ELSEVIER

Evaluation and Program Planning 28 (2005) 279–289

EVALUATION and PROGRAM PLANNING

www.elsevier.com/locate/evalprogplan

Responsive evaluation: Its meaning and special contribution to health promotion

Tineke A. Abma*

University of Maastricht, Healthcare Ethics and Philosophy, PO box 616, 6200 MD Maastricht, The Netherlands

Abstract

Responsive evaluation offers a vision and rationale for evaluation. In this vision, evaluation is reframed from the assessment of program interventions on the basis of policymakers' goals to an engagement with all stakeholders about the effectiveness of their practice. This approach is especially appropriate for the field of health promotion given emerging ideas and the congruency between the underlying values of responsive evaluation and health promotion. This article presents the theory and methodology of responsive evaluation and discusses several controversial issues among them the nature of evidence and the political question: *who* should determine what counts as evidence? The value and meaning of responsive evaluation are illustrated by a case example. It concerns the evaluation of an injury prevention program for students in two performing art schools.

Keywords: Responsive evaluation; Qualitative evidence; Politics; Health promotion; Injury prevention in performing arts

1. Introduction

'Evidence-based policy' and 'evidence-based medicine' have become catchwords in many countries around the world and are spreading fast. The evidence-based movement draws on experimental methods-advocating the wider application of random controlled trials (RCTs) and methodological rigour. There is, however, an increasing uneasiness among health researchers about the fact that RCTs and meta-analyses have become the norm in the field of health promotion (Koelen, Vaandrager, & Colomer, 2001; Mc Queen, 2000; Nutbeam, 1997, 1998; Tones, 1997; Morse, Swanson and Kuzel et al., 2001). Qualitative methods and more participatory forms of research are considered as an appropriate alternative and have long been utilized in health promotion/prevention circles in the US (Kahan & Goodstadt, 2001; Springett, 2001). In the Netherlands participatory approaches are less common in this field. Questions have been raised about the quality and usefulness of evidence collected through these methods.

In this article, I propose responsive evaluation as an orientation to evaluation that generates qualitative evidence about the effectiveness of programs. Evaluation criteria to asses the program's effectiveness are not only derived from the goals and intentions of policymakers, but include a wide range of *issues* of as many *stakeholders* as possible, including policymakers, managers, practitioners, community and target-groups (Guba & Lincoln, 1989; Stake, 1975; Stake & Abma, 2005). Responsive evaluation is a disciplined form of inquiry that results in qualitative-evidence. I will argue that this kind of evidence is important in the context of health promotion, because it enhances the understanding of human behaviour, it promotes holistic-thinking, offers contextual information and brings in the perspective of the community or target group. I start off presenting the theory of responsive evaluation. As an example of what responsive evaluation looks like in the context of health promotion, the second section describes a responsive-evaluation of an injury program at two performing art schools.

2. The theory of responsive evaluation

Robert Stake (1975) coined the term responsive evaluation. Egon Guba and Yvonna Lincoln (1989) rely on Stake's responsive evaluation and distinguish four generations in the historical development of evaluation: measurement,

* Corresponding author. Tel.: +31 43 3881132.
E-mail address: t.abma@zw.unimaas.nl.

0149-7189/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.evalprogplan.2005.04.003

description, judgement and negotiation. 'Measurement' includes the collection of quantitative data. 'Description' refers to the identification of the features of a program or policy. 'Judgement' is the assessment of the quality of a program based on a comparison between standards and actual effects. The authors identify several shortcomings related to the first-three generations. First of all there is a 'management-bias': goals and intentions of policymakers are taken over as standards for judgement. Secondly, evaluation findings are hardly used in decision-making. Thirdly, there is no dialogue *with* and *between* stakeholders, while their interests are at stake (value component). In addition no use is made of their experiences and expertise (knowledge component).

On these grounds the authors propose an alternative. The term '*negotiation*' characterises the essence of responsive evaluation.[1] In the following Section I present the core concepts of responsive evaluation followed by several validation strategies.

2.1. Core concepts

Criteria in a responsive evaluation are derived from the *issues* of various stakeholders. The issues gradually emerge in conversation with stakeholders and should be related to underlying value-systems in order to facilitate the negotiations and mutual understanding. Stakeholders are groups of people whose interests are at stake. In a responsive evaluation stakeholders should actively participate in the evaluation process; they are involved in the formulation of questions, the selection of participants and the interpretation of findings (Greene, 1997). Stakeholders become active and equal partners in the evaluation. Deliberate attention should be paid to the identification of 'victims' or 'silenced voices' (Lincoln, 1993), because they are often hard to find, for example, because they want to remain anonymous or because they fear sanctions.

Methodologically, plurality implies that the 'design' gradually emerges in conversation with the stakeholders. Metaphorically one may compare the designing process in a responsive evaluation with improvisational dance (Janesick, 2000). Whereas the minuet prescribes the definite steps, definite turns and foot and arm movements, improvisation is spontaneous and reflexive of the social condition. The evaluator charts the progress and examines the route of the study as it proceeds by keeping track of his or her role in the research process.

Besides the identification of issues the evaluator should create conditions for the interaction between stakeholders. This is a deliberative process. Deliberation refers to the interaction and dialogue between participants. They do not just accept each other's beliefs and persuasions, but will explore these. Listening, probing and dialogue characterise this process, rather than confronting, attacking and defending. Central features of dialogue are openness, respect, inclusion and engagement (Abma et al., 2001; Greene, 2001). Conditions for dialogue are the willingness of stakeholders to participate, to share power and to change in the process (Abma et al., 2001). Dialogue may lead to consensus. Absence of consensus is however not problematic; on the contrary, differences stimulate a learning process (Widdershoven, 2001).

In the exploration of issues the evaluator will concentrate on controversies, and end the evaluation with an agenda that will cover these controversies. The evaluator will not formulate conclusions or recommendations, because this prevents the input and interpretation by stakeholders. The evaluation report is a vehicle for dialogue (Abma, 1998). This 'working document' will portray the existing diversity and polyvocality.

In a responsive evaluation evaluator roles include the one of interpretator, educator, facilitator and Socratic guide. The role of interpretator indicates that the evaluator has to endow meanings to issues. The role of educator refers to the creation of understanding by explicating various experiences to involved groups. Facilitator refers to the organisation of the dialogue and the creation of required conditions. In the role of Socratic guide the evaluator will probe into taken for granted ideas, final truths and certainties, and bring in new meanings and perspectives (Schwandt, 2001a,b).

2.2. Validation strategies

A responsive approach will start with the collection of issues of stakeholders. It does so by using quantitative and qualitative methods (Stake & Abma, 2005). The combination of various methods leads to better, more comprehensive and more insightful understandings (Greene, Kreider, & Mayar, 2005). Qualitative methods are appropriate to gain an insight in the experiences and complexity of our social world. In order to guarantee the quality of knowledge generated, Guba and Lincoln (1989) have introduced the following trustworthiness and authenticity criteria and validation strategies:

The credibility of interpretations in the eyes of stakeholders as a validation strategy requires that respondents receive interpretations of (group) interviews with the question if they recognise the analysis (so called 'member check') (Meadows & Morse, 2001). Triangulation of sources and the use of mixed methods will help to include different perspectives and values and to prevent biases

[1] Stake's version of responsive evaluation focuses on re-directing data gathering and interpretative efforts around emerging issues of importance to program practitioners and other stakeholders in the evaluation setting. Guba and Lincoln (1989) rely on and explicitly refer to Stake's work, but they shift the focus to the 'negotiation' among diverse stakeholders towards a common consensus and acceptance. While Stake widens the evaluation scope to include a broad range of stakeholder issues, Guba and Lincoln promote a participatory and transformational process *with* stakeholders as co-owners of the evaluation. We follow Guba and Lincoln's 'negotiation' approach to responsive evaluation, but instead of aiming at consensus we opt for an approach aiming at reciprocal understanding and acceptance.

(Greene et al., 2005). Keeping a reflexive logbook or journal is a good way to keep track of the process and the evaluator's role in it. The fact that interpretations are always based upon and motivated by the evaluator's own questions and experiences is not problematic in itself. From a hermeneutic perspective, interpretation requires pre-understanding, which takes the form of prejudice (Widdershoven, 2001). This prejudice should, however, be open to correction.

Responsive evaluation results in context-bound knowledge. This local knowledge can be generalised from the studied context to the context of readers of the evaluation-report if it contains 'thick descriptions' (Geertz, 1973). Thick descriptions' not only reveal factual details, but also include meanings of experiences and events. Whether or not the results can be transferred to other situations is in large part to be decided by the reader.

The quality of the process is partly dependent on the created power-balance: all participants should be able to have 'a say.' Authenticity refers to the enhancement of personal and mutual understanding, changes in perspectives, and increased opportunities to act. In responsive evaluation one especially has to be aware of power relations (Koch, 2000). One should try to find means to give voice to people and groups that are less powerful. Giving voice means creating a safe space for people to express their issues and concerns. It also includes the representation of different voices in the evaluation reports. One way to give voice to people is to have in-depth interviews with them. Another is to organise homogeneous groups, in which people with the same position in the system can exchange experiences. These then can be introduced as issues in other stakeholder groups. By presenting such issues through stories, a climate of open discussion and dialogue may be fostered. Active engagement of as many stakeholders as possible and deliberation minimises the chance of bias and domination of one party. Afterwards, it needs to be checked whether the dialogical process was really open. A careful reading of the transcript can do this.

3. The case example

> 'It's all the same, basketball, tennis or ballet.' (Balanchine, ballet dancer)

Traditionally music and dance are considered to be performance arts. Yet, there are those who have compared dance and music with top-sport, like Balanchine. Hans van Maanen, a famous Dutch choreographer, even devoted a ballet to this theme (Schaik, 1997). While the demands placed on the physical and mental condition are in many ways comparable with what is expected from athletes, almost no attention is paid to the health condition of dancers and musicians. There is, for example, no team of (para-) medical experts to assist dancers and musicians. Warming-up and cooling down before a dance performance is not common, according to Rachel Beaujean an ex-ballet dancer of the Dutch National Ballet (Korteweg, 1997). Due to the lack of resources and the dependency of dancers on the choreographer—he is the one who chooses the cast—dancers feel obliged to go on, evening after evening, even if they suffer from injuries. Pain is accepted as an inevitable part of dancing. Musicians also experience the pressure to perform despite health problems. The Dutch violist Isabella van Keulen (1997) tells she had to be hospitalized with severe health problems before she realized that it is important to '*take better care of myself, to be less tough on myself*.'

The lack of attention paid to the prevention of injuries and health related problems in the professional practice of dancers and musicians does not differ from what is going on in the schools. Until recently no systematic attention was paid to the health condition of dance or music students. Lately and gradually this is changing (Jowitt et al., 2001). In 1990 the Dance Academy of the Higher School for the Arts in Amsterdam formed a special department within the school for health problems. This department developed an injury prevention program. The Conservatoire in Amsterdam also started a discussion on the relevance of injury prevention programs for their students. In both schools it was acknowledged that intensive training sessions and performances put great demands on the physical and mental health of the students.

3.1. Program description: injury prevention

The coordinator of the department for health promotion at the Dance Academy has developed an injury prevention program. The program consists of protective and preventive activities, curative care, regular consulting hours and a referral system (see for an overview of the activities, Box 1). The aim of the program is that students should become more aware of their body and their limits, and that they learn to use their body in a more appropriate way. Besides education, training and information, the program also has a function to detect injuries and health related problems.

Box 1. Overview of prevention program activities at the Dance Academy

- *Physical examination and advice*: During auditions for the school every student is examined physically by a physiotherapist and sports physician. Students are informed about physical (im) possibilities. Furthermore, they receive an advice how to deal with these impossibilities during the school period.
- *Regular consulting hours*: Three times a week a physiotherapist holds consulting hours. Once a week

an orthopaedic surgeon is available for students in the school. The coordinator of the program has a regular consulting hour for questions regarding postures and movements and more personal problems of the students. If necessary these experts can consult or refer students to a stable network of other experts.
- *Lessons and group conversations*: Lessons in anatomy and injury prevention are part of the regular curriculum. Subjects that are part of these lessons include healthy food habits, dealing with stress and the importance of warming-up and cooling-down. In addition lessons have been started to assist students with injuries to start up again with their training program. For those who are interested there is a conversation group about unhealthy food habits.
- *Teaching teachers*: The physiotherapist organizes lessons for teachers on the prevention of injuries and other health related problems.

After having worked with a program for several years the coordinator of the program and the Board of Directors wanted to gain more insight in the value of the injury prevention program for students and teachers. There were questions regarding the effectiveness of the program and information was needed to further optimize the program. The coordinator noticed for example, that despite the program the incidence of injuries remained high and that prevention, in the regular curriculum, was still a largely ignored dimension.

While the Dance Academy already had a full-fletched program the Conservatoire in Amsterdam had just started a discussion about health problems and injuries. Injury prevention was not part of the curriculum and the rate of injuries and related health problems were high at the school. The Conservatoire psychologist, for example, had seen almost 60% of the school population. The consulting students often had psychosomatic problems. Despite the severity of the situation at the Conservatoire, the subject was surrounded by taboos. The Director of the Conservatoire wanted to open up a public debate on the issue.

3.2. Responsive evaluation of injury prevention

Several years ago I was approached to evaluate the injury prevention program at the Dance Academy and to help to open up a public discourse on health at the Conservatoire in Amsterdam. The project was commissioned by the Board of Directors of both schools. Those who commissioned the project wanted to improve the quality of the injury prevention practice at both schools and information how to modify this practice. The aim of the evaluation was to motivate students, teachers and medical experts on injury prevention to reflect and think about ways to improve the quality of their practice.

The evaluation was carried out by a team of three evaluators, among them two psychologists. One of the team members was doing the study as part of her master thesis. The project lasted a year (April 1997–April 1998). The salary of the junior evaluator and her travel costs were financed by The Dutch Health Care Foundation for Students. The senior evaluators did their research work for free as community service. The junior evaluator, being a student and musician herself, could readily access the students and could easily identify with them. The senior evaluators, being university teachers, were in the position to identify with the teachers at the schools. The different social positions of the team members proved to be helpful to understand the issues of the different stakeholder groups.

A project group was formed to critically monitor the evaluation. The project group was composed of an executive manager of the Conservatoire, the coordinator of the injury prevention program of the Dance Academy and a staff member from the School for Higher Education of the Arts. In advance of the evaluation we talked with the project group about the design and who should take part in it. Other methodological issues (the relevance and amount of in-depth interviews, the recruitment of participants), ethical considerations (how to protect the privacy, anonymity and confidentiality of respondents) and financial aspects were also discussed in advance. Later on we renegotiated our course with this group, and decided to place more emphasis, for example, on the psycho-dynamics that prevented students and teachers from paying more attention to health issues. In the last phase the project group actively participated in discussing how results should be disseminated.

As evaluators we identified three groups of stakeholders: students, teachers and medical experts. This was not taken-for-granted by the project group. Initially the project group only wanted to use medical experts as informants. We convinced them that it was important to include other groups as well. The evaluation would affect the interests of students and teachers, and their participation might enlarge the scope of the knowledge generated and pave the way for action on the basis of information generated if they recognised themselves as the owners (Greene, 1988).

Over the course of a year, the junior evaluator worked for three to four days a week at the schools attending regular lessons, special body-awareness lessons and consulting hours as well as concerts and student performances. After some time students spontaneously approached her to talk about their experiences. This ‘prolonged engagement’ (Lincoln & Guba, 1985) enabled her to build up a relationship with the communities. In order to enhance our knowledge of the field we read several (auto) biographies and interviews with dancers and musicians. Once every two weeks we met as a research team to discuss methodological considerations and to reflect on how our particular position, research agenda, prejudices and main filters influenced the project. One of the prejudices we identified was that we

assumed that jazz musicians and modern dancers would be more willing to pay attention to health issues than classical musicians and dancers, because spontaneity and collaboration are more important in jazz music and modern dance. Although this was true in general, we had to correct this prejudice for those who had experienced an injury. One famous piano teacher told us how he had discovered due to an injury that mind and body are related, and that the quality of one's performance is related to one's health condition and wellbeing.

We started the evaluation with conversational interviews (Reissman, 1993) with two students (jazz and modern dance), two teachers and two (para-) medical specialists. One of the main principles that guided the selection of respondents was variety; we tried to gain a broad spectrum of meanings. So, we selected and recruited persons from each stakeholder group. The project group helped us with the recruitment, suggesting individuals who had suffered from injuries and other health problems. The interviews were not guided by our topics but by the issues brought to the fore by respondents. We started with broad opening questions, such as *'What happened when you were injured and had to stop (temporarily)?*' The interviews were all tape-recorded and transcribed. Our interpretations were presented to every respondent in order to give them the chance to comment on our findings ('member checks'). All the respondents, however, recognised themselves in our reconstruction. Unfortunately, one Conservatoire student decided that her interview could not be used as a source of information, because she was afraid that it would lead to sanctions. This was an indication of the lack of safety students experienced at the Conservatoire. We reported this to the project group without further details.

The personal interviews were used as an input for further dialogue via a series of storytelling workshops among groups of students and groups of teachers in both schools. We planned the workshops within the regular meetings and lessons. The groups that attended the workshops were small (six persons appeared to be the ideal size for a session of one or two hours). In the workshops, participants were invited to respond to story-fragments from the intermediary report. The presented stories were selected because they were like life and critical about the way self-care was approached in the schools. They were edited so that they could be read within a short time period. We decided to give students the stories of their teachers, and vice versa.

At this stage of the project we attempted to create safe and comfortable environments in which participants would show respect for each other's perspective. Two researchers facilitated the workshops. Permission was sought to tape the conversation. After a short reading pause of five minutes everyone was invited to introduce themselves and to relate their own experiences to the presented stories. After one or two hours we would end the workshop.

It is not possible to cover the richness of the dialogues during these workshops in this article, but the following vignette may provide some insight into the conversations.

> When we enter the room we find a noisy group of students, who are obviously having fun. We reshuffle the chairs in the form of a circle and then Margot, who is their mentor, introduces us to the group.
> Margot will not actively participate in the workshop. She is very interested in the student's experiences and when we talked about the safe environment necessary to get honest responses she persuaded us that her relationship with this group was very different from the one they have with teachers. Margot is not in the position to judge them. The students are not dependent on her for their degrees. Furthermore, she is their mentor and she wants to know how she can help them.
> One of us gives a short presentation, first in Dutch and later in English because there are several foreign students. We distribute the story fragments of student Johan and dance co-ordinator Margot—in English and Dutch-and after a reading pause, we invite the participants to respond in the form of a few statements.
> Four responses concerned the communication between students and teachers and we decide to talk about that topic first. Tamara says in her statement that teachers do not really communicate. When we invite her to explain what she means, she says that the teachers don't know the students personally and that older teachers in particular often act in an authoritarian way.
> One of her classmates responds as follows: 'You should not expect personal attention from a teacher. Things should not be too easy otherwise you may start feeling he or she will accept anything from you.'
> Tamara replies she understands that there needs to be some distance, but she is very disappointed that teachers do not take a personal interest in students.
> A conversation begins to develop about the way they relate to teachers and whether or not they will ask questions or talk about their problems. A third student says: 'It's very personal. Maybe she [points in the direction of the person sitting opposite her] finds it easy to ask about injuries or to be honest about pain, but it maybe very different for me. It all depends.'
> Everyone agrees personal preferences are important. At the same time, they recognise that most of the teachers with whom it is easy to talk are the ones who do not put themselves on a pedestal.
> The conversation then shifts to Johan's story. Rachel says she recognises his account and the feelings one has when one cannot practice because of an injury. She relates about a painful experience she has just had. At the end of her story Janice replies that she can

imagine that Rachel feels afraid because she may have to leave again. She shares her struggles with the pressure to perform: 'I do everything I'm supposed to do although I know it is not good for me (…) I do not want to stop, because I am afraid they will think I don't turn up because I don't want to study'.
Linda doesn't agree: 'You are now studying at a higher school for arts, you are responsible, not the teachers.' But the other students say they recognise Janice's fear that teachers and classmates will think negatively about them if they start to talk about injuries and difficulties.
Toni explains 'You're always thinking about the degree you will get in the end.'
When one of us asks if missing a few weeks will result in a poor degree, she replies: 'Yes that's what I feel.'

During the workshops, the evaluators acted as facilitators, paying deliberate attention to the development of trust and a respectful, open and comfortable climate. We acted as educators explaining to teachers how students thought and felt, and vice versa. As Socratic guides we raised thought-provoking questions, reflected on underlying value-systems and introduced voices that were not acknowledged. The data from the interviews, workshops and participant observation were brought together in an evaluation report. This report consisted of a series of stories of students, teachers and medical experts, emerging conversations within these groups and reflections by the evaluators.

3.3. Findings

The interviews and conversations with students and teachers revealed that an injury is not only a physical problem, but that it affects the person's whole well-being. Getting an injury is a dramatic episode in the lives of students and evokes intense feelings and emotions. Students often start at a very young age with a disciplined and monomaniac training program to work on their career. Their identity and future is directly connected with dance or music, and health problems are experienced as a threat to their identity and as long as possible denied. The injury evokes feelings of uncertainty, fear and powerlessness. The following vignette is illustrative:

Marilyn is 20 years old and has followed all kinds of trainings since the age of seven. During her first year at the Jazz and Show dance course at the Dance Academy she gets problems with her back. The cause appears to be an irritated bone in her spine. The weekly sessions at her physiotherapist offer her some relief, but in the first week of the second year something goes wrong. Marilyn has just started fanatically, but is untrained. A fall from the stairs does the rest. She cannot even walk normally and feels the pain in her left leg. The physiotherapist refers her to an orthopaedic surgeon. He tells her she has almost a hernia and that she should keep bed rest and use medicine for the pain and muscles. A manual therapist visits her at home on a daily basis. Marilyn describes this period as 'a hell'. She is extremely worried about her future: 'I think it will never work again, it won't work again.' After a few months, however, she is again dancing and she finishes with success her second year.

Marilyn's hell illustrates that the injury is affecting the student's whole life and identity. This is, however, not always recognized, according to students. Other studies also reveal that medical experts do not always pay attention to the psycho-social components of the injury (Lee, 1992). Students missed a safe space to discuss their problems with a neutral and independent person they can trust. One of the students put it as follows: '*Just someone with whom you can talk about your feelings*.'

Many injuries appeared to be directly related to the repetition of movements and the way teachers build up their lessons. Students found it, however, problematic to be critical and to raise this issue with their teachers, because they depended on them. Teachers evaluate their performance and decide whether or not they can go on. Students also found it difficult to consult their teachers when they experienced health problems. Teachers and students stated that students were themselves responsible for their own health. At both schools it was not done to consult teachers with questions concerning health. It was seen as a disturbance of the teacher. This interaction rule sustained the teachers' authority and maintained the distance between students and teachers. Several participants suggested, however, that the status quo did not sufficiently acknowledge that students need support from their teachers and classmates. It was suggested that teachers do not have to become therapists, but they might take care of their students' well being.

Another issue concerned the acceptance of physical limits (see also the above vignette of the conversation in one of the story workshops). Some participants emphasised that it is necessary to go beyond one's limits to develop professionally. Pain is the price that needs to be paid. This idea was quite dominant, and many students noticed they did not rest in time because they feared exclusion and social isolation. Others criticised these ideas. A 'time out' was said to be a meaningful period to recover physically and some participants emphasised the importance of prevention by paying serious attention to pain signals. It was suggested that teachers might assist students to heighten their body-awareness so that they would be better prepared to recognise their limits.

Still another issue we identified as evaluators was whether or not health problems should be actively prevented. Arguments for not developing an active injury

prevention strategy were grounded in the idea that the school is there to deliver top-talents and these select themselves out. The school has no responsibilities as far as prevention is concerned. Other participants stated that prevention and health promotion should be an integral part of the curriculum. Prevention should not only be a responsibility of experts, but of teachers in general. Furthermore prevention should not only focus on the injury, but on the student as an integral human being. It was also suggested that a good health condition is positively related with the quality of one's performance and it is therefore important to invest in self-care. Jan Wijn, a famous piano player and teacher at the Conservatoire told a compelling story:

> 'There was a moment in my life that I thought I didn't play very well anymore and that I didn't know what I should do with the music and that I longed for a sort of sabbatical to get off the stage and then my prayer was heard.'
> Mentally Jan Wijn longed for a time-out and a physical injury ultimately gave him a break. He considered it a sign that mind and body are intimately related to each other. He said he needed the time-out to 'get rid of bad ideas', ideas that stood in the way when playing the piano.
> His career went so fast that he continuously thought: 'Gosh, I am not yet ready for this, I am not good enough, everyone else is better than I am. I also had the idea that I cheated with the public, because I wasn't that good.' He experienced this fundamental uncertainty as a handicap. And then his body 'failed'. This had a physical component: the muscles in his right hand were weaker, because he is left-handed. Playing the piano does, however, put a great demand on the right hand. His mental conditions also played a role: 'The body responds, protests when there is something wrong in the head. That is always true.'
> His sabbatical has taken a long time. When his mind was again peaceful, he was able to play again, although he will never again be able to use one of his fingers. Jan overcame his handicap and grew personally and artistically.

The evaluation was responsive to both school communities and the various stakeholder groups within these communities. In order to further stimulate the dialogue on the above issues in the schools we presented a basic scenario that referred to a continuation of the actual educational practice within the schools and three alternative scenarios (Abma, De Jong and Van der Zouwe, 1998). The four scenarios were visualised in a two-dimensional scheme in which the horizontal dimension referred to the mission of the school (absolute top/individual development) and the vertical dimensions to the responsibility (individual/collective). Each scenario was further elaborated in terms of the practical consequences it would have for the school. After a discussion within the project group the report was distributed among the school communities and others who showed an active interest in it. It was also presented to the Board of Directors of both schools and formed the basis for a collaborative meeting between them. The Director of the Conservatoire expected that the report, in particular the story of the piano teacher, would play an important role in deciding the agenda to be followed by Board of Directors at the school. At that particular moment concrete actions were not yet formulated and the evaluators were not involved in further actions after the dissemination of the evaluation report.

3.4. Justification

In this particular example a responsive evaluation approach has proven to be beneficial. The approach illuminated the complexity of injury prevention and self-care in the participating schools. It showed that injury prevention is not only a matter of a lack of medical knowledge about risks, but related to human aspirations and fears, social interactions and exclusion and the organisational culture. Injury prevention only works if this complex context is taken into account. Responsive evaluation gives voice to persons otherwise not heard, in this case the students. Giving voice meant creating a safe space for students to talk about their experiences and concerns. This enabled them to see that their personal experiences were in fact issues that went beyond their personal situation. Giving voice also meant that the voices of the students were amplified via the presentation of their experiences to the teachers and school management. Furthermore, the evaluation stimulated a public discourse about issues that were taboo, created a space for reflection, fostered dynamics and motivated participants to think about ways to improve the quality of their teaching practice.

4. Lessons learned

An important lesson learned in the evaluation of the injury prevention programs was that the inclusion of the students stimulated teachers to re-think their teaching practice, values and points of view. Students also learned from teachers. After reading a story of a teacher, students began, for example, to realize that it was sometimes difficult for teachers to observe injuries or early signals of injuries. As such, the evaluation enhanced the mutual understanding among teachers and students. A second learning experience was that stories offer a way of reaching a deeper understanding of lived experiences and are an appropriate vehicle for reflective conversations because of their openness and ambiguity. Thirdly, engaging key decision-makers appeared to be an important strategy to gain acceptance for the findings.

Finally, we experienced that the conditions for a responsive evaluation were not optimal in the schools. The schools were characterised by asymmetrical relationships between teachers and students, while responsive evaluation requires a certain power balance to give all stakeholders a fair chance in the process. Health and self-care were sensitive topics in the schools and surrounded by many taboos. Furthermore the Conservatoire teachers were not very interested in joining the evaluation, while responsive evaluation requires the participation of as many stakeholders as possible. As evaluators we took these conditions into account by investing a lot of time in developing conditions of trust and safety. We, for example, respected the wish of one student not to publish her story because she feared sanctions. Furthermore, we decided not to bring students and teachers physically together in the evaluation process, but invited them to respond to each other via written stories. In order to increase stakeholder participation research activities were integrated in regular lessons and meetings, and not too time-consuming.

5. Special contribution to health promotion

Responsive evaluation offers a unique vision on evaluation given the link between an interpretive methodology and a democratic and emancipatory ideology. Responsive evaluation has been implemented in various policy sectors, including the field of health care. Case examples can be found in the evaluation of palliative care programs and units (Abma, 2000a, 2001; Groen, 2000), elderly care (Koch, 1994, 1996, 2000), nursing curricula (Koch, 2000), rehabilitation programs for psychiatric patients (Abma, 2000b; Wadsworth, 2001), and *Supported Employment* programs for mentally handicapped (Widdershoven & Sohl, 1999). The responsive approach to evaluation was also used as one of the first steps in planning positive youth development programs (Huebner & Betts, 1999) and the formulation of quality criteria in psychiatric care (Berghmans et al., 2001). Finally, Guba and Lincoln's work is applied in participatory forms of *medical technology assessment* (van der Wilt & Reuzel, 1998).

There are some special features and emerging ideas in the field of health promotion that make it especially amenable to responsive evaluation. Willy de Haes and Hans Saan (2002) relate these problems to the emerging approach in health promotion. According to this approach successful health promotion interventions require community participation. This means that the population or target-group is involved in the diagnostic process and in the preparation and further development of the interventions. In the emerging wisdom effective health promotion interventions also require co-ordination from various different angles and sectors (education, community work, socio-cultural work, urban planning, traffic and transport, social affairs, sport & recreation, etc.) on certain themes that relate to each party's own work. Responding to the current interests of the population is also considered a critical factor for a successful health promotion activity. This kind of approach, however, is difficult to incorporate in a study design drawn up in accordance with customary ideas. A RCT, or even a Community Intervention Trial, cannot be used because the purpose and content of the intervention are by no means fully established at the start of the activity, because the time schedule is undetermined and unpredictable and because only time will tell what outcome measurements are suitable. Stated differently, when health ideal and design meet the complicated and messy real world, outcome measures may no longer be appropriate.

The emerging 'design' of responsive evaluation allows it to be more responsive to these problems. Responsive evaluation uses a flexible methodology. The 'design' emerges on the basis of the 'issues' that appear to be important in daily practice (versus intentions). Responsive evaluation acknowledges that relevant outcome measurements cannot always be preordained and that practitioners who implement health promotion plans are confronted with the difficult task to adjust the design to the local context. It also acknowledges that during the implementation process more and more human, social and behavioural and cultural factors intervene. Responsive evaluation does not derive evaluation criteria from the idealised world of (policy) theory, but starts with the 'real time' actions and the lived experiences of practitioners. It acknowledges the inherent ambiguity of practice and the importance of adjusting abstract knowledge to local conditions and particular needs (Schwandt, 2001b). It honours the 'tacit knowledge' of practitioners and 'opens up' the narrative knowledge and wisdom developed in the process of implementation.

A second feature that is unique to health promotion concerns the fact that 'the medical literature on health promotion/disease prevention is long on ideas and systems and short on 'grounded theory,' i.e. theory which is generated by systematic observation of community-based clinical practice.' (Kuzel, in Lincoln, 1992). Anton Kuzel argues that there is a need for qualitative data, because health promotion interventions deal with human understandings, beliefs, fears, attitudes, prejudices, hopes, dreams and aspirations. The human, cultural and social side of health promotion is not well understood using quantitative methods, but can be assessed well and in all its complexity following a responsive approach. In the presented case responsive evaluation revealed that injury prevention is related to human aspirations (reaching the top) and fears (social exclusion), social interactions (fierce competition between students) and the organisational culture (not taking responsibility for the wellbeing of students). The evaluation showed that injury prevention only succeeds if all these intertwining factors and context are taken into account.

A final aspect of health promotion as a field which needs extensive evaluation work is in the arena of politics and ethics. Lynne Ray and Maria Mayan (2001) raise, for

example, the question: *who* determines what counts as evidence, the right indicators and appropriate standards in evaluation research of health programs? Their point of departure is that various stakeholders in a research study have various interests and diverging, sometimes conflicting, ideas about the most appropriate standards to assess their practice. In the medical sector the authors distinguish eight stakeholder parties, among them insurers, regulators, health care organisations, health professionals, the research community, the medical–industrial complex, the legal system and consumers. These parties have various agenda's, such as fiscal accountability, professional effectiveness, quality of care, safety and personal needs. Ray and Mayan argue that the general public has a small power base with regard to the production and the use of evidence. Responsive evaluation offers an approach to restore the power balance given the engagement, inclusion and active participation of various stakeholders, including the local community and target groups. Our evaluation of the injury prevention program, for example, gave voice to the students.

Responsive evaluation is not only responsive to the unique features and emerging ideas in the field of health promotion, it is also synergistic with health promotion.[2] Below three parallels between responsive evaluation and the emerging approach in health promotion are listed:

- The move from passive constructions of health to active and meaningful participation in the diagnostic process, in setting up and further development of the intervention,[1] from absence of disease to wellness, from sickness prevention to health promotion, reflects new understandings of the move from being a research object to a respondent and active participant in the evaluation process.
- The move from single causes to multiple, mutually interacting factors and the need for co-ordination from different angles and sectors reflects responsive evaluation's embrace of contextual interaction, mutually shaping forces, and webs of influence in human life and health.
- The move from a professional posture which focuses on disease, treatment and patient freed of accountability, to a professional posture where responsibility is taken for a whole patient and shared equally between practitioner and patient. This reflects the move to shared decisions, shared constructions and dialogue in responsive evaluation.

In short, the characteristics of responsive evaluation allow it to be more responsive to the features and developments in health promotion than any other form of evaluation. Responsive evaluation therefore has enormous potentials for addressing policy issues in health promotion. Moreover, the emerging ideas in health promotion are more congruent with responsive evaluation. There is a natural and meaningful 'fit' between these domains and a potential synergy.

6. Meeting new challenges

Implementation of responsive evaluation requires that evaluators are willing to give up some of their control over the process of the evaluation and develop a tolerance for ambiguity. Besides the usual analytical skills of a social scientist, a responsive evaluator requires additional interpersonal, communication and negotiation skills (Guba & Lincoln, 1981). These skills can be learned best by doing it, preferably as an 'apprentice' to an empathic and knowledgeable evaluator, in a climate of support and encouragement (Swenson, 1991). Furthermore as a responsive evaluator one must be willing to replace the expert role and adopt the role as interpretator, facilitator, educator and Socratic guide. Both the evaluator and policymaker should be willing to share their power with other stakeholders and to engage in more horizontal and joint collaborations with other stakeholders.

Although this article concentrates on the application of an evaluation model that promotes the use of mixed methods (Greene et al., 2005; Stake & Abma, 2005), a barrier in the implementation of responsive evaluation may relate to the use of qualitative methods. Evidence is often restricted to quantitative facts derived from large sample, randomised experimental designs, but does not capture the inherent complexity of health promotion (Mc Queen, 2000). Irena Madjar and Jo Ann Walton (2001) argue that a broad notion of evidence also includes qualitative evidence in the form of lived experiences, case histories and stories. This kind of evidence is important because it enhances the understanding of human behaviour; it promotes holistic thinking, offers contextual information and brings in the perspective of the community or target group. Qualitative data are more than just 'mere opinions,' because they are generated in a systematic way. Responsive evaluation is a form of disciplined inquiry and uses internal verification strategies such as the development of a research proposal and design in terms of the planned research activities and an indication of stakeholders, the working towards to point of 'saturation', and the striving for methodological cohesion. Validation strategies include the use of various methods ('triangulation') and 'member checks.' In short, responsive evaluation does produce evidence.

A third barrier for implementation is a concern about the practical application of such work. One of the strengths of responsive evaluation is that practitioners do

[2] The idea of 'fit' between the evaluation and intervention approach is not new. Others have argued that the approach an evaluator chooses to use ought to demonstrate parallels with the object and context of evaluation (Huebner & Betts, 1999; Lincoln, 1992; Schwandt, 1989).

not need to wait for research findings until the evaluation study is completed, but can in fact begin using findings along the process given the frequent communication and participation along the process. The recently developed notion of 'process use' is in interesting in this regard (Shulha & Cousins, 1997). 'Process use' refers to the acceptance of knowledge and the personal and organisational learning processes that occur during the evaluation process. Stakeholder participation, engagement and communication promote 'process use' (Greene, 1988). It will give participants more confidence in the quality of information and in the ability to use the information. Responsive evaluation does not only deliver evidence in time, but also evidence that is context-bound. It produces local knowledge that enables practitioners to use it in their context, in a specific case. As such, it acknowledges the fact that practitioners do not only require knowledge of scientific studies, but that they need information about the specific needs, life-style, preferences, problems, history and other particularities of the community or target group in order to make the right decisions.

The case presented here may suggest that responsive evaluation is only feasible within a relatively closed system, such as a school community. Good experiences have been noticed, however, about the possibility of responsive evaluation within the context of social renewal in the city of Rotterdam (Abma, 1997; Fortuin, 1993, 1994). In three disadvantaged neighbourhoods that differed with respect to socio-geographic space, the need for social renewal and leadership, the evaluator spent six months getting insight in the actors developing authority, resistance demonstrated and kinds of social action undertaken. Then stakeholders (citizens, civil servants, social welfare organisations, business people and politicians) were invited to list projects that envisaged social renewal. In each of these neighbourhoods the evaluator selected three of the nominated projects for further investigation. This example demonstrates that it is possible to conduct a responsive evaluation within an open environment like a city.

The concept of responsive evaluation has led to paradigm debates (Guba et al., 1990). Although it is important to discuss philosophical matters, such as the representation and legitimation of knowledge, a paradigm debate becomes unproductive if it is restricted to a discussion about methods. Discussions over methods keep us from issues that are more important, such as the idea of evaluation as a social practice and what it means to work for practitioners in the field of health promotion.

Acknowledgements

This is an edited text of a paper presented at the '*Evaluation in Health Promotion*' conference, Rotterdam, 17 February 2003. I would like to thank Willy de Heas, Hans Saan and David McQueen and the reviewers for their constructive feedback on an earlier version of this article.

References

Abma, T. A. (1997). Playing with/in plurality: Revitalizing realities and relationships in Rotterdam. *Evaluation*, *3*(1), 25–48.

Abma, T. A. (1998). Writing for dialogue, text in an evaluation context. *Evaluation*, *4*(4), 434–454.

Abma, T. A. (2000a). Responding to ambiguity, responding to change. The value of a responsive approach to evaluation. *Evaluation and Program Planning*, *23*(2), 461–470.

Abma, T. A. (2000b). Stakeholder conflict: A case study. *Program Planning and Evaluation*, *23*(2), 199–210.

Abma, T. A. (2001). Evaluating palliative care. Facilitating reflexive dialogues about an ambiguous concept. *Medicine, Health Care and Philosophy*, *4*(3), 261–276.

Abma, T. A., Greene, J., Karlsson, O., Ryan, K., Schwandt, T. S., & Widdershoven, G. (2001). Dialogue on dialogue. *Evaluation*, *7*(2), 164–180.

Berghmans, R, Elfahmi, D., Goldsteen, M., en Widdershoven, G. (2001). Kwaliteit van dwang en drang in de psychiatrie. Eindrapport. Utrecht/Maastricht: GGZ Nederland; Cluster Zorgwetenschappen, sectie Gezondheidsethiek en Wijsbegeerte, Universiteit Maastricht; Instituut voor Gezondheidsethiek, Universiteit Maastricht.

Fortuin, K. (1993). Gezonde stad stuit op adoptieproblemen. *Tijdschrift voor gezondheid en politiek*, *11*(4), 6–8.

Fortuin, K. (1994). Evaluatie-onderzoek in een paradoxale beleidscontext. In A. Francke, & R. Richardson (Eds.), *Evaluatie-onderzoek, kansen voor een kwaliatieve benadering* (pp. 155–179). Bussem: Coutinho, 155–179.

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York: Basic Books.

Greene, J. (2001). Dialogue in evaluation; a relational perspective. *Evaluation*, *7*(2), 181–203.

Greene, J. C. (1988). Stakeholder participation and utilization program evaluation. *Evaluation Review*, *12*(2), 91–116.

Greene, J. C. (1997). Participatory evaluation. In Linda Mabry, *Evaluation and the post-modern dilemma. Advances in program evaluation* (vol. 3) (pp. 171–189). Greenwich: JAI Press, 171–189.

Greene, J. C., Kreider, H., & Mayer, E. (2005). Combining qualitative and quantitative methods in social inquiry. In B. Somekh, & C. Lewin (Eds.), *Research methods in the social sciences* (pp. 274–281). London: Sage, 274–281.

Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation*. Beverly Hills: Sage.

Guba, E. G., & Lincoln, Y. S. (1989). *Fourth generation evaluation*. Beverly Hills: Sage.

Huebner, A. J., & Betts, C. S. (1999). Examining fourth generation evaluation: Application to positive youth development. *Evaluation*, *5*(3), 340–358.

Janesick, V. J. (2000). The choreography of qualitative research design: Minuets, improvisations and crystallisation. In N. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 379–400). Thousand Oaks: Sage, 379–400.

Jowitt, D., et al. (2001). *Not just any body, advancing health, well-being and excellence in dance and dancers*. Canada, Ontario: The Ginger Press.

Kahan, B., & Goodstadt, M. (2001). The interactive domain of best practices in health promotion: Developing and implementing a best practices approach to health promotion. *Health Promotion Practice*, *2*(1), 43–67.

Koch, T. (1994). Beyond Measurement: fourth generation evaluation in nursing. *Journal of Advanced Nursing*, *20*, 1148–1155.

Koch, T. (1996). Implementation of a hermeneutic inquiry in nursing: Philosophy, rigour and representation. *Journal of Advanced Nursing*, *24*, 174–184.

Koch, T. (2000). 'Having a say': Negotiation in fourth generation evaluation. *Journal of Advanced Nursing*, *31*(1), 117–125.

Koelen, M. A., Vaandrager, L., & Colomer, C. (2001). Health promotion research: Dilemmas and challenges. *Journal of Epidemiology and Community Health*, *55*, 257–262.

Korteweg, A. (1997). Interview met Rachel Beaujean. Halverwege de trap naar de hemel. *De Volkskrant*, *13 juni*.

Lee, S. A. (1992). Patterns of choice, bias and belief: report from a pilot study on dancers and health care. *Medical Problems of Performing Artists*, *June*, 52–57.

Lincoln, Y. S. (1992). Fourth generation evaluation, the paradigm revolution and health promotion. *Revue Cannadienne De Sante Publique, April*, *83*, 6–10.

Lincoln, Y. S. (1993). I and Thou: Method, voice, and roles in research with the silenced. In D. McLaughlin, & W. Tierney (Eds.), *Naming silenced lives*, 29–47.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills: Sage.

Madjar, I., & Walton, J. A. (2002). What is problematic about evidence. In J. M. Morse, J. M. Swanson, & A. J. Kuzel (Eds.), *The nature of evidence in qualitative research* (pp. 28–47). Thousand Oaks: Sage, 28–47.

Mc Queen, D. V. (2000). Perspectives on health promotion: Theory, evidence, practice and the emergence of complexity. *Health Promotion International*, *15*(2), 95–97.

Meadows, L. M., & Morse, J. (2001). Constructing evidence within a qualitative project. In J. M. Morse, J. M. Swanson, & A. J. Kuzel (Eds.), *The nature of qualitative evidence* (pp. 187–201). Thousand Oaks: Sage, 187–201.

Nutbeam, D. (1997). A conceptual model for health promotion, and the consequences for research. Report of the Expert Meeting Beyond RCT towards Evidence Based Public Health, February 13, 1997, Rotterdam, p. 37–46.

Nutbeam, D. (1998). Evaluating health promotion—progress, problems and solutions. *Health Promotion International*, *1391*, 27–44.

Ray, L. D., & Mayan, M. (2001). Who decides what counts as evidence?. In J. M. Morse, J. M. Swanson, & A. J. Kuzel (Eds.), *The nature of evidence in qualitative research* (pp. 50–73). Thousand Oaks: Sage, 50–73.

Schaik, E. (1997). *Hans van Maanen, Leven en Werk*. Amsterdam: Arena.

Schwandt, T. S. (1989). Solutions to the paradigm conflict, coping with uncertainty. *Journal of Contemporary Ethnography*, *17*(4), 379–407.

Schwandt, T. S. (2001a). A postscript on thinking about dialogue. *Evaluation*, *7*(2), 264–276.

Schwandt, T. S. (2001b). Responsiveness and everyday life. In J. C. Greene, & T. A. Abma, *Responsive evaluation*. *New directions for evaluation* (vol. 92), 73–88.

Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *Evaluation Practice*, *18*(3), 195–208.

Springett, J. (2001). Participatory approaches to evaluation in health promotion. In I. Rootman, M. Goodstadt, B. Hyndman, D. V. Mc Queen, L. Potvin, J. Spirngett, & E. Ziglio, *Evaluation in health promotion, principles and perspectives* (vol. 92) (pp. 83–106). WHO Regional Publications, European Series, 83–106.

Stake, R. E. (1975). To evaluate an arts program. In R. E. Stake (Ed.), *Evaluating the arts in education: A responsive approach, Colombus Ohio, Merrill*, 13–31.

Stake, R. E., & Abma, T. A. (2005). Responsive evaluation. In S. Mathison (Ed.), *Encyclopaedia of evaluation* (pp. 376–379). Thousand Oaks: Sage, 376–379.

Swenson, M. M. (1991). Using fourth generation evaluation in nursing. *Evaluation and the Health Professions*, *14*, 79–87.

Tones, K. (1997). Research for health promotion practice. A tale of three errors, illumination and the judicial principle, Report of the Expert Meeting Beyond RCT towards Evidence Based Public Health, February 13, 1997, Rotterdam, p. 46–60.

van der Wilt, G. J., & Reuzel, R. (1998). Assessment of health technologies, Which issues should be addressed? *Evaluation*, *4*(3), 351–358.

Wadsworth, Y. (2001). *Becoming responsive- and some consequences for evaluation as dialogue across distance New Directions for Evaluation*, vol. 92. San Francisco: Jossey Bass pp. 45–58.

Widdershoven, G. A. M. (2001). Dialogue in evaluation: A hermeneutic perspective. *Evaluation*, *7*(2), 253–263.

Widdershoven, G. A. M., & Sohl, C. (1999). Interpretation, action and communication: Four stories about a supported employment program. In T. A. Abma, *Telling tales. On narrative and evaluation. Advances in Program Evaluation* (vol. 6) (pp. 109–130). Connecticut: JAI Press, 109–130.