children, considering income generating activities to pay school fees, school feeding program, water and sanitation for school/staff, staff housing etc. (ibid.).

Furthermore, the cooperation with national and local education authorities is essential in view of the sustainability of the program.

In the rehabilitation/reintegration process, the following must be considered
• adapt program as necessary for returnees
• facilitate integration of war victims/child soldiers into programs

It is particularly important to consider the target groups' needs in emergencies and that this should be reflected in the curriculum. "Basic learning and basic knowledge for coping, physically and mentally, is given priority along with knowledge and practice that will help promote a peaceful and values/rights based rebuilding of the war-torn communities" (ibid.). Education in emergencies needs to be flexible and adapted to the local needs and situations. Many of the ordinary requirements in normal school situations must be reconsidered, for instance the teachers' academic background, the children's age, the school facilities and education material.

An organization's ability to act rapidly and flexibly is crucial in emergencies. However, it has to be analyzed "how rapid is rapid" ensuring that the actions are well planned and defined and that the organizations are capable of following up. Although the needs are many and acute it does not mean that planning needs to be superficial and hasty. It is necessary to agree upon steps that ensures quick organization and start of activities, for instance whether to start with lessons or recreational activities.

In the following we shall present a checklist which covers areas that are relevant for education projects in emergencies, followed by guidelines for classroom observation. The check-list may serve as a tool in the planning, implementation and evaluation of projects. A list of questions related to the areas is enclosed, including questions for the evaluators. The list may be further developed and supplemented and finally develop into a list of indicators. The areas/questions are not applicable to all projects but must be tailored to each specific case. The checklist is mainly based on findings and lessons learnt from education projects in emergencies. Some items have been developed according to "the questions for evaluators" presented in Hallam's book on evaluation of emergency projects, annex III (1998).

All items below must be specified according to gender (girls/boys, female/male teachers/supervisors etc.).
11.4.1 Checklist

The questions in appendix I supplements and specifies the checklist further. Some of the areas/questions need quantification, others need to be described in qualitative terms. The evaluation should focus on description of processes(major changes and turning points) and final results.

1. **Understanding the context**
   - assessment of the initial situation
   - defining and considering security implications
   - considering political/ideological conflicts and the possibility of new movements/sudden changes
   - research on the culture and history of the beneficiaries
   - considering the impact of the emergency program on host populations and the local environment
   - agencies that are involved in the analysis of the problem and the plans

2. **Specifying objectives, activities, outcome, priorities, restrictions and beneficiaries**
   - Specification of objectives, indicators, activities according to Logical Framework Analysis or other conceptual models (when applicable)
   - Prioritizing objectives and activities
   - Assumptions about the outcome
   - Considering lessons learnt from previous programs
   - Restrictions to implementation, like political, economic, cultural and logistical factors
   - Timing
   - Defining the target group/beneficiary groups, and their special needs (for instance children/youth who have never been to school, girls, disabled, ethnic minorities)

3. **Management and cooperation**
   - Management of the project at national and local level
   - Cooperation between agencies, between agencies and government or between government and other national forces
   - Basic agreements, lead agency, focal points
   - The use of national resources
   - Factors contributing to lack of cooperation/smooth cooperation
   - Division of responsibilities between cooperating partners
   - Lines of communication
   - Delegation of decision making
   - Regular information-sharing with agencies, government, donors, beneficiaries
   - The cooperating agencies' ability to meet needs rapidly
   - Cooperation and communication with
     - the Ministry of Education
     - local community, local leaders, parents, PTA, the church, religious and other local organizations
     - the camp administration
     - local and provincial education authorities
     - local schools

4. **The selection of target groups/beneficiaries**
   - the criteria for the selection of target groups
• controlling that the target groups are reached
• the number of beneficiaries the program intends to reach

5. The selection of teachers
• the criteria for the selection of teachers
• controlling that the teachers selected meet the criteria
• the number of teachers the program needs to mobilize
• the number of untrained teachers needed and their qualifications

6. The schools/classrooms
• the number of schools/classrooms needed
• selecting location of schools/classes
• the need for construction of new buildings, rehabilitation, or use of old buildings, classes under a tree
• provision of desks/chairs/tables/blackboards
• financing, organization and administration of construction/rehabilitation work
• cooperation between parents and agencies in the construction of schools/classrooms

7. The teaching
• national/local curricula
• if the needs of war traumatized children are considered
• subjects
• textbooks
• notebooks/pencils/chalk
• teachers’ guides/manuals
• lesson plans
• ready-made and local didactic material
• educational kits, their contents and use

For more details on qualitative aspects of teaching, learning climate and the relationship between teacher and students, see "Guidelines for classroom observation" below.

8. Extra-curricular activities
• type of extra-curricular activities
• who is responsible for such activities

9. The teachers
• number in-service/number needed
• academic level/experience/other qualifications
• training
• training needs
• work load/shifts
• distance to school/classes
• means of transportation
• living conditions
• salaries, incentives
10. Teacher training
• Plan/curriculum for training of teachers
• Duration of the training
• Components of training
• Subjects
• Selection and qualifications of trainers
• Teaching methods
• Education and didactic material
• Teachers' guide/manuals
• Trainers' guide/manuals

10. Monitoring and supervision of teachers
• criteria for selection of supervisors
• controlling if the criteria are met
• supervisors' academic background and experience
• frequency and duration of monitoring and supervision
• distances and means of transportation
• administration of supervisors
• supervisors' workload/other jobs
• definition of the supervisor's tasks
• supervisors' reports
• remuneration/incentives

11. Management and administration of teachers/supervisors/schools
• plans for the number of teachers needed
• plans for the transfer of teachers to the government's payroll
• transfer of children to ordinary schools
• placement of teachers
• control of teachers' attendance and work
• daily management of schools/classes
• distributing/supplementing/controlling the kits
• regular collection and processing of statistical information
• conflicts between teachers/headmaster, between teachers, teachers/supervisors/headmasters

12. The children
• statistical information on the enrolment/regular attendance/dropping out/results from tests and exams/age distribution/gender/number of children in class
• children with special needs/learning difficulties
• distance to school

13. Development of plans/curricula/education material
• existing plans/curricula/textbooks and their quality
• development of plans/curricula/textbooks
• existing didactic material and their quality
• development of didactic material

14. The parents and the community
• the parents' and communities' involvement in the schooling
PTAs (parent-teacher associations) and their responsibility
financial contribution from parents and the community

11.4.2 Guidelines for classroom observations

- Camp/village/city………
- Name of school…………
- Grade……………
- Number of girls in the class (present)………..
- Number of boys (present)……………………
- Number of pupils registered in the class according to the attendance list (girls/boys)…………
- The pupils' age (boys/girls)………………
- How young is the youngest child and how old is the oldest? (boys/girls)
- The teacher (sex, training)…………………
- Take a look at the attendance list

The school's/classrooms' physical conditions
- Descriptions of the physical conditions: the tent, building, roof, floor, windows, doors, latrines (how many and if they have separate latrines for boys/girls), access to water, light/electricity
- Describe the environment (distance to the village, if it is a dry or a wet area etc.)
- Ask how the situation is during the rainy season

The physical conditions in the classroom
- Desks or tables/chairs for all the pupils, chair/desk for the teacher, position and quality of blackboard, chalk, drawings/photographs/maps, decorations

Practical arrangements
- How the children are seated, if there is enough space to organize the pupils in smaller groups, if all the children hear and see well

Educational material
- Text-books, notebooks, pencils, pens
- Ready-made and local didactic material
- Teacher's manuals
- Other equipment
- Material from the educational kit

The teacher
- Describe the teacher's appearance (looks, if she/he seems eager and alert/interested or tired/unmotivated etc.)- first impressions

The children
- Describe the children's appearance - first impressions

The teaching and the teacher-student interaction
- Describe the activities going on when you enter the classroom
- How the teacher uses his/her voice, addresses the children
• If the teacher uses the local or the national language
• If the teacher has a written plan for the lesson which is in line with the teacher's guide
• If he/she follows the teacher's guide and the lesson plan (TEP)
• If the teacher knows the subject, teaches correctly
• How he/she motivates the children
• How he/she explains the lesson
• If the language he/she uses is adapted to the students' level
• Describe the teaching methods used
• The use of participatory methods (group work, working in pairs, discussions, other)
• How much the teacher talks compared with the pupils
• If the teacher uses questions and what type of questions
• If and how the students respond to the questions
• If the children ask questions and what type of questions
• How the teacher responds to their questions
• If the children seem eager and willing to learn
• How the teacher manages to catch the children's attention
• If there are some children who do not pay attention. Describe what they are doing
• If the teacher takes notice of these children and what he/she says to approach them
• If the teacher praises children who succeed and how she/he treats those who do not succeed
• If the teacher ridicules children in front of their colleagues
• How the teacher assists weak learners
• If some of the clever students are encouraged to help weaker students
• If the teacher explains clearly wrong and correct answers and makes sure that everybody understands
• The use of peer instructors
• If the teachers ask boys more than girls and how the teacher encourages the participation of girls
• If the teacher treats girls and boys differently
• If the teacher calls upon individual students or ask for volunteers (boys and girls)
• The use of time (how much time is spent on each task)
• How the teacher handles discipline problems
• Corporal punishment
• How the teacher uses didactic material and if the children use didactic material
• If the teacher uses the material in the educational kit correctly. Describe.
• If the teacher has developed her/his own didactic material
• If she/he manages to create a playful and warm learning atmosphere
• If the teacher is friendly and if the children seem to like their teacher and feel relaxed in the classroom. Describe
• If the relationship between the children is friendly and supportive
• If the teacher seems engaged and motivated and likes the teaching and the children
• How the children's performance is being evaluated
• How does the teacher check what the children have learned? (evidence of learning)
• The use of and results on tests and exams.
12 WRITING THE REPORT

The quality of the data gathered during the field work is important, but the results must be presented convincingly in the report. Often, it is not sufficient time to prepare the report.

NORAD and Sida have developed a standard format for evaluation reports as follows:
1. Executive summary. The summary should be brief (2-6 pages) and include findings, conclusions and recommendations.
2. Introduction, describing the background of the evaluation, brief description of the project and its history, the structure of the report.
3. The methodology. Approaches and methods used, the duration of the field work, what was done by whom, when, where and how. Reflections on the limitations of the methodology and the study.
4. Findings. Can be presented in several sections addressing the issues in the Terms of Reference.
5. Conclusions which must be substantiated by the evidence and analysis presented
6. Recommendations (as to future actions)
7. Lessons learned (Sida). A short chapter of a more general nature on lessons that the evaluation have generated.
8. References/literature
9. Appendices:
   - terms of reference
   - list of persons contacted
   - tables, diagrams, statistical information (optional)
   - itinerary
   - Acronyms and abbreviations (often put after the table of contents to facilitate the reading)

Following the above format does not guarantee a good report; the quality depends on the substance. The above-mentioned standard format may not be appropriate for all types of reports particularly not if the data are mainly qualitative and based on case studies. It is however necessary to find a clear structure that fits the report's contents.

A major problem with many reports is that they lack clarity. The structure is not logic, the concepts used are not defined, the language is diffuse or unnecessarily difficult. Another problem is that the writer does not take the perspective of the reader, but takes it for granted that the reader knows the context of the project and the evaluation. If we want the report to be read and understood, the writer must have the reader in the back of her/his mind and describe the project and findings to the reader who does not have any previous information.

The methodological part is often weak. The methods used are not clearly stated and why they were chosen, and the results are not discussed in light of the methodological limitations. Furthermore, findings may be written as general facts without referring to observations and other data that substantiate them.

Tables and diagrams presented are not always clear and explained. They are meant to clarify the presentation and facilitate the reading, not to complicate it. If they don't clarify, drop them.
If the data are qualitative, it may be appropriate to present excerpts from interviews and observations that illustrate the findings. Photographs, maps, and illustrations may also be included.

The use of numerous appendices should be avoided. If the evidence, like tables and diagrams, is attached to the main report, it is not studied seriously. It is better to include them in the text. The text should be complete and understandable without the appendices. The appendices are supplementary information.

The length of evaluation reports varies. Sida recommends 50-60 pages at the most. Often reports are too long and contain numerous repetitions. It is of course better to write a short report that is to the point, but again, the time factor is decisive. In general it takes longer to write a short report than a long. The length of the report also depends on how well the writer has analyzed the data and therefore knows how to separate main findings from minor findings.

It is common to write a first draft of the report which is presented to all stakeholders for comments. The evaluators may be asked to present a draft already before they leave the field, or at least the main conclusions and recommendations. In my opinion this puts a lot of stress on the evaluators who are always in short of time. It is particularly negative to come up with conclusions and recommendations before the evaluators have had a chance to look critically and analytically at the totality of their findings. Instead, the evaluators may gather some of the people they have talked to or observed for an informal meeting where they present some of their preliminary observations and engage them in a discussion.

It is of course important to communicate closely with those who are giving information to the evaluators, but that has to be secured in the evaluation process. When we did the TEP evaluation in Angola, we gathered teachers, supervisors, educational authorities and other stakeholders for informal meetings after the fieldwork and discussed some of our observations with them. We did not pretend that they were final or ready, but asked for their opinion. In this way we could also verify whether our observations were in line with theirs.

Another way of communication during the evaluation process, is to organize a reference group. The reference group consists of people from the organization that has ordered the evaluation and representatives of other donors and stakeholders, who are not directly involved in the evaluation. During our TEP evaluation we had regular meetings with representatives of UNESCO and NRC where we presented our work and preliminary findings, which were discussed. Thus the final report did not contain any big surprises.

Anyhow, a draft version should be presented before the final report allowing them to make comments and corrections. Corrections of factual information should be made by the evaluators, but as a rule, the perspectives, findings and discussions in the report have to be accepted. If the organization that commissioned the report or others have critical comments, they should be added as attachments. Such comments generally enrich a report and highlight what I have mentioned earlier: that there are several perspectives on the same phenomena. Those who have been evaluated may feel that their perspective has not been understood. The Norwegian Save the Children always ask their local program offices for written comments on the evaluations they have commissioned.
Save the children Norway has made an evaluation summary form that presents a short version of evaluation reports. It covers the following:

1. Title and date of the evaluation report
2. File number
3. Description of project (project no. and name, partner org., year stared, expected life span, overall objective, issue codes, accumulated cost, main activities)
4. Purpose of the evaluation
5. Summary of main conclusions and findings
6. Summary of main recommendations
7. Comments by the programme office
8. Description of the evaluation team or consultant
9. Method for information gathering/analysis
10. Costs of evaluation
11. Terms of reference
13 EVALUATING THE REPORT

The contents of the report has to be checked against the issues in the Terms of Reference to find out whether they have been covered, and if not, if the evaluators have given a convincing explanation of why it was not possible to do so. If it is required to use a specific standard format for the report, it should be checked if it has been followed.

It is further recommended to examine the following:
• If the methods used are presented and argued
• If the methods are appropriate
• If limitations of the methods and the study are discussed
• If findings are presented with reference to evidence (data)
• If there is consistency between findings, conclusions and recommendations
• If the language is clear and understandable
• If important concepts are explained
• If there are facts that are not correct
• If references and sources in the text are correct
• If the literature list is correct
• If tables, diagrams, illustrations are clear and easy to understand
• If the structure is logic and easy to follow
• If gender issues (and environmental issues when relevant) are considered
• If the recommendations are realistic
• Whether the report is balanced and analytic and includes critical as well as positive aspects
• If the report is too long. If that is the case, it may be suggested how the draft version can be shortened down

Furthermore, it is important to encourage beneficiaries and local program offices to evaluate the evaluation report. Although a report may fulfill all the above-mentioned requirements, they may feel that it is biased and does not correspond with their perceptions. A written report is an authoritative judgment of a project and although they have been communicating with the evaluators in the process of evaluation, the findings look differently when they are written in a final report. When the evaluators have left, they have to face the conclusions and recommendations and do the hard work.

Some reports are too critical. They focus solely on the negative aspects and forget the positive. Ideally the report should include examples of both successes and failures. If we want conclusions and recommendations to have impact locally, the recipients must recognize ways of working that are successful and not only be told what has not worked.

On the other hand, a report may be too positive and not present critical judgments. The evaluators have to consider how to present critical issues, but the report is not complete if they leave them out. The reason why some evaluators find it difficult to present criticism may be that they are too loyal to the beneficiaries or organization that has commissioned the evaluation or it may be related to cultural issues. It many cultures criticism is conveyed in disguised ways. How to handle this may be discussed with the evaluators before they write the report.
14 EVALUATION STEP - BY - STEP

Here we shall summarize some of the main points by guiding you step-by-step through an evaluation process. The numbers in brackets refer to the chapters.

1. Why evaluate? (chapter 3 and 4)
2. When to evaluate? (chapter 8)
3. What type of evaluation is needed? (chapter 3, 8)
4. The composition and selection of the team of evaluators (chapter 9)
5. Drafting of the terms of reference (selecting criteria of evaluation, elaborating evaluation questions, considering qualitative and quantitative issues) (chapter 5, 6, 7, 11, appendix I)
6. Gender issues (chapter 5)
7. Discussing and finalizing the terms of reference together with the evaluators and other stakeholders (chapter 5 and 6)
8. The evaluators start planning their work (time-schedule, methodology, design) (chapter 7, 10, 11, appendix I)
9. The organization that commissions the evaluation provides all relevant documents regarding the project (chapter 10, 10.2.1)
10. Fieldwork (chapter 10, 11, appendix I)
11. Continuous dialogue with reference group, beneficiaries and stakeholders (chapter 10, 12)
12. Presentation and discussion of preliminary observations (chapter 12)
13. Presentation of draft (chapter 12)
14. Comments to draft (chapter 12)
15. Finalizing the report (chapter 12)
16. Evaluation of the report (chapter 10.2.7, 11, 13, appendix I)
15 REFERENCES

APPENDIX I

List of questions related to the checklist

1. Understanding the context (cf. Hallam 1998, annex III)
   • How was the initial situation assessed? What data was collected, and how was it analyzed?
   • Have contingency plans been made for new movements?
   • Are other agencies included in analysis of the problem and the drawing up of response plans? Are response plans "holistic" or very much determined by the mandate of individual organizations?

2. Specifying objectives, activities, outcome, priorities, restrictions and beneficiaries (cf. Hallam)
   • Does the operational/conceptual model/logical framework underlying the intervention make sense?
   • Are the program objectives sufficiently clear and explicit?
   • Do the activities correspond to actual needs?
   • Will the activities proposed achieve the desired outputs?
   • Will these outputs lead to the desired changes (impact)?
   • Are constraints to implementation adequately assessed from the outset?
   • How rapid is rapid?
   • How have beneficiary groups been identified?
   • Are the special needs of certain groups considered?

3. Management and cooperation
   • How does the cooperation function when rapid action is needed?
   • What happens if the agency is not able to act rapidly?
   • How can different agendas (between agencies, between agencies and government or between government and other national forces) complicate the response?
   • Have measures to mitigate the negative impact of different agendas been considered?
   • Is the best-placed organization (in terms of knowledge of the country, skill profile, capacity etc.) used to carry out the program?
   • Are the activities well coordinated?
   • Are plans well communicated (to other partners, governments and beneficiaries)?

4. The selection of target groups/beneficiaries
   • Does the project reach the target group and how is this being controlled?
   • Who selects the target groups?

5. The selection of teachers
   • Who decided the criteria for the selection of teachers?
   • Who controls that the criteria decided upon are being followed?
   • Are teachers from some political groups/ethnic groups being favored over other groups?
   • Do the teachers need to pass an academic test before they are accepted in the program?
   • Who makes the final decision regarding the selection of teachers?
6. **The schools/classrooms**
   - How many schools/classrooms were planned?
   - How many are functioning?
   - How many take place under a tree/in a tent/in existing buildings/in new buildings?
   - How many schools were planned to be constructed/rehabilitated?
   - How many have been finished?
   - When should they have been finished?
   - Do they have roofs and windows, desks?
   - Who are constructing/rehabilitating the schools?
   - Who pays the costs involved in the construction/rehabilitation of schools?
   - What type of remuneration and/or incentives do the constructors get?
   - What are/have been the main challenges regarding the construction of schools
   - The use of local schools - advantages and disadvantages

7. **The teaching**
   - Does the program follow the national curriculum or a locally developed curriculum?
   - Have the needs of war traumatized children/child soldiers been assessed?
   - Which subjects are taught during the week and for how many hours?
   - Which subjects have textbooks and which have not?
   - How many of the children in the class have textbooks and in which subjects?
   - Are the textbooks free of charge or do they have to pay for them?
   - How much is the cost of a textbook? Are they available?
   - What is the quality of the textbooks? (layout, text, content etc.)
   - Does the teacher have textbooks in each subject?
   - Does the teacher have a teacher's guide/manual in each subject?
   - The quality of the guide/manuals
   - How many of the children in the class have notebooks?
   - How many have pencils/pens?
   - Does the teacher have chalk?
   - Does the teacher use a lesson plan?
   - Is the plan in line with the teacher's guide?
   - Which type of ready-made didactic material is in use?
   - Which type of local didactic material is in use?
   - Is there a need of an educational kit?
   - How many of the teachers have received the educational kit?
   - How many have not?
   - Do they receive it when they start teaching?
   - Are there examples of delays?
   - What do they do in the meantime if they have not received it on time?
   - Why do the kits not arrive on time?
   - Was the kit complete when it arrived?
   - If not, what was missing?
   - What do the teachers do if the kit is not complete upon arrival?
   - Who is responsible for supplementing/controlling the kits in use?

   *See guidelines for classroom observation above for more questions regarding the quality of the teaching*

8. **Extra-curricular activities**
   - Are there any organized recreational activities after ordinary school hours?
• Which activities?

9. **The teachers**
   • How many teachers (exclusive of supervisors) are working in the program?
   • How many were working when the program started?
   • How many teachers are still needed?
   • How many teachers have a teacher training background?
   • What type of teacher training and for how long?
   • What is their academic level?
   • How many have no teacher training?
   • What is their academic level?
   • How many of the trained and untrained teachers are refugees/IDPs?
   • Have the teachers' training needs been assessed?
   • Are there teachers traumatized by the war and how does this affect the teacher-student interaction?
   • Is the teaching organized in shifts? How many?
   • How many shifts does the teacher teach per day?
   • How many hours?
   • How many classes daily and weekly?
   • Does the teacher teach other classes in the regular school?
   • How far is the distance between the school and each teacher's residence?
   • Do the teachers have means of transportation? Which?
   • How much does a teacher in the program (with teacher training) earn compared with teachers who have a similar background in the ordinary school?
   • Is there a difference in remuneration between those who are teacher trained and those who are not? What is the exact difference?
   • Do the teachers receive incentives? What type of incentives and who distribute them?

10. **Teacher training**
    • Does a plan/curriculum for the training of teachers exist?
    • How many teachers have totally been trained and how many are in the process of being trained?
    • How many hours of training do they get and for how may weeks?
    • Which subjects are covered in the training?
    • Which pedagogical and psychological topics are part of the training?
    • Are "children with special needs/disabled children" part of the training?
    • Are "Children traumatized by the war" part of the training?
    • Describe the teaching methods used in the training course. How much time is approximately spent on each teaching method?
    • Who is responsible for the training?
    • What are the qualifications of the trainers?
    • How have they been selected?
    • Who trains and supervises the trainers?
    • Have educational and didactic material been developed to accompany the training?
    • Are the teachers trained to use this material? Approximately how many hours during the training are devoted to the use of didactic material?
    • Is the material distributed to the teachers during the training?
    • Are the teachers trained in how to develop local didactic material?
    • Are the teachers trained to find solutions to unexpected educational problems?
    • Are the teachers trained in how to plan a lesson?
• Are the teachers asked to present their own experience and educational material they may have developed?
• Is it ensured that their ideas and material are collected and brought to the attention of the project's education advisers?
• Is there a teacher's guide that is distributed to the participant teachers?
• The quality of the guide?
• Does the guide tell the teachers how to plan a lesson?
• Other material which is distributed?
• Is the teachers trained in how to collect statistical information and write reports?
• Describe the main challenges during the training course
• How many of the teachers who entered the training program pass the final exam?
• How many leave the training and for what reasons?
• Are all the teachers who finish the training course offered a job in the program?
• Describe main changes and turning points in the training program

10. Monitoring and supervision of teachers
• Does the program have a monitoring/supervisory component?
• Who decided the criteria for selection of supervisors?
• Who controls that they are followed?
• Who selects the supervisors?
• How many supervisors are working in the project?
• How many were working when the program started?
• What are the supervisors academic background/training/experience?
• How many schools/classes does the supervisor visit each month?
• How many schools were visited last month?
• Are there schools/classes that have never been visited?
• How often does each school/class meet a supervisor?
• How long does the supervisor's visit last in each school/class?
• Does the supervisor have access to means of transportation?
• If motorized vehicle, who pays the fuel?
• Distance between the supervisor's residence and the schools/classes
• How is the supervisory system organized?
• To whom do the supervisors report?
• Who controls the supervisors?
• Do the supervisors have other jobs to attend to?
• What does the supervisor do during the visit? Is there a job description?
• How much time does the supervisor spend on each task? (control of attendance lists, control of tests and exams, observation, conversation with teacher, inspection of the school and the class, pedagogical support, other)
• Are there guidelines for supervision? Guidelines for observation in the classroom? Are they used?
• What are the most common problems the teachers present?
• What do the teachers master/not master?
• Are there guidelines for supervisor's reports?
• How often do the supervisors write a report?
• Who gets the report?
• Have the supervisors been trained in how to record statistical information?
• Do the supervisors get feedback on their report?
• Are the reports summarized, analyzed and presented to higher levels in the project organization?
• Do the supervisors get incentives or a higher salary than the teachers?

11. Management and administration of teachers/supervisors/schools

• When did the program start?
• Who initiated the program?
• How was the program announced?
• How many schools/classes exist totally today?
• Where are they located?
• What is the distance between the center of the district and the different schools/classes?
• Is there a long-term plan for the number of teachers needed?
• Is there a plan for the transfer of teachers to the government's payroll?
• How many teachers have been transferred?
• How is the placement of teachers decided?
• How are the daily management and control of the schools/classes/teachers carried out?
• Who is responsible for the distribution/supplement and control of the educational kits?
• Who is responsible for the statistics of the program at national, provincial and local levels?
• Are there guidelines and instructions as to how to record statistical information?
• Have the people in charge at national and local levels been trained in how to use the guidelines and instructions?
• How many children have been transferred to regular schools?
• To what grades have they been transferred?
• Where are those who have not been transferred?
• Is information on the children's progress in ordinary schools available?

12. The children

• How many children were enrolled in the program in the beginning of the school year? (based on information from each class)
• How many should have been enrolled?
• How many entered during the school year? (in each class)
• The age distribution in each class
• The number of children in each class today
• How many have dropped out during the school year and for what reasons?
• At what time of the year did they drop out?
• Have the drop outs been reported?
• Have the local education authorities asked about the drop out rate and the reason for it?
• How many children presented themselves for the tests during the school year?
• How many did/did not sit in for exams? Reasons?
• How many passed/failed? (broken down by each subject)
• How is the distribution of marks in each class (in each subject)
• Are there children with special needs/disabilities/traumatized by the war in the classes?
• How have they been detected?
• How are these children's needs being met?
• The distance from the children's residence to the school

13. Development of plans/curricula/education material

• Describe main points in existing plans and curricula and how they relate to the national curricula
• Consider the quality of the plans/curricula
• Quality of textbooks
• Describe how the plans/curricula and textbooks have been developed
• Describe existing didactic material and their quality
• Describe how they have been developed

14. The parents and the local community
• Are there PTAs (parent-teacher committees)?
• What are their tasks?
• Do the parents visit the school regularly?
• Does the teacher have meetings with the parents?
• What are the subjects discussed with the parents?
• Is the local community engaged in the school? In what ways?
• Do the parents contribute financially to the running of the school?
• Are the parents or other members of the community invited to visit the schools?

15. Other evaluation questions
• Were plans followed? If not, why?
• Were the objectives clear enough and were they relevant?
• Have the objectives changed during implementation? Why? On basis of what analysis and what events? Describe the major changes and turning points.
• Did the activities meet the target population's essential needs? Were needs overlooked or neglected? Describe.
• Have the activities changed? In what direction and why?
• Were the interests of beneficiaries adequately taken into account in the design and implementation of the program?
• Have the activities discouraged initiative and led to dependence?
• Are any groups (unintentionally) favored over others as a consequence of the way the program was set up?
• Are beneficiaries unfairly assisted in relationship to host/other populations?
• Do the program plans take into account longer-term (re)-integration of beneficiaries?
• Have the needs of girls and female teachers, been taken into consideration?
• Could activities have been more cost-effective?
• Were measures to provide security and protection adequate?
• Were those most in need prioritized for assistance?
• Will activities continue once agencies pull out?
• Were resources effectively and efficiently used? Do the results justify the costs?
• Were is improvement needed?
• Unintended positive/negative effects