

# Evaluating community based psychosocial programmes: why, what & how?

***Guus van der Veer***

*Evaluation of psychosocial programmes can be carried out for a variety of reasons. It is the nature of these reasons that determines what is exactly evaluated, what criteria and which methods are used. In this article, the focus is on evaluation as a pathway to learn from experience and develop expertise. Some evaluation criteria are discussed, and a step-by-step plan is described.*

**Keywords:** effectiveness, evaluation criteria, impact, indicators

## **Why do we want to evaluate psychosocial programmes?**

There are several reasons to carry out evaluations. Sometimes these reasons are externally motivated (by donors for example) and sometimes they are driven by the wish of the project implementers themselves. Evaluations often have two elements: accountability and lessons learned from the experience.

*Accountability.* Community based psychosocial programmes cost money. The people making the funding available want to know whether the funds have been well spent. *Well spent* can be translated into two questions: 1) have the planned activities been carried out, and 2) have the activities resulted in the desired effects? In the latter query we are referring to accountability. *Accountability* is one of the most important reasons for evaluation. The donor wants the people carrying out the psychosocial intervention to account for what they have done. Some donors can

be satisfied by telling them that the activities have been carried out in line with the project proposal, and by describing a few case histories that illustrate the impact of the psychosocial intervention on the life of the beneficiaries. All we have to do to convince such donors of the effectiveness of the project is to interview a few beneficiaries and write down their stories. Other donors want a different type of evidence. They want to see quantitative results that can be analysed through statistical methods. This kind of evidence is generally considered 'harder' evidence than the evidence gathered from a few case histories. The call for so-called hard evidence seems to become louder and louder with each passing day. I find it interesting that double standards are often applied for psychosocial programmes in Western countries and psychosocial programmes in areas of armed conflict. For example: during the 35 years I was attached to mental health institutions in the Netherlands no one ever evaluated the effect of the interventions offered by my colleagues and I, although these centres cost millions of Euros per year. However, if the non governmental organisation (NGO) I work for supports a psychosocial programme for children affected by war in the Palestine territories, the donor wants *hard evidence* that the invested 50,000 Euros have resulted in effective interventions.

*Lessons learned.* Another important reason for evaluating psychosocial programmes is that

we want to learn from our experience. Evaluation can help us to develop our practical know-how on how to carry out psychosocial interventions, within a particular context, in a way that makes a real and tangible difference to the beneficiaries. Evaluation is then part of the process of developing contextual expertise. In this type of evaluation, it is important to know exactly how the activities have been implemented. For example, we do not learn much if we only report that we have successfully trained a particular target group. To learn something, we need to report how the training was carried out. We should also include the content of the training in detail: what subject matter was introduced during the training and which knowledge and skills were featured in the curriculum. We also need to report which educational methods were used for different parts of the curriculum, and how the participants received the various items in the training programme during the training (Box 1).

**Box 1: Report of a training (format)**

Describe all items in the training programme. Per item, specify:

*Name of the training item:* (topic, title)

*Purpose:* (in terms of knowledge, skills, or attitude)

*Content:* (in terms of information or concepts discussed)

*Method:* (describe what the trainer did during this item and quote the instructions given to the participants)

*Interaction:* (describe how the item was received by the participants)

The above format sums up what is minimally required to report interaction during training.

**What to evaluate?**

Evaluation reports on psychosocial projects show great diversity with regard to the matters that are evaluated; different evaluators seem to use diverging criteria for judging whether a project is a success or a failure. One criterion that is considered to be important is *efficiency*. In order to say something about efficiency we have to measure the outputs – both qualitative and quantitative – achieved as a result of inputs. This generally requires comparing alternative approaches to achieving a particular output, to see whether the most efficient approach has been used. The question we seek to answer is: did the implementers do it correctly and cheaply? *Cheaply* means: with as few resources as possible; as little effort as possible, in as little time as possible, with as little money as possible, with as few people as possible, and as few materials as possible.

Another important criterion is *effectiveness*. Effectiveness is about the extent to which an activity achieves its purpose, and whether this can be expected to happen on the basis of the outputs. The question we seek to answer is: did the project achieve its objectives? In other words: did the implementers do the thing properly?

A third criterion is *impact*. Impact looks at the wider effects of the project on individuals, gender and age groups, communities, and institutions. Impacts can be both those that are intended and/or unintended. The impact can be positive and/or negative. The impact can be macro and/or micro. For example: a project carried out in a few schools can have an impact on the whole of the educational sector; that is a macro effect. A programme aimed at children may have an impact on the household in which the child is living, that is a micro effect. The question here is: have the implementers not only done the right thing, but also brought about profound

changes in the situation or in the longer term? So impact includes *sustainability*. Sustainability measures whether an activity or an impact is likely to continue after donor funding has been withdrawn.<sup>1</sup>

The fourth criterion I would like to mention is *relevance* or *appropriateness*. Relevance assesses whether the project is in line with local needs and priorities. In other words: have the implementers of the project been doing the right thing within the context of that community?

In sum, many different things can be evaluated. We have to make choices, and we may have to set priorities. When we start to evaluate a project, we have to transform each criterion that we choose into one or more *core questions*. In relation to these core questions we have to define *key indicators* and to decide about *data collection*.

### **An example**

*The project.* An NGO, IBISS (an organisation that fights the impacts of severe poverty and illness on children in Rio de Janeiro), wants to help children and youth between 8 and 23 years of age that have been recruited as 'soldados' (soldiers used by drug traders in Rio de Janeiro). There are currently about 12 500 of them. They have seen a lot of violence and were often forced to kill others. Of the *ex-soldados* seen by IBISS, 54% suffer from severe trauma related symptoms like nightmares and anxiety attacks.

These children refuse to talk about their experiences and feelings, but they feel able express their emotions in various activities, such as hip hop music. The local mental health institutions have no experience in dealing with them. The psychosocial workers of IBISS want to experiment with a variety of means of expression that are attractive to the target group, and thereby develop a form of therapy.

So, a training team will be formed. This training team will start sessions with 50 *ex-soldados* and develop a form of therapy. They will then put together a curriculum for training 12 psychosocial workers in carrying out this form of therapy. This training will then be evaluated and adapted. Subsequently, another 12 psychosocial workers will be trained. After this, the impact of the first training course in the field will be evaluated. Afterwards, the curriculum of the training will be finalised and a training manual published.

*Efficiency.* With regard to the criterion of efficiency, a *core question* could be: is this approach cheaper or more expensive than other approaches (e.g. psychiatric treatment) used for working with this kind of target group? The cost of the IBISS therapy then would be one *key indicator* and the local costs of individual psychiatric assistance would be another of the key indicators. *Data collection* could take place by analysing the IBISS budget and comparing it to the local rates of psychiatric facilities.

*Effectiveness.* Thinking of the criteria of effectiveness, we could ask the following *core questions*: did the change in approach of the psychosocial workers after the training result in a decrease in symptoms? Do more *soldados* go to school or have a job than before the training? *Key indicators* would be the symptoms as reported by the *ex-soldados*, the aggressive behaviour reported by psychosocial workers, and the rate of school attendance or employment as reported by the *ex-soldados*. *Data collection* could take place through individual interviews as well as through focus groups with both *ex-soldados* and psychosocial workers.

*Impact.* With regard to the impact criterion, *core questions* could be: do family members report a positive change? Do other community members report a positive change? Has the crime rate decreased? Are the *ex-soldados*

doing more work for the community? *Key indicators* would be: opinions of family members, police statistics, and opinions of community leaders. *Data collection* could take place by interviewing family members, local authorities and community leaders.

*Relevance.* With regard to relevance, a *core question* would be: how does the IBISS approach concur with what the literature suggests about helping *ex-soldados*, or youth with a similar background? *Key indicators* are what the professional literature has written about helping young people with similar problems to the *ex-soldados*, such as former child soldiers and delinquent adolescents. *Data collection* could take place by collecting and reviewing relevant books and articles. For example, the literature on former child soldiers suggests that the frequency and intensity of trauma related symptoms is very much dependent on the amount of stress in the current life situation of these former combatants. Support, therefore, must have various components aimed at dealing with current stress, such as support in removing stress factors in the community and skill training in stress management. Stress management training could include relaxation exercises, as well as exercises for prevention and control of inadequate reactions to stress such as aggressive outbursts, and mobilisation and strengthening of protective factors in the social environment. In addition, psycho education about trauma related symptoms and advice on coping strategies for these symptoms has proved to be effective. The literature on delinquent adolescents suggests that training in social skills, practical skills (ranging from arithmetic to solving conflicts through negotiation, and dealing with one's own emotions when overwhelming), as well as impulse control is usually more effective than expression of emotion and exploration of traumatic childhood experiences.

## **How to carry out an evaluation?**

*Start planning the evaluation before the projects starts.* I strongly believe that evaluation of psychosocial projects has to be planned before the projects starts. We can distinguish three activities that are components of evaluation: monitoring, making explicit descriptions of activities, and assessment of effects. The first two activities always have to be carried out from the very first day.

*Monitoring* is an ongoing process during the course of the project and a method of checking whether activities are indeed completed as planned (in terms of coverage, duration and other quantitative criteria) and also examines the causes of each and any deviation from the plan.

*Making explicit descriptions of the activities* is also an ongoing process, during which a manual is developed describing the details of each activity that forms an individual part of the project. Explicit descriptions are necessary if one wants to replicate the project. It can also contribute to the sustainability of the project after the initiators have left. It is also a prerequisite for any significant assessment of effects: it is useless to know that an intervention is effective if one cannot describe that intervention in detail.

*Measuring effect and impact* can be derived from two fundamentally different approaches. One approach focuses on effects in terms of the mental health of the individual people that are reached by a psychosocial programme. *Effect* is measured in terms of satisfaction of the client with the services delivered to him/her, through questionnaires covering complaints, symptoms and ways of coping with symptoms. It is easier to prove effect with baseline data on these measures, so that means the effect measurement may also have to start on day one. The other approach focuses on variables that represent the quality of community life, such as the

presence of basic conditions for the normal health development of children, the availability of traditional forms of collective coping with stress and trauma, the presence of new forms of collective coping, new institutions that replace traditional ways of coping, and so on. Effect is measured with so-called participatory tools: the target population is involved in the planning and development of the project and in determining the desired effects of those planned activities. Examples of this approach that also start at day one have been published in *Intervention* (Bragin, 2005; Hart, Galappatti, Boyden & Armstrong, 2007).

When planning an evaluation there are a series of steps involved that are discussed below.

*Step 1: deciding the purpose.* Step one is discussing the overall purpose of the evaluation. Is the evaluation for learning lessons from the experience, or for accountability, or both? We need to formulate the overall purpose in a core question. We need to decide how and by who the results should be used. We need to make a plan for dissemination of the results.

*Step 2: identifying the stakeholders.* When we start planning an evaluation, it is helpful to discuss which persons and organisations are involved in the project, and which persons or organisations may be affected by the project. We can represent these stakeholders on a map: who are central to the project and who are in the periphery?

*Step 3: choosing a type of evaluation.* There are many types of evaluation.<sup>2</sup> With regard to a single training, I strongly believe in *real time evaluation*: an evaluation of the events during the training as they unfold. Another type involves someone in the implementing organisation carrying out an *internal* evaluation. This method has both advantages and disadvantages (Box 2). Somebody not previously

### **Box 2: Internal or external evaluation – advantages and disadvantages**

*Advantages of using internal evaluators:* they know the organisation, understand organisational behaviour and attitudes, are known to staff, are less threatening, often have a greater chance of adopting recommendations, are less expensive, build internal evaluation capability, and contribute to programme capacity.

*Disadvantages of using internal evaluators:* their objectivity may be questioned, their personal gain may be questioned, they accept the assumptions of the organisation, full participation in the evaluation may be constrained by normal workload, they may not be trained in evaluation methods, they may lack special technical expertise, their involvement may lead to the evaluation not having acceptable levels of outside credibility and they may have difficulty avoiding bias.

*Advantages of using external evaluators:* they are more objective, may have fresh perspectives and a broader experience, they can serve as an outside expert, they are not part of the power structure, they can bring in additional resources, they are trained in evaluation and experienced in other evaluations, they are regarded as an *expert*.

*Disadvantages of using external evaluators:* they may not know the organisation, they may not know of constraints affecting recommendations, they may be perceived as an adversary, they are expensive, contract negotiations may take time, they often cannot provide follow up on their recommendations, and they are unfamiliar with the environment.

known to the implementing organisation carries out an external evaluation.

When evaluating a particular project as a donor, a *joint evaluation*, that is an evaluation conducted jointly with the implementing organisation, sometimes may be indicated. In other situations, an evaluation by the donor only is indicated.<sup>3</sup>

Participatory evaluation is an evaluation that is carried out with, or by, the primary stakeholders, usually the project beneficiaries.

### Evaluation instruments

A project can be evaluated by interviewing the direct and indirect beneficiaries and/or involving beneficiaries in focus group discussions (Olij, 2005) about the desired effects of the various activities carried out during the project on individuals. Traditionally, the interviewers usually were looking for signs that symptoms, complaints and disturbed or dysfunctional behaviour of individuals have been diminished, as well as for signs that indicate a positive development (like participation in youth clubs).

Recently, researchers have become more interested in evaluating the effect on communities, by using *so-called* participatory tools. Members of the community are involved in discussing the consequences of the conflict for the local community and for the functioning of young people in particular (Bragin, 2005; Hart et al., 2007). This approach focussed more on community coping strategies.

For the evaluation of single activities (such as a training of psychosocial workers in empathic listening, or a training of trainers on the use of interactive participant oriented training methods) one can use scales. Many trainers use scales that measure satisfaction of the participants; these scales, of course, do not measure behaviour change as a result of the training. If the goal of training has been

described in terms of details of behaviour in practice after the training, Goal Attainment Scales can be constructed.<sup>4</sup> These scales have to be filled in by judges who observe the participants both before and after the training.

### Conclusions

Evaluation of psychosocial projects can be done in many ways, using diverging criteria, and for different reasons. An evaluation plan should be based on explicit choices with regard to purpose, criteria and evaluation methods. In most cases, evaluation can best be planned before a project is started. An evaluation report ends with recommendations, but an evaluation process may continue until all the recommendations have been implemented and the results of this have been observed.

### References

- Bragin, M. (2005). The Community Participatory Evaluation Tool for Psychosocial Programs; a guide to implementation. *Intervention*, 3 (1), 3-24.
- Hart, J., Galappatti, A., Boyden, J. & Armstrong, M. (2007). Participatory tools for evaluating psychosocial work with children in areas of armed conflict: a pilot in eastern Sri Lanka. *Intervention*, 5 (1), 41-60.
- Olij, J. (2005). Trauma awareness, healing and group counselling in secondary schools *Intervention*, 3 (1), 51-56.

<sup>1</sup> Many humanitarian interventions, in contrast to development projects, are not designed to be sustainable. They still need assessing, however, in regard to whether, in responding to acute and immediate needs, they take the longer term into account. Larry Minear has referred to this as connectedness; the need 'to assure that activities of a short term emergency nature are carried out in a

context which takes longer term and inter-connected problems into account.

<sup>2</sup> Source: Development Co-operation Directorate (DAC) (2001) Glossary of Evaluation and Results Based Management Terms DCD/DAC/EV (2001) <sup>3</sup> Working Party on Aid Evaluation, Paris: OECD; see <http://www.oecd.org/dataoecd/29/21/2754804.pdf>.

<sup>3</sup> For example, when there is suspicion about the truthfulness of reports send in by the implementing organisation.

<sup>4</sup> see for example <http://www.regional.org.au/au/apen/2006/refereed/5/3068lukiesr.htm> or <http://www2.uta.edu/sswmindel/S6324/Class%20Materials/measurement/Goal%20Attainment%20Scaling%20presentation.pdf>.

*Guus van der Veer, psychotherapist and mental health consultant, is Editor in Chief of Intervention. email: [guusvanderveer@hetnet.nl](mailto:guusvanderveer@hetnet.nl)*