

Feedback from local staff

The pitfalls of psychosocial evaluations: a critical perspective from a field worker

Anica Mikuš Kos

Evaluations of psychosocial projects are meant to contribute to better projects. However, in practice, the evaluation process, in particular when done by external evaluators, can pose its own difficulties. Based on a wide field experience, the author presents arguments about how evaluations can cause problems, and even produce negative effects in project staff and recipients of assistance. Psychological processes triggered by external evaluators, or appearing in the process of evaluation, can create uneasiness in field workers. Also, some evaluation methods have questionable cultural acceptability. The author argues that evaluation designers, publishers of evaluation reports and researchers have an ethical and social responsibility.

Keywords: evaluation, psychological obstacles, ethical problems

Introduction

A thorough evaluation of psychosocial programmes is considered a legitimate requirement of donors and decision makers. It is seen as an important learning tool, and regarded as a mechanism for improving the programmes, as well as a treasure box for data serving research. It is also important as a

feedback to field workers informing them about the impact of their efforts towards the assisted population. Evaluation also helps to answer questions, such as; ‘would it not be more justified to spend money feeding hungry children or covering other basic needs?’

Positive evaluation feedback has an important motivational impact on implementers. Psychosocial helpers usually work in difficult circumstances, under high stress and can suffer emotional distress. The emotional aspect is important; it is encouraging to know that one’s work is perceived as useful, and that there is some evidence that your programme is well done and makes a difference.

Evaluation also fosters a culture of critical reflection. Honestly, we as field workers, have a tendency to embellish the positive results of our work, mainly for emotional reasons, but sometimes we are also guided by a desire to obtain funding for the continuation of the programme.

Community based psychosocial programmes trigger complex processes and are supposed to have broad effects. Even in so-called normal situations, it is not easy to evaluate psychosocial programmes in a scientific way. In areas affected by major disasters this becomes a huge challenge. In circumstances

of armed conflicts or in its aftermath, there are specific obstacles affecting the feasibility of evaluation or that may act against a productive cooperation between field workers and evaluators.

I have 15 years of field experience in territories affected by armed conflicts. I have also experience in both *'being evaluated'* and in *'being an evaluator'* of community based psychosocial programmes. The aim of this feedback is to reflect some concerns and reactions triggered in field worker by the evaluation procedure, which can prevent productive cooperation and hamper the process, cause unease, and even affect motivation in field workers.

The discourse presented in this feedback concerns mainly the evaluation procedures performed by external evaluators. I consider that the internal evaluation process should be built into each programme design. However minimal, it should combine qualitative and quantitative methods. The built in internal evaluation is not only a safety measure for the protection of the quality of the programme, the process has an important immediate impact on workers through making the situation more transparent and reducing their doubts and uncertainty. It is advisable to combine the internal assessment with psychological support to the staff, although it might not please scientists who believe that that the evaluation should be an autonomous process detached from emotions of involved human beings.

I want to make it clear here that I am not *'against'* external evaluation of programmes. By describing difficulties, or as somebody once said; *'horror stories in evaluation'*, I only intend to contribute to bridging of the gaps between the world of field workers and that of scientific evaluations. I will present a list of arguments. Some of them might seem irrational, but this should not diminish their importance.

The standards for evaluation of psychosocial projects are unfairly high

During my 35 year clinical career, my work has never been so rigorously evaluated as it was in the framework of psychosocial programmes, sometimes worth only 40 000 Euros. When I hear persistent requirements for scientific evidence that psychosocial interventions are helpful, I wonder about the evidence of various psychotherapeutic methods in general. Indeed, we, the mental health professionals, work for years spending huge sums of money on our services, using fashionable therapeutic methods without any evaluation of the impact of our work. The assumption is that theories and practices that are officially accepted by the professional community are reliable enough to be applied without evaluation. We, as mental health workers (at least the older ones) can remember many theories that we preached and practiced and were in the course of time found not to be efficient. For example nowadays the value of critical incident debriefing is very much questioned. I wonder why psychosocial projects in difficult circumstances are comparatively more harshly evaluated than is usual in other areas of mental health care.

Some evaluation methods could be questioned

Let me give two examples. The first is the presentation of cases of successfully treated children. Cases of recovery influenced by the programme are convincing and frequently used illustrations of its quality. We all can find beautiful cases of improvement of the inner mental situation or functioning of our *'clients'*. However, how representative are individual stories when applied for the whole *'target'* population? One of the crucial questions for an evaluation is; *'how many*

people have been helped within the programme? Quantity is often the basis of evaluation of the economy of the programme, i.e. how much money is spent to help how many people?

Another evaluation method where value is sometimes overestimated is the questionnaires filled in by participants of different workshops, training and other educational or psycho-educational activities. According to my experience, and I have run at least 500 workshops and trainings for teachers, medical workers, mental health workers, volunteers, etc. All of those resulted in the average mark 4–5, given by participants in a scale in which the mark 5 was the highest, indicating *‘excellent’*. Questions in the evaluation questionnaires were about satisfaction, gains in knowledge, applicability and other issues usually asked at the termination of psycho-educational events. There are many possible explanations of this high appreciation: participants were indeed so satisfied; they felt obligated to praise the organisers, lecturers and/or group leaders; through high marks they expressed their gratitude that somebody is caring for them; or it is mentally the most comfortable to assign high marks. Such questionnaires do not actually tell us much, if anything, about the implementation of acquired knowledge and know how.

Can we measure what really is important?

Even with a well planned evaluation design, we will only be able to measure a part of effects of the programme on the target population and on the community as whole. According to Adjukevic (2008) *‘the assumption is that project interventions should yield observable or measurable outcomes (consequences, results) that can be used to demonstrate through evaluation indicators that the project is reaching the goal and objectives.’* Are all impacts of psychosocial

programmes measurable? The inner state of mind (sorrow, suffering) is not measurable. You can measure symptoms, but you can be very sad without having any of the symptoms listed in medical classifications. Suffering is not a measurable symptom of disorders, although alleviating suffering should be the top priority of programmes in disaster circumstances. Suffering can be accompanied by withdrawal, apathy, but this is not a necessity. In Bosnian refugee children, professionals were astonished by the lack of such symptoms, and good social functioning. Sadness showed in the children’s eyes when they heard a Bosnian song, or when they talked about their longing for relatives, friends, houses, domestic animals, and when they spoke about their sadness and suffering. A comparative study measuring symptoms of depression in Slovenian and Bosnian refugee children showed a higher rate of symptoms of depression in Slovenian children than Bosnian ones. One could hardly imagine that Slovenian children have suffered more other refugee children. However, evaluation is usually focused on symptoms. Evaluators neglect phenomena covered by *‘vague’* terms such as sorrow and suffering. Field workers often consider such an attitude as disregard for affected people, lack of understanding and lack of empathy.

Of course, one cannot evaluate all impacts of the programme, but the choice matters. Many evaluation methodologies concentrate the impact of the programme on symptoms of traumatisation, neglecting less *‘scientific’* categories of benefits such as joy, increase of activity, or empowerment.

The view of the beneficiaries is not given enough priority

There are differing views concerning the quality of the programme in evaluators, implementers and the target population.

For instance, a Danish programme in one of the poorest villages of Kosovo combining some psychosocial activities in schools (with teachers and children) with giving a sandwich three times a week to all children, was remembered as something very helpful by local people long after its termination. I do not know whether the programme was evaluated, but I have never heard any other psychosocial programme to be so much praised by local people, children, parents and teachers. A project can be very useful without obvious results in fulfilment of its goal, for instance reducing symptoms of trauma. It can simply demonstrate to children that somebody is caring, that there is goodness and humanity in the world as opposed to the man made evil to which they were exposed, restore hope and make life more bearable. It may be popular, demanded, and appreciated by the local population, yet it still may have great difficulty showing that its approach is effective regarding the goals defined by the project proposal.

Sometimes the assisted population becomes reluctant of attempts to be 'evaluated'. This happens primarily when people have the feeling of being used for the scientific promotion of researchers and when they get fed up with answering questions. When we gathered a group of socially well functioning Bosnian youth who had experienced important losses and traumatic events, during their third year of refuge in Slovenia, considerable efforts had to be invested to convince some of them to participate in a group work aimed at identifying protective factors contributing to their resiliency. The procedure also served as an attempt to evaluate the impact of psychosocial programmes. The Bosnian youth wanted in depth explanations and answers to questions such as for whom and to what purpose would the collected information serve. Political suspicion can be

another obstacle for gathering data. All field workers know that in some circumstances it is better not to ask some questions. My example is from Chechnya. Running a programme for Chechen medical workers, I wanted to gather some baseline information about traumatised of the population. Very soon I could sense that my questions were not welcome. Some people asked me; *for whom are you spying? For the Americans or for the Russians?* Such things are more likely to happen at the beginning of the programme, when trust is not yet established. Therefore, it is sometimes not advisable to insist on collecting baseline data. And finally, there are so many costly screenings and other data collection methods, which should assist in action planning and as a baseline for evaluation of interventions, which, in the end, do not serve anything.

Evaluations often measure only short term effects and neglect longer term impacts, indirect effects and spin off effects

Impact is an elusive aspect of evaluation. First, impact frequently requires time, and may not necessarily happen within a project's time frame. Some effects will appear only in a time distant from the period of evaluation. This happens on an individual level and on a broader community level. Social processes in communities usually have a latent period and an 'underground' phase in which they cannot yet be detected, and become observable only after a lapse of time. In most cases, we measure the impact immediately after the termination of the programme, or even during the run of the programme, so no one really knows what are the long term impacts. There is also the overlooked indirect impact of interventions, which is difficult to accurately document. Synergy is an important interactive process producing unexpected

effects. Some subthreshold effects of the programme can contribute in combination with some other influences – accidental circumstances or systemic changes totally unconnected with the programme, – to considerable positive changes. An intriguing question is whether positive changes caused by the intervention are due to a specific component which is qualified as ‘*professional techniques or therapeutic methods*’ or to non-specific ‘*just*’ human and social attributes provided by the programme, such as care, concern, empathy, etc.

Sometimes very important interventions are not evaluated well enough. For instance, after the Beslan tragedy, European countries guided by empathy and other good intentions, invited many children to spend their holidays within those countries. They did everything possible to cheer the children up and to assist them in overcoming a horrifying and traumatic experience. However, now, three years after the event, you can hear in North Ossetia Allania that the grief process for children and families might have been shorter if the children had not been detached from their family and community during the phase of most intense distress.

Sometimes, positive effects stemming from other social or material influences, or from natural recovery processes are attributed to the programme. On the other hand, a programme may seem without particular positive impact because of counteracting new negative influences and to deteriorating circumstances and life conditions. In such cases, the question should be posed: ‘*what would happen, how much and how many children would be affected by adversities without the programme?*’ The question is not easy to answer. Evaluation procedures sometimes do not discover unexpected spin off effects. For instance, in a United Nations Children’s Fund (UNICEF) programme in Kosovo,

children without parents received within the frame of a psychosocial programme, school bags with the inscription ‘UNICEF’. So they were marked as children without fathers. Children did not want to carry the school bags and developed a negative attitude towards the programme. Close to this issue is the question, which is rarely asked: ‘*what have we destroyed with our programme, what natural support system and social processes are we hampering?*’ Certainly there are also many unplanned positive impacts of psychosocial programmes that escape from the net of data gathered through evaluation.

The most difficult is to evaluate the impact of the programme on the community as a whole – mobilisation of human resources, building social capital, synergic processes, etc. We know very little about what it means to people in affected areas that someone cares and is trying to help them. Or, more precisely, we know about it from our general insight into human experience and the psyche, but little from a scientific perspective. Does this then mean that the perspective matters so much, is it so important?

Research interests, ethic and social responsibility

There are many ethical questions stemming from the evaluation *procedure*. Applied evaluation tools can affect the responders psychologically. The best example are questionnaires containing the assessment of experienced traumatic events, traumatic reactions, post traumatic stress disorder (PTSD), etc. Such questions can trigger intense emotional and somatic reactions. Usually there is no one in the vicinity who could provide psychological assistance. Even if there were someone, have we the right to re-traumatise a person? On the other side, some questionnaires can trigger positive reactions, for instance questions on coping

strategies. They raise the responders' awareness about the importance of coping strategies and the insight into their own strength, and can activate some positive processes that contribute to healing.

Evaluation data often serves as research material. Sometimes, the main reason for evaluation is research. This is understandable as situations of disasters of any kind are a tragic natural occurrence that provides incredible possibilities for research. Areas affected by armed conflicts are a kind of 'promised land' for research. Research activities should be submitted to the rules of ethical codex. During armed conflicts the ethic control of professional instances are lacking, or insufficient. From my experience, I can say that questionable research procedures do happen. Furthermore, what often happens is that the researched population has no benefit from the research.

There is also the issue of social responsibility in interpreting and publishing results. Widely published research findings can influence public opinion. This can be harmful and enhance existing stereotypes. For instance, in the 1990s many researchers in Croatia and Bosnia found extremely high numbers of people with PTSD. They deduced from this, that the majority of children (or adults) from territories affected by war would have long lasting traumatic disorders, primarily PTSD (and consequently need specialised treatment, which was not available!). This view spread rapidly among professionals and the lay public, and was for evident reasons, widely accepted and sometimes welcomed by the politicians of attacked countries. The generalisation of the image of helpless victims who are not able to cope with the horrors and losses experienced without professional assistance, had some positive effects, such as making the need for psychological and psychosocial and therapeutic programmes evident, for

mobilising resources for such programmes, spreading the knowledge about psycho trauma and its acknowledgement. However, it had also an adverse impact, namely contributing to the learned helplessness of numerous people who readily accepted the role of a traumatised person, especially adolescents, unable to cope and passively waiting for assistance from others.

A questionnaire was prepared for Albanian school children in Kosovo asking them what they would do to their enemies. The evaluation was to be run in villages where mass killing of the population took place during the armed conflicts. The few Albanian children, responders forming the pilot group, answered that they would do to their enemy what the enemy did to Albanians in their village. As well as other objections to posing such questions in a questionnaire to children, publishing this 'evidence based' research finding would contribute to the image of wild, aggressive and vengeful Albanians, which was in those times was spread throughout Europe, particularly among those sympathising with the Serbs. Fortunately the research was never realised.

The views of mental health professionals can have social or even political impact, that can be highly relevant to the image of refugees in asylum countries, and can shape the attitudes of the host country toward refugees. Widely, publicly disseminated data about the behavioural problems caused by traumatised adolescent refugees, can raise fear of refugees and nurture xenophobic attitudes in the asylum country.

Evaluations cost too much

Evaluation requires valuable resources that could be used within the project itself. *How much does it cost?* is a question frequently posed (but often not loudly) by field workers. External evaluation costs quite a lot of money.

The better it is designed (more extensive, more in depth, combining different methods, elucidating the effects from different perspectives), the more it will cost. There are travel costs, fees to scholars, etc. Valuating the evaluation is a tricky issue. If the evaluation is positive and the programme receives money, there is the *'we knew that the programme is good even without evaluation'*. In the reverse case, field workers could consider the evaluation unjust. The problem is that a satisfactory dialogue between evaluators and field workers about the evaluation findings rarely happens, and the field workers are not acquainted with some of the more remote benefits of evaluations.

The most important questions concerning the implementation and the impact on the assisted population are the most difficult to be evaluated in a valid quantitative way. Such an evaluation could cost as much as the intervention programme, which was the case in my working history. It happened during wartime with a psychosocial programme run in Mostar (Bosnia Herzegovina). Decision makers becoming aware of the costs of a thorough evaluation, changed their mind and agreed to a less *'scientifically reliable'* evaluation.

Fieldworkers sometimes wonder why so much money is spent on consultants from fashionable consultancy firms who, for a fee of several hundreds of dollars per day, travel for a week or so to a post disaster area, rearrange all of the project activities and end up writing a report that will gather dust in the archives.

The power relationship between evaluators and field workers is asymmetric

We can repeat over and over again that evaluation should be a partnership and an egalitarian relationship between field workers and evaluators, but the fact is that

the future of the programme (further funding or support) depends on the opinion of the evaluators. Evaluators are appointed by donors to provide an opinion about the value of the programme, upon which the decision concerning the future of the programme or replication of similar programmes, depend. So the distribution of power is not at all equal. Field workers have, for a variety of reasons, a huge interest in the continuation or replication of their programme. The most common reason is that they strongly identify with their programme, and with the assisted population, and they want a programme that they consider helpful to people to be continued or spread. Of course, there are also less altruistic reasons. The income of field workers, or the survival of a non governmental organisation, can depend upon the continuity of funding of the programme. Therefore field workers are always a bit worried about the arrival of evaluators, as are teachers with the arrival of inspectors.

Evaluators can impose methods that seem inappropriate or even harmful to field workers. It can happen that, for instance, evaluators do not take into account the acceptability or feasibility of some evaluation procedure. Instead of explaining, I will give an example. An evaluator is proposing (and in my practice some were even requiring) teachers included in the psychosocial programme to write reports on children with special needs and psychosocial difficulties on a biweekly basis. Those reports would first of all serve as documentation for evaluation of effectiveness of outreach activities of local mental health workers visiting schools. The programme was run in remote villages of Kosovo in which it would not be acceptable to provide a written report about the behaviour of the child to other people, especially to strangers. The second issue was that teachers entering in the programme, in which

they took over the role of psychosocial helpers on voluntary basis without any fee, were earning about 150 Euros per month. In their free time they used to do manual work, such as building, in order to earn additional money to feed their large families. Writing a report biweekly and exposing oneself to the dissatisfaction of parents should not be requested from them for the sake of the evaluation of the programme. Many of them would simply drop out from the programme.

The process of evaluation can have negative effects on field workers

Programme staff can feel quite uncomfortable about how external evaluations are carried out. According to my own experience, some common concerns and negative feelings experienced by field workers are: evaluators do not always show respect for the efforts invested in the programme; the evaluation can be harmful, demotivate them and damage the programme. Field workers, both foreigners and especially local field workers, sometimes do not feel understood and have the impression that evaluators are depreciating their work or humiliating them in the process. Such feelings can have a damaging impact on people who work at the margin of their forces, who are not at all or very poorly financially rewarded, or work under dangerous circumstances. This is particularly relevant for local workers without whom the programme cannot be realised.

Evaluators sometimes have behaviour and attitudes that could be perceived as arrogance by field workers or beneficiaries. Some examples: 1) Not sufficiently exploring contextual or cultural feasibility and limitations of the required evaluation. 2) Having insufficient basic information and understanding of what was going on in the area in which the programme is run. 3) Using terminology

that cannot be understood by responders. For instance, a Bosnian evaluator of a programme run in Bosnia was asking elderly local people whether the programme respected the 'gender issue'. He used the term 'gender', which no one understood and it took a while before someone dared to tell the evaluator that people had no idea about what she was asking. On top of that, the gender issue was not the focus of the research. It had simply become fashionable and a concern of psychosocial programmes in the time period after the research had started.

A particularly common feeling among evaluated field workers is that the evaluators understood neither the extremely difficult circumstances, suffering, nor the misery of the population, nor the situation of psychosocial helpers who strongly identified with the assisted population. This feeling is similar to the feeling of traumatised people who consider that others who have not had the same experience cannot understand their state of mind. For instance, I had such a feeling during the first two days of the visit of my two colleagues from the Netherlands to North Ossetia-Abkhaz Republic. Both were deeply sensitive and compassionate women who had prepared for their visit to Caucasus by even learning some Russian. They came for an evaluation of the psychosocial programme mainly aimed at alleviating the consequences of the Beslan tragedy. We took them directly from the airport of Vladikavkaz to the cemetery of Beslan, where all victims of the terrorist act are buried under same reddish marble monuments, with toys and bottles of water on the tombs of the children. When walking together, I suddenly felt that somehow it was not appropriate may even be unfair, to bring them here, to the sacred place of suffering to which I felt entitled to be a part through my involvement in the region. Later I learned that my colleague

Katja, working with me in the region, had the same feelings. Besides, I wondered how Ossetian friends felt about foreign visitors coming to visit the graveyard and the destroyed school not only for compassion, but were also led by interest and asking questions that had already been answered hundreds of times before by local people. It took me two days to get rid of the feeling: *'they cannot understand'* and to start communicating in a normal way with our guests. This is not a place to analyse in depth such feelings, but field workers and guest evaluators should be aware of their existence and work on them.

Conclusions

Some of the examples described in this paper could be qualified as *'horror stories'*. Certainly beside those horror stories, there are many more cases of productive cooperation in the evaluation process. Those cases are not described in the paper.

I do not think that evaluation data should serve research. Data about people affected by armed conflicts and other disasters, can contribute to the development of new paradigms, and to the quality and cost effectiveness of provided psychosocial assistance. However, evaluation data should only be used under the condition that the researchers do not exploit people affected by war, maintain the highest ethical standards, and are aware of the social responsibility of researchers in a concrete political and social context.

Although quantitative measures are mostly desired by and appreciated by donors – and

they are indeed of crucial importance for the value of the programme – there are many other issues. These include as process dynamics proper to the programme and social processes supported or triggered by the programme, which need qualitative methods of assessment. The dissatisfaction with quantitative measures in social sciences resulted in the development and spread of participatory research and action research. How familiar are evaluators of psychosocial programmes in areas affected by armed conflicts with those methods?

The times when evaluations were perceived as reflecting lack of confidence in the value of a programme and its implementors, are long behind us. However, psychological obstacles for the evaluation of psychosocial programmes, especially when programmes are run in difficult circumstances and based on local human resources and good will of people, have to be taken into account. Evaluators need more than technical expertise. Common sense, integrity and honesty are no less needed.

Reference

- Adjukovic, D. (2008). Development of evaluation indicators in psychosocial projects: Balance between a creative challenge and research rigor. *Intervention* 6(1), 29-38.

Dr Anica Mikuš Kos, pediatrician and child psychiatrist, has worked in: Bosnia, and Herzegovina, Kosovo, Macedonia, Serbia, Ingushetia, Chechnia, North Ossetia-Allania, Pakistan, and Iraq. email: kos.a@siol.net.