



# **Considerations in the Design of a Mixed-Method Cluster Evaluation of a Community Programme for 'At-Risk' Young People**

JAYNE C. LUCKE

*University of Manchester Medical School, UK*

MARIA DONALD

*University of Queensland, Australia*

JO DOWER

*University of Queensland, Australia*

BEVERLEY RAPHAEL

*New South Wales Health Department, Australia*

This article discusses the design of a comprehensive evaluation of a community development programme for young people 'at-risk' of self-harming behaviour. It outlines considerations in the design of the evaluation and focuses on the complexities and difficulties associated with the evaluation of a community development programme. The challenge was to fulfil the needs of the funding body for a broad, outcome-focused evaluation while remaining close enough to the programme to accurately represent its activities and potential effects at a community level. Specifically, the strengths and limitations of a mixed-method evaluation plan are discussed with recommendations for future evaluation practice.

**KEYWORDS:** community; mixed-method evaluation; self-harming behaviour; young people

## **Introduction**

Evaluators have struggled with the many difficulties inherent in the evaluation of community development programmes (Dixon and Sindall, 1994; Duignan and Casswell, 1989; Gruenewald, 1997; Nee and Mojica, 1999). Demonstrating that

these programmes have brought about change is often challenging. Funding organizations commonly require interventions to demonstrate their efficiency in terms of costs saved and outcomes achieved. This is often difficult for community development programmes because outcome-focused quantitative approaches to evaluation may not be sensitive enough to provide a true account of the effects of a programme on a community. Developments in more abstract outcomes such as 'community capacity' are widely used (Abatena, 1997; McLaughlin et al., 1997; Poole, 1997) but there is considerable debate about definitional issues (Goodman et al., 1998). The time required for community development programmes to reach an efficient level of functioning is often not considered by funding organizations. Therefore, evaluations may occur prematurely, resulting in an inappropriate assessment of programme effects.

This article discusses the design of a comprehensive cluster evaluation of a community development programme for young people 'at-risk' of self-harming behaviour. The multi-method evaluation design aimed to fulfil the needs of the funding body for a broad, outcome-focused evaluation while remaining close enough to the programme to accurately represent its activities and potential effects at a community level.

## **The Programme**

The programme was a suicide prevention programme implemented in four separate areas of Queensland, Australia and funded by the state government over three years, at 1.5 million Australian dollars per year. Suicide among young people has become the focus of growing concern in Australia. During the past 30 years, rates of suicide among young males have trebled (Harrison et al., 1997). The increase in rates of completed suicide has been less dramatic among young females but attention has focused on high rates of suicide attempt and depression (Patton et al., 1997).

The programme was intended to use a primary health care approach to empower communities to become active in the prevention of self-harming and suicidal behaviour amongst young people aged 10–24 years. It was co-ordinated from central government offices, and multi-disciplinary programme teams were located in four different areas of the state, corresponding to regional health authority areas. The programme teams were intended to provide an area-wide service and undertake a range of activities, including education and training, support and referral. In particular, the programme teams were intended initially to establish Local Action Reference Groups (LARGs) and then to support their functioning within the regions. LARGs were to be local community coalitions made up of local people bringing about change for young people.

The LARGs were intended to consist of a core group of informed key people at a local level, such as teachers, general practitioners, police and other community members, who would be trained by the local programme team to act as a source of information and support for the community when dealing with young people. Voluntary community coalitions are becoming increasingly popular as a method of addressing difficult community health problems' (Bogensneider,

### *Evaluation 7(1)*

1996; Butterfoss et al., 1993). They have been applied to various issues affecting young people, such as substance abuse (Fawcett et al., 1997) and unplanned pregnancy (Vincent et al., 1987). The rationale behind this strategy is that coalitions can enhance access to existing community resources and increase longevity of community commitment to prevention by increasing the participation of a broad range of community members, thus building links between members and giving a sense of local ownership (Stevenson et al., 1996).

The programme was based upon a number of underlying principles of community development (Chavis and Pretty, 1999; Checkoway, 1998). Programme documentation articulated these principles by prescribing: (a) a prevention focus; (b) accessibility to young people; (c) a high emphasis on community participation in, and ownership of, processes; (d) the inclusion of community reporting mechanisms to inform and communicate with the local community; (e) the maintenance of strong health/mental health links; (f) the encouragement and facilitation of collaboration across all sectors including the community, community agencies and government organizations; and (g) a focus on operational outcomes with the provision of timely and effective on-the-ground service.

The programme was based on the recognition that health is determined by a broad range of social, environmental and biological factors and that effective service delivery requires a regionally specific response to local factors. Consequently, each of the four programmes developed and implemented a regionally specific model based on community consultations. The four specific areas selected to participate in the programme corresponded to existing health authority regions at that time. They were selected to ensure inclusion of areas which were rural (Area 1), metropolitan (Area 4), rural/metropolitan mixed (Area 3) and coastal/provincial mixed (Area 2). This approach was adopted to provide the opportunity for the development of a number of different models which could then be transferred to other similar communities.

A summary of the four regional programmes follows. Area 1 was a rural/remote region covering over 322,517km<sup>2</sup> with an estimated population of 29,872. Around 21 percent of the population were aged between 10 and 24 years and the region was characterized by an above average unemployment rate among young people, a high proportion of young people identifying as being Aboriginal or Torres Strait Islander, and difficulty in attracting and retaining appropriate services. The programme in Area 1 commenced in March 1995 and the team included a co-ordinator, four young people's support workers, two young people's mental health workers, a resource worker and funding for administrative support. The focus of the Area 1 programme team was on community development activities, although approximately 20 percent of the young people's support workers' time was spent on direct service activities.

Area 2 was a coastal/provincial region of 52,150km<sup>2</sup>. An estimated 22 percent of the total population of 193,991 were young people aged between 10 and 24. The area had experienced rapid population growth with demographics changing from predominantly retired/aged pensioners to be younger and family based. In addition, the region had a high unemployment rate and, along with Area 1, was one of the most socio-economically disadvantaged regions in the state. The Area

2 programme commenced in May 1995 and the programme team consisted of a co-ordinator, three project officers and an administrative officer. This team focused primarily on the development of intersectoral collaboration as well as the identification and filling of gaps in services and training.

Area 3 was a smaller region on the outskirts of a large city. It included both metropolitan and rural areas with a total population of approximately 150,000. Young people made up approximately 21 percent of the total population and the region was characterized by a high youth unemployment rate. The Area 3 programme also commenced in May 1995 with a programme team consisting of a co-ordinator, a counsellor, a support and education officer, an information and resource officer and an administrative officer. The focus of this team was primarily on education, training and resourcing within the community. The counselling position, however, contained a small clinical component.

Finally, Area 4, the smallest region, was a metropolitan area undergoing rapid population growth with an estimated population of 168,013, of whom almost 47 percent were under the age of 25. The area was characterized by a high degree of cultural diversity and unemployment for young people. The Area 4 programme commenced in January 1996 and the programme team consisted of a co-ordinator, a community development/support officer, a resource development/education officer, a migrant liaison officer, an indigenous peoples liaison officer and an administrative support officer. This team initially provided a counselling service; however, it was refocused during 1996 to undertake a training, education and resourcing function.

The evaluation of the programme was carried out by an independent evaluation team from a local university. The evaluation team were asked to provide a comprehensive, outcome-focused evaluation of the whole programme.

## **Programme Goals and Objectives**

Clear goals and objectives are vital for the successful implementation and evaluation of any intervention programme (Hawe et al., 1990). Although existing programme plans and documentation outlined broad goals, they were not specific enough to use for the development of a realistic evaluation plan. For example, the main goal stated in programme documentation focused on the reduction of suicide as the main outcome measure. However, although it was politically necessary that the programme be focused to some extent on suicide and self-harming behaviour, a three-year community development programme is unlikely to be able to demonstrate a significant impact on this outcome, particularly since suicide is a relatively rare event (Moscicki, 1995; Patton et al., 1997; Silverman and Maris, 1995). Therefore, before a comprehensive evaluation plan could be developed, extensive consultations occurred in an attempt to determine the precise nature of the programme, and clearly identify its objectives and strategies. Table 1 shows the final programme objectives and some examples of the general strategies employed by programme teams to meet them.

The objectives were developed over several meetings with each programme team and refined throughout the project. For example, an original objective

Table 1. Programme Objectives and Strategies

TARGET POPULATION	PROGRAMME OBJECTIVE	TYPES OF PROGRAMME STRATEGIES
SERVICES	To facilitate optimal service provision to meet mental health needs for young people.	Education and training
	To develop strategies to fill identified gaps in the continuum of prevention, intervention and postvention services.	Strategies appropriate to specific gaps, e.g. Education and training Provision of support and referral Application for funding
	To facilitate strong, effective and ongoing intersectoral links, communication and planning between all relevant services.	Facilitation of collaboration and communication between services
YOUNG PEOPLE	To increase young people's awareness of, and ability to access, prevention, intervention and postvention services	Provision of information, support and referral
GENERAL COMMUNITY	To increase the ability of the community to respond to young people at risk.	Develop Local Action Reference Groups

relating to surveillance and monitoring self-harm was abandoned when it became evident that such a task was beyond the resources of the programme. The programme teams employed diverse strategies in order to meet the objectives. However, although the focus of the teams varied according to the perceived needs of their areas and the expertise of the staff, there was considerable overlap in the actual day-to-day activities of the teams over the life of the programme.

Area 1 focused heavily on direct contact with young people because of the severe lack of services in their remote location; Area 2 emphasized the need to link relevant organizations together and raise awareness of issues relating to young people; Area 3 focused on the provision of information and training for services who may deal with young people at risk of self-harm; Area 4 was delayed in starting and therefore the programme team did not participate in this consultation process, although consultations did take place with key personnel involved in programme planning. Despite the disparate emphases and strategies of each programme team, they were each concerned, to some extent, with targeting three distinct populations: (a) young people themselves; (b) services and organizations working with young people; and (c) the general community in the area. The evaluation plan focused on these areas of commonality between the diverse programme teams.

### **The Evaluation Agreement**

The evaluation team was required to provide a comprehensive evaluation of the programme at both state and area levels. Although the evaluation was to contain both qualitative and quantitative components, there was to be a heavy emphasis on quantitative data collection. Extensive baseline data were to be collected to inform the development and implementation of the programme. The evaluation agreement required an extensive report, to be submitted to government officials, which would be used in determining the level of ongoing support for the programme.

### **The Evaluation Plan**

The identified goals and objectives of the programme were used as the framework for the development of a comprehensive evaluation plan (final version shown in Table 2). The evaluation strategy attempted to address formative, process, impact and outcome evaluation considerations and focused on the three populations targeted by the programme: services, young people and the general community. This was achieved through five major components: three interview studies (focusing on key service providers, young people and programme team members) and two large surveys (with service providers and young people). Areas 1, 2 and 3 participated in all the evaluation components; however, Area 4 was not included in some components of the plan because of various difficulties. These included an early change of focus which delayed the start of the programme, and continuous rapid turnover of staff. Unfortunately, these problems led to the parent organization deciding on the early closure of the programme in this area.

Table 2. Programme Evaluation

TARGET POPULATION	EVALUATION STRATEGY			
	FORMATIVE	PROCESS	IMPACT	OUTCOME
SERVICES	Interviews with Service Providers (Times 1 and 2) Survey of Service Providers (Times 1 and 2)	Area Teams conducted their own process evaluation.	Survey of Service Providers (Time 1 and Time 2)	Option for future repetition of Survey of Service Providers
YOUNG PEOPLE	Interviews with Young People (Times 1 and 2)	Area Teams conducted their own process evaluation.	Interviews with Young People (Time 2)	Survey of Young People (baseline data) (Option for future repetition of Survey of Young People)
GENERAL COMMUNITY	Interviews with Area Team members (Times 1 and 2)	Area Teams conducted their own process evaluation.	Not Applicable	Not Applicable



## **Evaluation Components**

### ***Interviews with Key Service Providers***

Interviews with key service providers were conducted in order to provide early formative feedback to programme staff. Three out of the five programme objectives related to services. These interviews were also used to gain preliminary information to guide the design of evaluation for these objectives, specifically the Survey of Services.

**Method** A total of 70 service providers were interviewed about their service provision and their contact with the programme. The interviews took place twice, once as soon as the programme teams were established, and again one year later. The service providers included mental health services (e.g. psychiatrists, psychologists and social workers), schools and youth services (e.g. guidance officers and school support centres, youth shelters, youth drop-in-centres, youth employment and training services), medical services (e.g. hospital directors of nursing, general practitioners, medical centres, women's health services and the Ambulance Service), other community services (e.g. community health, community support groups and neighbourhood centres), Aboriginal and Torres Strait Islander services, places of worship (e.g. priests and church youth workers) and legal services (e.g. police and juvenile justice services). These service providers were identified by the programme teams as relevant to the programme and likely to be (or already) involved with programme activities. However, some services were contacted directly by the evaluation team. This generally occurred where the programme team had not yet developed contacts with a particular type of service. The rationale was to gain a good spread of service types. We were not concerned about bias caused by interviewing contacts of the programme. This was for two reasons: (a) in order to maintain a collaborative approach to evaluation; and (b) a forthcoming Survey of Services would be more likely to provide unbiased data for evaluation purposes.

The interview aimed to: (a) provide an outline of the function of the service; (b) establish the extent of the service's contact with young people, particularly young people at risk of suicide and self-harming behaviour; (c) ensure that the service was informed about the programme; and (d) identify opportunities for assistance from the programme.

**Use of Results** The findings of these interviews were not used in the impact evaluation of the programme. Rather, the results were discussed with the programme teams as formative evaluation. The results were also used to develop the major source of data for the impact evaluation: the Survey of Services.

Programme teams found the information from these interviews useful in developing better intersectoral links with service providers. The first round of interviews served to initiate contact with many service providers and raised awareness of the programme and its aims. This was also the case in the second round of interviews. It had originally been intended that the second interview would provide an opportunity to evaluate the service providers' perceptions of



### *Evaluation 7(1)*

the impact of the programme. However, due to the high rate of staff change, the second interview was often conducted with someone who was new to the organization or position and therefore not able to comment extensively on the impact of the programme.

The results of the interviews were useful in developing the questionnaire for the Survey of Services. The process of arranging and conducting the interviews highlighted difficulties such as high staff turnover and poor communication within and between services, which would have implications for the design of the survey. In particular, the survey would need to be sent to a specific position within the organization, rather than to the organization itself. Also, it would need to be recognized that the information received was from the point of view of this person only, and not necessarily reflective of the entire service, particularly in large organizations.

### ***Interviews with Young People***

Interviews were conducted with young people in Areas 1, 2 and 3 in response to requests from programme teams for more qualitative components in the evaluation plan. The interviews were designed to provide recommendations about programme functioning from the perspective of young people involved in the activities of the programme.

**Method** As the programme activities differed substantially between areas, the young people selected were diverse. However, in general, young people were either programme clients receiving therapy or support, or they were participants in community development activities. Each area team selected young people who had been in contact with the programme in their area. This was the best way to identify appropriate young people; however, the biased nature of the sample must be considered in interpreting the results.

Overall, 59 young people aged between 10 and 24 were interviewed. Most were female ( $N=37$ ) and still at school ( $N=47$ ). The interviews were conducted, either face-to-face or by telephone, by a member of the evaluation team of the same gender as the interviewee.

**Use of Results** This component was useful as most young people who were interviewed reported that they had felt they had benefited from their involvement with the programme. No significant problems were identified with young people's access to the programme or their involvement in the clinical/support components or the community development activities. They felt that the programme had helped them and they were willing to recommend the programme to their friends. This provided encouraging feedback to the programme teams, as well as useful evaluation data demonstrating the satisfaction of young people with the programme, despite the small, biased sample.

### ***Interviews with Programme Team Staff Members***

Interviews were conducted with 20 programme team staff members from Areas 1, 2 and 3 during June, 1996 (Time 1) and 21 members in June, 1997 (Time 2).

The aim of these interviews was to (a) provide recommendations about future programme implementation and functioning at Time 1 and Time 2; and (b) provide specific evaluation of the LARG concept.

**Use of Results** These interviews resulted in many useful recommendations for programme functioning. Programme team staff discussed strengths and weaknesses of the programme in reaching young people, services for young people and the general community. Specific recommendations for future programme implementation were made regarding the structure of the programme, the staffing of the teams, programme funding and administration, resources, research and evaluation. The programme team staff members provided many useful insights into programme functioning. They were enthusiastic about programme strengths and gave constructive suggestions for addressing weaknesses.

In particular, the LARG concept was found to be problematic. Programme team members reported limited success in establishing LARGs and it was generally felt that they were an interesting concept but not practically viable. Practical barriers to forming LARGs included confusion about the role of a LARG, and difficulty establishing and maintaining a cohesive, enthusiastic group, particularly with respect to the amount of time and resources required by programme teams to do this.

### ***Process Evaluation (Monitoring)***

It was intended that an extensive process evaluation be conducted. The rationale for this was to: (a) document the programme; (b) provide formative evaluation to improve and develop the programme; and (c) monitor the programme. Monitoring is particularly important where young people at risk of suicide and self-harm are targeted by untested programmes (Mrazek and Haggerty, 1994). Regular collation and analysis by the evaluation team were intended to provide regular progress reports as successfully demonstrated elsewhere (Francisco et al., 1993; Paine-Andrews et al., 1997).

However, when this was negotiated with programme teams, it emerged that extensive data were already required from the teams as part of their operational planning procedures within the organizations in which they worked. There was strong resistance to the suggestion that any further evaluation-specific data be collected. In order to maintain a collaborative approach to the evaluation, it was decided to abandon the idea of a specific process evaluation and to rely on regional reports to document the programme, interviews and surveys to provide formative feedback, and the existing data collection procedures to monitor the programme.

### ***Survey of Services***

The aims of this survey were to: (a) provide formative evaluation data for future programme development; and (b) provide data for the impact evaluation of the programme.

**Research Design** Pre and post surveys of service providers were conducted twelve months apart (Time 1 – January 1996 and Time 2 – January 1997) in the

### *Evaluation 7(1)*

four Programme Areas and one Control Area. It should be noted that these surveys were cross-sectional in nature, rather than longitudinal. Due to methodological difficulties, in particular high rates of attrition, a longitudinal design was considered to be inappropriate for a survey of service providers.

**Participants and Sampling Procedure** For both the pre and post surveys, an attempt was made to identify all services in each of the Programme Areas who had the potential to come in contact with a young person at risk aged 10 to 24 years. The types of services matched those targeted for interviews and included mental health services, schools and youth services, medical, other community services, Aboriginal and Torres Strait Islander services, places of worship and legal services.

Identification of services for the Time 1 sample was achieved through previous regional visits by the evaluation team, programme proposals, local 1995 phone books, and the Youth Services Information System (YSIS) database. The lists were then added to, and amended by the programme team members in Areas 1, 2 and 3 and the health department office of the Control Area, to provide a comprehensive list of services in each area.

The database compiled for the Time 1 sample was used as a starting point for identification of services at Time 2. The database lists of services for each region were added to, amended and updated using the most recent available lists of regional services provided by team members in Areas 1, 2, 3 and 4, and the 1996 local telephone books. The outcome of this procedure was the identification of a cross-sectional sample of services at Time 2, which was similar though not identical to the lists of the services identified at Time 1. All of the identified services received a mailed questionnaire, with the exception of places of worship. Due to the large number of churches and other places of worship identified, a randomly selected group of 85 percent was surveyed.

**Data Collection Procedure** The data collection process was identical for both the 1996 and 1997 surveys. All of the identified services in the four Programme Areas and the Control Area were mailed a survey package which consisted of a survey form, a covering letter, an information sheet and a reply-paid envelope. The services were requested to fill in the survey form and post it back in the reply-paid envelope. Return rates for mail surveys are typically low unless an effort is made to follow up non-responders. Two follow-up procedures were implemented in order to maximize response rate. A reminder letter was sent to non-respondents one month after the survey mail-out. After a further month, non-respondents received another reminder letter and another copy of the survey and reply-paid envelope. Services that did not plan to complete the questionnaire were requested to return it blank in order to avoid being sent reminder letters and a replacement questionnaire. Each service also received a letter to thank them for their participation.

A total of 116 services in Area 1, 265 in Area 2, 184 in Area 3, 164 in Area 4 and 271 in the Control Area (Total  $N = 1000$ ) were mailed a survey package at Time 1; 137 services in Area 1, 377 in Area 2, 194 in Area 3, 178 in Area 4 and 348 in the Control Area (Total  $N = 1234$ ) were mailed survey packages at Time 2.

**The Questionnaire** The questionnaire contained five sections. Section 1 covered questions about the type of service provided for young people. Section 2 included questions about the young people who came in contact with the service. Section 3 asked about opportunities for, and participation in, training activities, specifically crisis intervention for young people and ongoing counselling for young people. Section 4 covered questions about support for staff, opportunities for networking and informal debriefing and protocols for dealing with young people at risk of suicide or self harm. Section 5 asked questions about the programme. As the programme had only recently commenced when the initial survey was conducted, Section 5 of the Time 1 survey contained questions regarding areas such as awareness of the programme, potential of the programme and expectations regarding what the programme had to offer the service. For the subsequent survey, Section 5 included questions which covered the services' experiences with the programme in their area, such as how successful they thought the programme had been and how it had assisted them.

**Participation Rates** Across the four regions, 25.0 percent and 25.4 percent of services did not return the questionnaire at Time 1 and Time 2 respectively. In addition, 16.1 percent and 22.9 percent of the service providers at Time 1 and Time 2 respectively returned the questionnaire blank. Overall, 541 services (54.1 percent) at Time 1 and 578 services (46.8 percent) at Time 2 returned a completed questionnaire.

**Use of Results** Detailed results are not discussed here but general patterns are identified. The survey was useful in identifying the target areas of the programme teams in each area. Across the four areas, they had concentrated most heavily on working with schools, youth services, mental health services and community services. Although this focus was considered to be appropriate due to the short time-frame, it was important to acknowledge that the original goal of including a wide range of services would take more time to reach. In particular, medical services needed more attention, given that young people in crisis are often identified in these settings. Mental health services appeared to have somewhat unfavourable views of the programme and it was recommended that more attention be paid to building good links with them. In general, services in Programme Areas were more likely to have been offered training in dealing with young people at risk compared to services in the Control Area. The survey also usefully identified barriers to provision of training for service providers which would need to be addressed by programme teams trying to increase the number of service providers participating in training opportunities.

The survey also reflected the different emphases of separate programme areas. For instance, Area 2 showed an increase in opportunities for informal debriefing and networking and Area 3 showed an increase in services with a formal procedure for dealing with young people at risk. The survey also provided the opportunity for services to comment on their interactions with the programme. Both positive and negative comments provided useful formative evaluation to the programme teams. Positive comments were encouraging and highlighted areas to

### *Evaluation 7(1)*

maintain, whereas negative comments were useful in providing recommendations for future improvements.

### ***Queensland Young People's Mental Health Survey***

**Rationale** This study was designed to target long-term improvements in young people's mental health. Although unable to provide specific information about the programme at this early stage, it was intended as a baseline with which to compare future data collection exercises. It was also intended that the results of the survey would provide information to the programme teams and other services about young people in Queensland.

**Study Design and Participants** The study was designed as a cross-sectional household survey, involving a multi-stage sampling approach. Telephone recruitment was followed by an anonymous self-report postal questionnaire with telephone follow-up. The target population for the survey was 15 to 24 year old males and females living in the state of Queensland.

Young people were recruited into the study over a five-month period. A total of 78,108 telephone calls were made to 35,509 households during the study. Of the 4594 eligible households that were contacted, a returned completed questionnaire was received from 3092 young people, resulting in an overall participation rate of 67.3 percent.

**Questionnaire** The survey asked about young people's: (a) access to and use of mental health services; (b) preferred sources of help; (c) past suicide attempts and suicidal ideation; (d) mental health issues such as depression, delinquency and mental illness; and (e) predisposing and precipitating factors associated with past suicide attempts and suicidal ideation.

**Use of Results** The results of this study will be published separately. They do not contribute directly to this evaluation, although they provide baseline data for comparison with future surveys. The results provided information to programme staff about young people in general. As well as using the information themselves, programme staff were able to build links with other service providers by disseminating the information informally.

### **Strengths and Weaknesses of the Evaluation Plan**

#### ***Cluster Evaluation***

This evaluation took the form of a cluster evaluation, an '... evaluation of a programme that has projects in multiple sites aimed at bringing about a common general change' (Sanders, 1997: 397). A basic assumption underlying this design is that each project is relatively autonomous in the planning and implementation of their programme. However, the key to success is in good relationships between the parent organization, each programme team and the evaluators. It is also imperative that each of these organizations or individuals has a clear idea of the



concepts relevant to the programme (Sanders, 1997). Cluster evaluations focus on broad issues and multiple perspectives and have the advantage of being able to be responsive to stakeholder needs. However, there is a danger of evaluators losing their independence because of the amount of interaction necessary between parties. There is also the problem of coping with constant evolution and development of the programme, the need to have full co-operation of all concerned and the limited ability to look for long-term change (Sanders, 1997).

The importance of a shared clarity of understanding of programme concepts prompted us, as evaluators of the programme, to take a proactive role in the articulation of specific aims and objectives. We considered the need for clarity of aims to be so crucial for the design of the evaluation that we decided to become more involved with programme direction than we would have ideally preferred. The benefit of this strategy was a set of agreed aims on which to base the evaluation plan although critics may argue that the independence of the evaluation team was unsuitably compromised by this process (Scriven, 1997). However, given that the programme was about empowering the community, we viewed our role as facilitators, empowering programme teams to shape their own programmes and have a say about evaluation matters.

### ***Mixed Methods***

The use of mixed methods in evaluations is widespread, to the extent that Riggan (1997) questions the value of treating them as a distinct type of evaluation. Cook (1997) presents the position that both qualitative and quantitative methods are legitimate in evaluation, the frequency of their joint use limited by financial and logistical considerations rather than logical difficulties. More extensive dissemination has been recommended regarding practical, contextual and particularly consequential results of mixed-method evaluations (Datta, 1997).

Greene and Caracelli (1997) summarize three major stances regarding mixed-method evaluation. They discuss the purist stance, which argues that different paradigms of inquiry represent fundamentally distinct assumptions about knowledge. Thus, the mixing of methods of inquiry within one study is not acceptable. On the other hand, two alternative positions advocate the use of mixing methods. The first is the dialectic position which acknowledges the importance of different paradigms of investigation and argues that they should be used together in order to enhance understanding of social phenomena. The second position is the pragmatic view (Patton, 1990; see also Datta, 1997), which sees the philosophical assumptions entailed in different paradigms of inquiry as logically independent. Thus, it is appropriate to mix different paradigms and methods in order to meet the practical demands of the situation. The pragmatic position argues that a combination of different methods of inquiry is necessary to investigate complex social phenomena comprehensively. Datta (1997) proposes that a pragmatic approach means that '... the essential criteria for making design decisions are practical, contextually responsive, and consequential' (p. 34).

The philosophical underpinnings for the evaluation of this cluster evaluation began from the dialectical position, but quickly engaged a heavy pragmatic outlook in response to consultations with programme teams. It was clear from the outset



### *Evaluation 7(1)*

that a mixed-methods design would be required in order to fulfil the extensive requirements of the evaluation agreement. The design included distinct components in a predominantly expansion framework (Greene and Caracelli, 1997; see also Wholey et al., 1994) where different methods were used for different elements of the evaluation, for example formative, process, impact and outcome aspects. However, there were also elements of complementarity (Greene and Caracelli, 1997), where qualitative information was used to clarify and contextualize quantitative results. Similarly, quantitative data were used to supplement qualitative data.

Chen (1997) suggests that mixed methods are appropriate where there is a '... context requiring information that is both intensive and extensive, offers high access to some information but low access to other information and has characteristics of both open and closed systems' (p. 64). This was unquestionably the case for this evaluation setting.

### ***Evaluation Purpose***

The potential use of the evaluation report and the collected data was a consideration in the design of the evaluation plan. Chelimsky (1997) discusses evaluation purpose in terms of the way the findings will be used. She outlines the accountability perspective, where the purpose is to measure results in connection with the funds expended, and generally forms the basis for policy, governmental or agency reform. There is also the knowledge perspective where the purpose is to generate insights about problems, programmes and processes used for research, education or knowledge base construction; and the developmental perspective where the purpose is to strengthen institutions or build capability in some area, often used internally as part of the evaluative process.

These three perspectives differ in key dimensions, such as the role of the evaluator and the independence and objectivity of the evaluation. Evaluations taking the accountability perspective prize high objectivity and a distant, independent evaluator, whereas uncertain objectivity may be acceptable for evaluation for a developmental or formative purpose, with the evaluator taking the role of a 'critical friend' who may even be part of the team. The challenge for the evaluation of this programme was that the evaluation agreement alluded to all three of these purposes. Thus, although measures could be taken to maximize the validity of distinct components of the evaluation in a way that was appropriate for the particular methods employed, particular challenges were presented regarding the general role of the evaluators and their relationships with the programme teams.

### ***Evaluator Role, Independence and Objectivity***

There is much debate about the nature of truth and the value of objectivity in research in general and in evaluation specifically (Scriven, 1997; Shufflebeam, 1994). While a strength of mixed-method evaluations relates to their ability to draw on the strengths of differing paradigms to address diverse research questions, a problem arises in terms of meeting standards of methodological rigour relating to diverse strategies (Chen, 1997). The presence of systematic errors and/or random inconsistencies may be difficult to address when diverse methods are implemented.

Scriven (1997) argues that validity is seriously compromised by a close relationship between evaluators and programme teams, because sources of bias such as personality clashes, prejudice and other personal feelings inevitably influence the evaluation. On the other hand, it has been argued (Fetterman, 1996) that less distancing between evaluators and programme staff results in enhanced data quality, more likelihood that the evaluation will be accepted and less likelihood that the evaluation will have any destructive consequences. Scriven (1997) counters by asserting that 'tempering validity with mercy (or the like) is a violation of validity' (p. 483).

Although the extensive requirements of the evaluation itself were specified in the evaluation agreement, the expected role of the evaluators was not. As outlined above, we were involved from the beginning in developing specific aims and objectives. We endeavoured to maintain a consultative and collaborative approach to the evaluation throughout, discussing the evaluation plan and its rationale regularly with teams, and considering any objections raised or modifications suggested. Programme teams also had input into several evaluation components. For example, they provided data about possible research participants, in some cases choosing interviewees. Programme staff also provided direct evaluation data through the interviews conducted with present and past programme staff. Although Scriven (1997) argues that this type of evaluation activity presents particular difficulties for validity and objectivity, we believe that our inclusive approach increased the relevance and acceptability of the evaluation.

Scriven (1997) suggests that obtaining non-interactive feedback from senior members of programme teams may provide an opportunity for comment, but reduce the likelihood of personal factors influencing the evaluation design. Certainly, our face-to-face contacts with some teams were occasionally hostile and often lively; however, at other times, discussions resulted in benefits for all parties, whether in the form of literature exchange, increased contacts or useful background information about aspects of the programme. Of particular benefit was the improved access that the programme staff gained to current literature about suicide prevention strategies and research studies through the evaluation team.

Problems with bias in evaluation may also occur where the evaluator role overlaps with a helping role, or where evaluators are called upon to provide recommendations regarding the programme (Scriven, 1997). We provided assistance to the programme teams, both informally as outlined above and formally (for example, formative evaluation results, consultation regarding monitoring). The evaluation agreement called for recommendations about all aspects of programme functioning. Moreover, we occasionally took a more pro-active role (for example, facilitating articulation of specific aims and objectives), particularly where our interpretation of scientific literature indicated an ethical difficulty.

This was particularly relevant to matters relating to youth suicide where 'common sense' prevention strategies may, in fact, increase the problem – for example, there are concerns that certain types of strategies may normalize or romanticize suicide (Mason, 1990; Schaffer et al., 1990), or 'common sense' indicators may not be relevant (for example, comparing suicide rates, as mentioned

### *Evaluation 7(1)*

above). In this way, the evaluation team was not independent and could be seen to be biased towards the programme; however, we were more concerned with pragmatic and ethical issues. There would be no credit in an objective evaluation which concluded that the programme was not a success because the suicide rate of young people increased by 33 percent<sup>1</sup> if there had been an opportunity to prevent the suicides in the first place. Ethical standards required for general research concerning human beings must also be applied to evaluation research.

## **Lessons Learned and Recommendations for Future Evaluation Design and Implementation**

### ***Focus Broadly***

A key feature of the programme was breadth. It aimed to cover a huge geographical area, deal with controversial and emotive issues and engage diverse organizations and individuals. From the perspective of the evaluation team, this was both a strength and a weakness. It was a strength because it provided a large pool of activities on which to focus the evaluation. But it was also a weakness because the theoretical effect of programme activities was diffuse and difficult to measure empirically.

Our solution was to focus the evaluation broadly. A range of discrete, focused outcome measures were identified, for example, proportion of services offered training. This allowed the collection of specific empirical data. However, the type and number of these outcome measures were such that they allowed a broad view of the programme. Similarly, a wide range of people were interviewed. This allowed the formal presentation of qualitative interview data, giving a broad range of the views of young people themselves, service providers and programme team members.

Although numerous measures were examined, potentially useful indicators of programme success have been omitted. One limitation of our approach is that there was no attempt to directly measure community capacity or the sustainability of programme effects. One reason for this was difficulty in identifying and defining the 'community'. There was no one community, rather the programme addressed an overlapping set of various communities.

### ***Collaborate***

Theoretically, an independent evaluation team has the advantage of an objective view of the programme to be evaluated. In practice, any evaluator has many pre-conceived ideas which may influence the design, implementation and outcome of the evaluation. We found our consultations with programme team members invaluable in designing a suitable evaluation plan. It is vital that the philosophy informing community development programmes is applied to their evaluation (Israel et al., 1998). Programme teams and community members should be empowered to participate in the evaluation of their activities and outcomes. In our experience, consultation with programme team members was time-consuming and at times frustrating but ultimately led to the development of a more relevant,

useful and successful evaluation than would have emerged from a distanced, objective approach.

### ***Be Flexible***

It is not always possible to adhere strictly to philosophical and theoretical considerations when confronted with the reality of a living programme to evaluate. In real life, pragmatic and ethical concerns may overthrow original plans. Flexibility was an essential element in the design of this evaluation. A consultative or collaborative approach means being willing to change designs and plans in response to feedback (Kelly, 1990). This does not mean that every suggestion from anyone is followed because this would be impractical and may lead to a confused evaluation with no overall structure or plan. However, it was important to be flexible in response to feedback, for instance regarding modification of survey instruments or the addition of qualitative interviews.

As well as internal programme influences, there were external factors which necessitated a flexible approach. Evaluating a broad community development programme over several years means having to modify plans in response to structural and political changes in the community. A particularly influential structural change which occurred mid-programme was a change of state government and the subsequent reorganization of regional boundaries. Although programme functioning was not disrupted substantially, the boundary changes had considerable implications for the proposed survey of young people's mental health. Originally intended to provide detailed regional data, the survey was expanded to the entire state because, although more expensive, the state-representative data were considered more valuable than information sampled on the basis of out-of-date regional areas.

### ***Use Multiple Methods and Approaches***

The use of multiple methods was necessary because of the many types of information required (Israel et al., 1998). The evaluation provided formative evaluation data as well as collecting impact and outcome measures. The need for qualitative and quantitative data necessitated the use of multiple methods.

At times, the requirements of the programme teams for formative evaluation data were in conflict with the requirements of the government department providing the money. Whereas the programme teams primarily found qualitative data most useful, it was clear that further support for the programme would be based on success in demonstrating a measurable programme impact. The application of an outcome-driven approach to the evaluation of community development programmes is generally not realistic in the short term, and this must be acknowledged in the requirements of accountability.

Therefore, this evaluation plan represents a marriage between an outcome-driven approach and a participatory action research approach (Penuel and Freeman, 1997; Rains and Ray, 1995). It has endeavoured to pursue both perspectives in order to provide objective indicators of success to funding bodies, while involving community partners (in this case programme teams and, to some

### *Evaluation 7(1)*

extent, service providers) in the evaluation. Although the result of this hybrid approach was satisfactory, there are potential pitfalls. Funding bodies may view the time and resources spent in community consultation as excessive, while community members may resent the use of objective indicators of programme success. Clear communication is needed to build trust on all sides (Israel et al., 1998).

## **Conclusion**

In conclusion, the multi-method approach discussed in this article has offered a potentially useful model for evaluating community development programmes. The careful identification of realistic, broadly focused outcomes measured with a quasi-experimental pre-post design allowed the demonstration of programme effects on service provision for young people. On the other hand, the extensive consultation process and detailed interviewing of key stakeholders provided in-depth accounts of features of programme functioning that would otherwise have been missed.

## **Acknowledgements**

This evaluation was funded by the Queensland Department of Health. We would like to thank the many programme staff, service providers, community members and young people who contributed to the programme and its evaluation. We would also like to thank two anonymous reviewers for their helpful comments on an earlier draft of this article.

## **Note**

1. Given that one area recorded only 12 suicides during the baseline period, this could have been a conclusion if the comparison period showed a rise of just 4 individuals.

## **References**

- Abatena, H. (1997) 'The Significance of Planned Community Participation in Problem Solving and Developing a Viable Community Capability', *Journal of Community Practice* 4(2): 13–34.
- Bogensneider, K. (1996) 'An Ecological Risk/Protective Theory for Building Prevention Programmes, Policies, and Community Capacity to Support Youth', *Family Relations* 45: 127–38.
- Butterfoss, F. D., R. M. Goodman and A. Wandersman (1993) 'Community Coalitions for Prevention and Health Promotion', *Health Education Research* 8: 315–30.
- Chavis, D. M. and G. M. H. Pretty (1999) 'Sense of Community: Advances in Measurement and Application', *Journal of Community Psychology* 27(6): 635–42.
- Checkoway, B. (1998) 'Involving Young People in Neighbourhood Development', *Children and Youth Services Review* 20(9–10): 765–95.
- Chelimsky, E. (1997) 'The Coming Transformations in Evaluation', in E. Chelimsky and W. R. Shadish (eds) *Evaluation for the 21st Century: A Handbook*. Thousand Oaks, CA: Sage.
- Chen, H. (1997) 'Applying Mixed Methods Under the Framework of Theory-Driven



- Evaluations', in J. C. Greene and V. J. Caracelli (eds) *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms*, No. 74 in the series 'New Directions for Evaluation', a publication of the American Evaluation Association. San Francisco, CA: Jossey-Bass.
- Cook, T. D. (1997) 'Lessons Learned in Evaluation Over the Past 25 Years', in E. Chelimsky and W. R. Shadish (eds) *Evaluation for the 21st Century: A Handbook*. Thousand Oaks, CA: Sage.
- Datta, L. (1997) 'A Pragmatic Basis For Mixed-Method Designs', in J. C. Greene and V. J. Caracelli (eds) *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms*, No. 74 in the series 'New Directions for Evaluation', a publication of the American Evaluation Association. San Francisco, CA: Jossey-Bass.
- Dixon, J. and C. Sindall (1994) 'Applying Logics of Change to the Evaluation of Community Development in Health Promotion', *Health Promotion International* 9(4): 297–309.
- Duignan, P. and S. Casswell (1989) 'Evaluating Community Development Programmes for Health Promotion: Problems Illustrated by a New Zealand Example', *Community Health Studies* 13(1): 74–81.
- Fawcett, S. B., R. K. Lewis, A. Paine Andrews, V. T. Francisco, K. P. Richter, E. L. Williams and B. Copple (1997) 'Evaluating Community Coalitions for Prevention of Substance Abuse: The Case of Project Freedom', *Health Education and Behaviour* 24(6): 812–28.
- Fetterman, D. M. (1996) 'Empowerment Evaluation: An Introduction to Theory and Practice', in D. M. Fetterman, S. J. Kaftarian and A. Wandersman (eds) *Empowerment Evaluation: Knowledge and Tools for Self-Assessment and Accountability*. Thousand Oaks, CA: Sage.
- Francisco, V. T., A. L. Paine and S. B. Fawcett (1993) 'A Methodology for Monitoring and Evaluating Community Health Coalitions', *Health Education Research* 8(3): 403–16.
- Goodman, R. M., M. A. Speers, K. McLeroy, S. Fawcett, M. Kegler, E. Parker, S. Rathgeb Smith, T. D. Sterling and N. Wallerstein (1998) 'Identifying and Defining the Dimensions of Community Capacity to Provide a Basis for Measurement', *Health Education and Behaviour* 25(3): 258–78.
- Greene, J. C. and V. J. Caracelli (1997) 'Defining and Describing the Paradigm Issue in Mixed-Method Evaluation', in J. C. Greene and V. J. Caracelli (eds) *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms*, No. 74 in the series 'New Directions for Evaluation', a publication of the American Evaluation Association. San Francisco, CA: Jossey-Bass.
- Gruenewald, P. J. (1997) 'Analysis Approaches to Community Evaluation', *Evaluation Review* 21(2): 209–30.
- Harrison, J., J. Moller and S. Bordeaux (1997) 'Youth Suicide and Self-Injury, Australia', *Australian Injury Prevention Bulletin, National Injury Surveillance Unit* 15(March): 1–8.
- Hawe, P., D. Degeling and J. Hall (1990) *Evaluating Health Promotion: A Health Worker's Guide*. Sydney: MacLennan and Petty.
- Israel, B. A., A. J. Schulz, E. A. Parker and A. B. Becker (1998) 'Review of Community-Based Research: Assessing Partnership Approaches to Improve Public Health', *Annual Review of Public Health* 19: 173–202.
- Kelly, J. G. (1990) 'Changing Contexts and the Field of Community Psychology', *American Journal of Community Psychology* 18(6): 769–93.
- McLaughlin, M. J., P. E. Leone, S. Meisel and K. Henderson (1997) 'Strengthen School and Community Capacity', *Journal of Emotional and Behavioral Disorders* 5(1): 15–23.
- Mason, G. (1990) *Youth Suicide in Australia: Prevention Strategies*. Canberra: Department of Employment, Education and Training (Youth Bureau).



- Moscicki, E. K. (1995) 'Epidemiology of Suicidal Behaviour', *Suicide and Life-Threatening Behaviour* 25(1): 22–35.
- Mrazek, P.J. and R. J. Haggerty (1994) *Reducing the Risks for Mental Disorders: Frontiers for Preventive Intervention Research*. Washington, DC: National Academy Press.
- Nee, D. and M. I. Mojica (1999) 'Ethical Challenges in Evaluation with Communities: A Manager's Perspective', in J. L. Fitzpatrick and M. Morris (eds) *Current and Emerging Ethical Challenges in Evaluation*, No. 82 in the series 'New Directions for Evaluation', a publication of the American Evaluation Association. San Francisco, CA: Jossey-Bass.
- Paine-Andrews, A., K. J. Harris, S. B. Fawcett, K. P. Richter, R. K. Lewis, V. T. Francisco, J. Johnston and S. Coen (1997) 'Evaluating a Statewide Partnership for Reducing Risks for Chronic Diseases', *Journal of Community Health* 22(5): 343–59.
- Patton, G. C., R. Harris, J. B. Carlin, M. E. Hibbert, C. Coffey, M. Schwartz and G. Bowes (1997) 'Adolescent Suicidal Behaviours: A Population-Based Study of Risk', *Psychological Medicine* 27: 715–24.
- Patton, M. Q. (1990) *Qualitative Evaluation and Research Methods*. Thousand Oaks, CA: Sage.
- Penuel, W. R. and T. Freeman (1997) 'Participatory Action Research in Youth Programming: A Theory in Use', *Child and Youth Care Forum* 26(3): 175–85.
- Poole, D. L. (1997) 'Building Community Capacity to Promote Social and Public Health: Challenges for Universities', *Health and Social Work* 22(3): 163–70.
- Rains, J. W. and D. W. Ray (1995) 'Participatory Action Research for Community Health Promotion', *Public Health Nursing* 12(4): 256–61.
- Riggin, L. J. C. (1997) 'Advances in Mixed-Method Evaluation: A Synthesis and Comment', in J. C. Greene and V. J. Caracelli (eds) *Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms*, No. 74 in the series 'New Directions for Evaluation', a publication of the American Evaluation Association. San Francisco, CA: Jossey-Bass.
- Sanders, J. R. (1997) 'Cluster Evaluation', in E. Chelmsky and W. R. Shadish (eds) *Evaluation for the 21st Century: A Handbook*. Thousand Oaks, CA: Sage.
- Schaffer, D., V. Vieland, A. Garland, M. Rojas, M. Underwood and C. Busner (1990) 'Adolescent Suicide Attempters: Response to Suicide-Prevention Programmes', *Journal of the American Medical Association* 264(24): 3151–5.
- Scriven, M. (1997) 'Truth and Objectivity in Evaluation', in E. Chelmsky and W. R. Shadish (eds) *Evaluation for the 21st Century: A Handbook*. Thousand Oaks, CA: Sage.
- Shufflebeam, D. L. (1994) 'Empowerment Evaluation, Objectivist Evaluation, and Evaluation Standards: Where the Future of Evaluation Should Not Go and Where it Needs to Go', *Evaluation Practice* 15(3): 321–38.
- Silverman, M. M. and R. W. Maris (1995) 'Epidemiology and Risk Factors: 1. The Prevention of Suicidal Behaviors: An Overview', *Suicide and Life Threatening Behaviour* 25(1): 10–21.
- Stevenson, J. F., R. E. Mitchell and P. Florin (1996) 'Evaluation and Self-Direction in Community Prevention Coalitions', in D. M. Fetterman, S. J. Kaftarian, and A. Wandersman (eds) *Empowerment Evaluation: Knowledge and Tools for Self-Assessment and Accountability*. Thousand Oaks, CA: Sage.
- Vincent, M. L., A. F. Clearie and M. D. Schluchter (1987) 'Reducing Adolescent Pregnancy Through School and Community-Based Education', *Journal of the American Medical Association* 257: 3382–6.
- Wholey, J. S., H. Hatry and K. Newcomer (1994) *Handbook of Practical Evaluation*. Thousand Oaks, CA: Sage.

*Lucke et al.: A Mixed-Method Cluster Evaluation Design*

JAYNE LUCKE is a research psychologist with interests in programme evaluation, particularly in the areas of young people's mental and sexual health. Currently working in the University of Manchester Medical School, she lectured in psychology at the University of Huddersfield between 1997 and 1999. At the time of this study she was employed in the Department of Psychiatry at the University of Queensland. Her other research interests include gender roles and the use of narrative methods to investigate sexual experiences. Please address correspondence to: Dr Jayne Lucke, 4 Railton Avenue, Whalley Range, Manchester, M16 8AU, United Kingdom. [email:jayne.c.lucke@man.ac.uk]

MARIA DONALD joined the Centre for Primary Health Care, Department of Social and Preventive Medicine, at the University of Queensland, as a lecturer in 1997. She is a community psychologist and her principal areas of research include young people's health behaviours and the processes involved in building resilient communities to promote health and well-being. The focus of Maria's recent work has been on building the capacity of communities to respond to young people's health-related issues and increasing the sustainability of community-based health programmes. She is interested in the development of databases and measurement tools, and the design of evaluations for health interventions, particularly for community-based mental health promotion and prevention initiatives.

JO DOWER is a community psychologist who is currently employed by the Centre for Primary Health Care in a research capacity. She has conducted extensive research and evaluation, particularly in the areas of mental health and community health. Jo has a particular interest in the area of health promoting environments and the evaluation of health-related initiatives. Jo's research interests focus on efforts to identify and address the wide array of social and behavioural determinants of health. Her research draws on both qualitative and quantitative methods and recognizes the value of integrating both approaches to address complex public health problems.

BEVERLEY RAPHAEL is Director of the Centre for Mental Health for New South Wales and Emeritus Professor in Psychiatry from the University of Queensland. She also holds professorial appointments at The University of New South Wales and The University of Sydney. She has a long-term interest in prevention in mental health, with special reference to child and adolescent mental health, family issues and issues affecting consumers in mental health care systems, and in a public health approach to mental health. She has also been involved in the development of Aboriginal Mental Health Policy and has contributed in the field of trauma, grief and disasters.