

CHAPTER 13

MEASURING EFFECTIVENESS

Gabrielle Maxwell

Perhaps the greatest research challenge in the field of violence intervention is to measure the effectiveness of programmes. In part this is because of our limited knowledge about what leads to and sustains violence and how to mitigate its impact on others. Designing research to further our understanding of family violence is a much simpler task than determining if it has changed in response to an intervention. In part the latter is more difficult because of the complex matrix of factors that affect research designed to answer questions related to the operational goals of service funders and providers.

A central dilemma is that research designed to measure effectiveness must probe beyond current understandings and be open to alternative explanations. These may differ from those that have, of necessity, been adopted by providers and funders who have made a commitment to a model they believe will enable them to change behaviour. It is essential that those who are providing programmes or funding them already have their own views on the reasons for family violence, the impact it has on others, how best the violence can be prevented and how those on whom it has impacted should be responded to. However, these theories take for granted the issues that evaluation research must question. And the perspectives driving programme provision will often affect the way that those offering or funding programmes construe the appropriate role of the researcher, how they should or should not collect data and what opportunities should be arranged for evaluators to have access to clients and their records. Providers must develop guidelines that enable them to protect clients' interests, but these can interfere with the needs of research to obtain data and contact clients independently of the programme.

There are more practical and obvious problems facing the researcher measuring programme effectiveness. These include the specific research

objectives set, what data can be obtained, what design should be used for assessing change over time, and the formal framework in which the research will be conducted. Each of these three major topics is considered in turn.

Research objectives

It is not enough that programmes are formed according to specified guidelines, or that they meet the legislative or institutional objectives set down for the process and funding of services. What is crucial is that outcome objectives should be met. In the area of family violence, the targets are often broad, complex and difficult to operationalise. For instance, in the three programmes we evaluated over the period 1998 to 2001, the following objectives were set for the evaluations:

1. To determine the extent that specialist services for child victims and witnesses of family violence are meeting the children's needs and to describe key features of successful practice and best practice (Shepherd and Maxwell 1999). Key aims for all programmes were to:
 - (a) Successfully prevent and ameliorate any likely harm to children resulting from their experience of family violence;
 - (b) Provide a range of successful therapeutic services to children to respond to individual needs;
 - (c) Deliver the service in a culturally appropriate way;
 - (d) Involve parents so that there is a transfer of learning to the home environment; and
 - (e) Assist in removing barriers to the child's development that result from family violence.
2. To determine whether programmes for women (defined as adult protected persons under the Domestic Violence Act 1995) contribute to the protection of victims of family violence (Maxwell et al 2001).
3. To determine the effectiveness of community based violence prevention programmes for men and to inform policy development and service provisions including providing an assessment against what is currently regarded as 'best practice' (McMaster et al 2000).

The breadth, complexity and, at times, the near impossibility of the expectations of the programmes is immediately apparent from these evaluation goals. For example, the first project required an evaluation of success in meeting needs of children. This meant identifying relevant needs and assessing children with respect to them on at least two occasions

to determine change. The aim in objectives a) and b), that the needs identified should be those linked to the experience of family violence for that individual child, did not seem possible given current knowledge or techniques of assessment. The extensive therapeutic goals seemed beyond the capacity of relatively short programmes, two of which were attended by groups of children either for seven or ten weeks.

In addition, each of the evaluation specifications required information on a wide variety of other process issues such as assessment and selection procedures at entry, programme details including content and attendance, cultural issues around programme delivery, throughput and costs of programme provision.

What data can be obtained

The second major issue is about the nature of the data that will be obtained. The usual options are to:

- Examine programme records
- Interview providers and other community stakeholder organisations
- Interview clients and significant others
- Obtain data from other relevant agencies.

In the three studies described above, the expectation was that programme records would be obtained and clients would be interviewed at entry to a programme, on exit and after a follow-up period. Interviews with family, key community stakeholders and providers were also part of the requirements. This would seem to cover all the options possible for such a study. Yet even such a comprehensive approach may be insufficient. Some of the obstacles to obtaining satisfactory data are described below, along with some of the problems encountered:

1) Will the base records be available? In practice we found that many of the programmes had not kept records that enabled us to determine:

- How many clients entered and exited the programmes
- How many sessions and which ones were attended by each client
- Information on needs at entry, assessments of progress and evaluation of outcomes on exit.

These problems arose despite apparently satisfactory visits at an early stage of the research to consult with providers, to request co-operation and to discuss record keeping needs. The project where we obtained the best records was one where a condition of funding was co-operation with the research and where a member of the provider group was part of the

advisory committee during the research. However, even in this project one provider dropped out of the research because of difficulty in meeting any of the research requirements and had to be replaced at a late stage by another programme on which, because of the limited time available, less data could be collected.

2) Can clients be successfully recruited, located and interviewed?

We are proud of our reputation for successfully locating, recruiting and interviewing people who are generally among the most mobile, traumatised and alienated groups in society. Nevertheless, these difficulties are an ever-present part of our work and these three studies were no exception. In the evaluation of the children's programmes, the goal was to collect data on 30 children at six sites (n=180). In practice, only 65 parents agreed to be interviewed when asked by the providers if they would take part, and only 52 were available at exit and 48 at follow-up. These difficulties often reflected the life situation of the mothers who had also been victims of family violence and they are characteristic of the attrition rates in such samples. They also related to the optimistic expectations of programmes about recruitment and retention of clients.

Similarly the target of 120 men in the stopping violence evaluation was not reached. Smaller than expected programme numbers and the need to replace one programme at a late stage meant that only 83 men were assessed at entry. The target interview number for women who partnered these men was 80, but only 41 were successfully interviewed. The programmes did not always know how to contact the women and were cautious about recruiting them, so as to avoid jeopardising their safety.

In the third programme, smaller targets were successfully met by extending the recruitment period but, as in all the samples, attrition from entry to exit and follow-up limited any useful comparisons between programmes which could have informed conclusions about the relative success of different programme models or identify best practice factors. These recruitment difficulties inevitably limited the ability of the research to successfully detect change over time and to draw reliable conclusions about the effectiveness of programmes.

3) What data should be obtained? If we interview people, what are the right questions to ask? Can we obtain suitable and reliable data on behaviour? Surprisingly, asking the right questions of those involved in family violence is one of the easier tasks because of the wealth of published research and because this is the sort of research task where professional

researchers who know the particular area can be selected on their ability to demonstrate their skills in their research applications. Furthermore, previous research has shown a relationship between self-reported violence and violence assessed by other means, such as the reports of partners or criminal notifications.

On the other hand, obtaining information on violence behaviour that can validate self-report is not often practical, especially with small samples and with such short timeframes. Data from police records on reported breaches of non-violence orders or assaults will inevitably be selective and could be skewed depending on the circumstances of the offence and the willingness of a partner to report violence to the authorities. Self-report is, therefore, usually the only practical alternative for relatively small-scale studies over relatively short periods of time.

4) Will the data gathered be reliable? We have already seen that obtaining information on a sufficient number of clients to answer all the questions posed by the research was not possible in any of these studies. But there are other important questions about the quality of data. The study of men's stopping violence programmes was able to compare men's and women's responses on scales designed to assess the degree and type of violence in their relationships and changes over time. These results showed that perceptual difference did occur on some scales for men and women but that, on other scales, there were similar responses:

A comparison of men's and women's responses indicated that the men were reporting less involvement in emotional abuse, controlling behaviour and very serious abuse than the sample of women partners. However, the men's reports of using physical and sexual violence were not significantly different from those of the women. Furthermore, the reports of the women endorsed the men's reports of decreased abuse and violence over time: most of the women reported positive changes in the men's behaviour which they attributed to the programme and which were maintained when they were interviewed three months later.

(McMaster et al 2000)

These results indicate how important it can be to make comparisons that check the perceptions of different players in evaluation studies in the area of family violence. Unfortunately, too often the resources available for the research preclude this.

5) Will the type of data collected be useful and sufficient to meet the research objectives? Not all the data gathered in evaluations

actually proves to be particularly useful in determining how effective the programmes are. In the three studies examined here questions can be raised about the value of aspects of the detailed programme descriptions and the key stakeholder interviews.

In each of the studies considerable information was collected about the programmes, the qualifications of staff, the content of sessions, the history of the programme, the interagency connections of the programme and the programme philosophy. All this data is important from a descriptive point of view and much of it is necessary as a context to the research. However, in practice, the details obtained did not enable possible best practice issues indicators to be identified with the specificity that would enable them to be related to outcome criteria. Our experience is that empirical data on best practice related to outcomes can only be collected from observations of critical aspects of the programme in the context of research designed to compare a variety of practice over relatively large numbers of clients. An example is research currently being finalised on 'Achieving effective outcomes in Youth Justice' which examines the outcomes for over 1000 young people over a period of more than two years who were involved in family group conferences with 24 specified co-ordinators on whom information was available about routine practices and programme philosophy. In addition, observations on practice were available for 18 of these practitioners (Maxwell et al 2002). Such large-scale research is likely to be the only effective option for defining objectively the key best practice factors in family violence intervention.

The only practical alternative is to gather information by self-report on the views of practitioners, stakeholders and clients about views on best practice. The risk is that such information may identify features about which there is consensus but which are not crucial to outcomes. For example, the psychological literature has identified aspects of therapist behaviour which lead to positive ratings but which do not appear to predict successful outcomes.

6) Is all this information necessary? A related issue is that of collecting data that proves to be of little use for any purpose. In all three of the studies reviewed here, key stakeholder interviews were carried out. The stakeholders represented the main agencies making and receiving referrals from the programmes. Many were nominated as agencies with whom the programmes had the greatest contact. However, many of those interviewed were unable to answer the questions most relevant to issues

of effectiveness with clients. The comments that were made tended to reflect the interviewees' views of the personnel in the programme or the philosophy underpinning the programme rather than provide an objective assessment. Comments on outcomes for clients tended either to reflect experiences with speed of processes or to be anecdotal and based on one or two cases. In terms of value for money, this aspect of the research probably produced the least useful data.

Design issues

The usefulness of the data collected depends on other design factors that will impact on the value of the research. The most important of these are questions of comparison.

1) The use of control and comparison groups. Ideally, a control group should be selected of people who are randomly assigned to a programme or a no-programme group. Both groups should be assessed over a period equal to at least six months after programme completion to determine whether or not there are changes in longer-term outcomes, but such designs are ethically problematic as they deny the possibility of assistance to those in need.

A more acceptable control group strategy is to assign randomly some volunteers to a waiting list for a shorter period, for example a period equal to the length of the programme so that changes can be assessed at the end of the programme period for both groups. However, there are cost issues around this solution and the ethical issues are not entirely resolved by the period of delay.

Alternatively, comparison groups of people can be selected from other areas where programmes are not available or from people who have not taken up the option to enter a programme. In our evaluation of groups for abused women, 38 women who met the criteria for programme referral under the Domestic Violence Act 1995 agreed to be interviewed as a comparison sample. In terms of basic characteristics, the two samples were similar in most respects. However, the comparison women reported similar changes in perceived safety over time and reported using similar safety strategies to the women on the programme, indicating that the programmes may not be the only way in which the women can develop safety strategies. This could be explained by the finding that the comparison women were more likely to report receiving support from others, especially family and friends, were more likely to have left their abusive partner and were

more likely to be in new relationship. Further, some of the comparison group had individual counselling. Nevertheless, it remains true that the programme groups reported that the programmes had helped them, they provided for the needs that led them to choose to attend and they believed the programme had been responsible for them learning how to keep themselves and their children safe (Maxwell et al 2001).

The above comparisons between the programme and the comparison group are ambiguous. Questions about programme effectiveness remain, in some senses, unanswered. On the one hand, the programme group attributes their changes to the programme and, given the needs that led them to join and attend the programme, they are almost certainly right. On the other hand, the comparison group were able to achieve many of the same outcomes in other ways. Given the supports that they had, the fears some held attending a group programme and the dislikes of others for the particular philosophy of the programme, these women probably made the right choice not to attend.

These findings indicate that finding the right road to recovery from family violence will be an individual choice. The search for ‘the effective solution’ is not realistic. Different options will meet the needs of different women. Such a finding is reinforced by other findings in the same study which suggest that:

- some, but not all, of the Māori women would have preferred a programme based on tikanga Māori
- some of the women on the programme still felt the need for other support such as individual counselling or ongoing group meetings.

Given these complexities, it is clear that a lot can be learnt from comparison group designs but that they will not necessarily give a definitive answer about the effectiveness of the programmes.

2) Designs that make comparisons over time. Another strategy is to compare the responses of programme clients over time. In all three studies reviewed here, the methodology required assessments of clients at entry to the programme, at the end of the programme and after a follow-up period. These designs can be criticised because they raise the question of whether the changes are due to the programme, changes in other life circumstances or simply the passage of time. The strength of these designs is that they allow for comparisons of clients of different backgrounds, preferences and patterns of need and they can be an effective way of comparing programmes with different approaches, providing clients do

not self-select and that there are adequate numbers in the samples in each type of programme.

In our studies we struck a variety of complicating factors, a major one being that of programme homogeneity. In both the men's and women's programme studies, all but one adopted similar overall philosophies and presented themselves as operating on a modified Duluth model, using the power and control wheel as a central tool and emphasising the safety of the victims of family violence as a primary outcome.

Programme heterogeneity can also be a problem. When programmes are different in several respects, it is difficult to be certain about what any differences in outcomes can be attributed to. The programmes we examined varied in the characteristics of their clients and in several aspects of practices, e.g. in the number of sessions that they offered, in session length and in whether or not family/whānau were involved. Thus when there are differences in outcomes it is not clear what the decisive factors are. For example, one men's programme was distinctive in general approach and in adopting tikanga Māori, but as its clients were different in many important respects, it was difficult to know whether it was the approach or the clientele that determined the different outcomes. Answering questions about the relative value of these many different potential programme features would only be possible with large scale designs that allow for multivariate analysis.

3) Designs that compare outcomes for different client groups. The discussion so far has emphasised the centrality of obtaining information on the effectiveness of programmes in achieving their goals of stopping family violence and protecting and strengthening victims. Another goal that was part of the requirements of the studies reviewed here was to determine what types of clients are most responsive to various programme types. While this is yet another complicating factor for small scale research designs, answers can sometimes be suggested by examining self-reports of programme effectiveness for clients with different characteristics. In the men's programme study, the tikanga-based approach of the Māori programme was endorsed by all the Māori men attending it and the programme showed a greater success than other programmes in recruiting, retaining and enabling change for some of the most violent men and men from the most troubled backgrounds. At the same time, Māori men and women in other programmes did not always express a desire for a tikanga-based programme. It is difficult to determine if self-selection is

not operating at the point people elect to enter a programme or if the lack of a comparison point led people to prefer the type of programme that had proved successful for them.

4) Recruitment and retention of clients. A central problem for all studies of programmes is that not all clients who enter a programme actually attend from start to finish. Some leave the programme before it is complete and others will miss substantial parts of it. If a critical aspect of effectiveness for programmes is the planned and graded sequential aspects of their sessions, drawing conclusions from research samples who do not necessarily experience the whole of the intended programme is problematic.

Similarly, those recruited to take part in the research may decide to opt out of the research or be untraceable at the time of the second and third interviews. Both these factors can limit the pool of data available at the end of the project. Carrying out additional interviews at the point of entry to the programme is a strategy that can ensure sufficient end-point data if project budgets allow it. Alternatively, we have found that comparing the responses of samples of clients at entry and at exit usually appears to show the same overall patterns even if not all the same individuals were assessed both at entry and at exit. Thus recruiting extra clients to the research at programme conclusion is an option for making up for attrition in samples over time.

Support issues

Some very practical issues are central to the life of a research project. These may seem trivial compared to the larger issues of deciding which data to collect, how, when, from whom and how to interpret it. Yet these housekeeping issues make up much of the life of the evaluation researcher.

1) Adequacy of funding and appropriateness of timeframes. These two aspects of any research project will limit what is possible. It is unusual for all the goals of the project to be readily met within the cost and time budgets set by agencies who themselves have limited budgets to contract providers and evaluate services. It is always tempting to see the provision of more services as preferable to extended evaluation. Both those funding and those undertaking research have a responsibility to attempt to negotiate a realistic contract in which expectations about what can be achieved are balanced with the resources available.

2) The management process. There is normally a management or advisory committee for a research project and a lead person who is able to make most of the day-to-day decisions as the evaluation proceeds. Both timeframes and costs of research can be placed under great pressure if the management process is not consistent, timely and supportive. Processes that involve large numbers of people sharing and making day-to-day decisions about research processes and instruments can damage the quality and timeliness of evaluation, particularly when key decisions are delayed or reversed. At the same time it is essential that evaluators manage costs and timeframes within agreed guidelines and keep the lead representative of the management group fully informed about relevant events that could impact on the contracted goals and targets.

3) The relationship with provider organisations. Providers and evaluators come from different perspectives. Providers, to be effective, must be committed to their professional philosophy, be clear about how programmes can best be delivered and passionate to ensure that they provide a service to the best of their knowledge and belief. They will be protective of the best interests of their clients, especially when the safety of those clients can be placed at risk by any breaches of confidentiality.

Evaluation researchers must be sceptical. They need to be committed to, indeed obsessive about, collecting quality data on as many of those involved in the programmes as possible at the optimal time to meet the requirements of the design. They need to scrupulously eliminate inaccuracies and verify as much of the data they collect as possible.

In many ways, the two roles appear fundamentally at odds. Yet both have common ground in a desire to achieve the programme objectives and better understand how to improve current practice and both will share a commitment to avoiding harm to the programme clients. A mutual commitment to working together is essential to a successful evaluation project and often also to continued programme funding. Each party must, therefore, accommodate the needs of the other. Providers will need to ensure that researchers can contact and recruit clients to the research, and researchers must do so respectfully and confidentially. Providers will need to allow researchers access to their records, but researchers must minimise the data-gathering burden placed on providers.

4) Interviewers. Obtaining quality data means recruiting, training and supporting appropriate and effective interviewers who can treat those who participate with respect and obtain quality data. We have found that

best practice is to have at least two people interviewing in the same area at the same time to provide support for one another and, depending on interviewee preference, a choice of interviewer is possible. In general, only women should interview women who have been victims of family violence or abuse. Other things being equal, it is preferable for Māori to interview Māori and for Pacific peoples to be interviewed by those of similar ethnicity. But data recently obtained (Maxwell et al 2003) indicates that an experienced interviewer who can create a rapport with interviewees is more important than an interviewer of the same ethnicity.

Protocols and guidelines are needed for interviewers to ensure that they meet best practice standard and that providers can be reassured about the sensitivity with which the evaluation responds to the critical issues of respect, safety and confidentiality. Examples of guidelines we developed for the three evaluation studies of programmes for men, women and children are presented in the appendices of the three reports (cited at the end of this chapter).

5) Cultural consultancy. Obtaining quality input from cultural advisors or consultants is not a simple matter. The relative scarcity of people with the appropriate skills and background to provide quality comment means that the demands on their time are considerable. Yet such advice is important to ensuring that the needs of those of different cultures are accurately interpreted. Building the cultural expertise of relevant agencies and making it available to evaluators may be one way to improve the quality and relevance of advice. Providing adequate resources in evaluation budgets for purchasing such advice is also important.

THE RESULTS OF EVALUATIONS: THREE CASE STUDIES

Difficulties and possible solutions can be seen by examining the three evaluation studies focusing on children, women and men that have already been referred to. A short description of their methodology and conclusions is presented as a case study before returning to the dilemmas faced in programme evaluation and, most importantly, trying to answer the question: ‘What is good enough?’

1. Children

Six programmes catering for child victims or witnesses of family violence were evaluated. Three were primarily group programmes, providing weekly sessions for 7 to 10 weeks for children aged 5 to 12. Three were individual programmes for a wider age range over a variable period that

averaged six months.

Information for the evaluation was obtained from the six programmes themselves, five key agencies involved with each programme and interviews with parents and children. At four of the sites interviews were carried out with 65 parents and 45 children as the children entered a programme. When possible, additional interviews were conducted on exit (or three months after entry whichever was sooner) and at follow-up after another three months. Because of a relatively low recruitment rate it is not possible to generalise the results of this evaluation to all those who took part in programmes. Results are presented below under headings related to evaluation objectives.

Meeting children's needs

At exit, two-thirds of parents indicated that the programme had definitely helped. The few parents who responded negatively said that their child had made little or no improvement or that they were worse. The majority of parents (83%) reported positive changes. Parents described improvements in behaviour, self-esteem, expressing feelings, ability to communicate, controlling anger, understanding family violence and relationships with family and friends.

At follow-up, parents reported that over half of children had generally improved after the programme had finished. There was, however, a small proportion of parents who felt their children had not improved or were worse.

A comprehensive needs checklist was used to assess the needs experienced by children and to assess change from entry to exit and to follow-up. The checklist covered six developmental areas: feelings and behaviours, relationships, education, identity, health, and social presentation and was used to examine overall need as well as specific areas of difficulty. A comparison of parents' reports at entry and exit indicated a significant reduction in children's overall needs. There was also a lesser reduction from exit to the time of the follow-up interview three months later, although this did not reach significance. There was no significant difference in the amount of reduction for Māori and non-Māori. Children's replies also indicated that overall their needs had decreased over time — however, the number of replies was only 20 and the amount of improvement was not statistically significant.

Data was also presented on the specific areas of need which were most responsive to the programmes. The greatest improvement, for about

two-thirds to three-quarters of those with the problem, was in relation to being worried or anxious, feeling frightened, and aches and pains. The number reported as having problems related to sadness, anger, aggression and destructiveness was also substantially reduced, by about half. The one problem reported relatively frequently which showed virtually no improvement was trusting others. There was also relatively little change in difficulties concentrating, getting into trouble at school or home, considering others' feelings and being restless or fidgety. Perhaps this is not surprising as these difficulties may relate to experiencing violence as a very young child (Fancourt 1999). Such difficulties are likely to require more extensive intervention to be resolved. It is also possible that trust in people could be slow to return if there are still some unresolved issues of safety for these children.

Items from the needs assessment checklist which were least likely to be reported for these children included substance abuse, stealing, getting in trouble with the police and running away or truanting. The low incidence of these items may reflect the relatively young age of the children in the sample. In addition, the lack of problems with sexualised behaviour may reflect the fact that most children's exposure was not to sexual abuse. There was also a low incidence of more severe signs of pathology such as self-harm, suicidal feelings, soiling and a lack of interest in eating or learning, which suggests that most of these children were not severely disturbed and hence unlikely to respond to a group programme or to relatively short-term counselling.

Key features of successful programmes

Interviews with parents and children asked about features of the programmes identified as important for successful outcomes.

- Two-thirds of parents reported that they were involved in the programme
- Over half helped with planning
- 56% learnt something themselves
- 98% trusted the staff
- Eight out of ten indicated that their child had enjoyed the programme
- Of the ten parents who identified as Māori or indicated that their children identified as Māori, seven said that the programme was good for Māori children. The remaining parents said they did not know or did not answer

- At least 70% of children said they had learnt about safety issues and not to blame themselves for family violence
- About a third of children said they had learnt about family violence; in particular, they described learning about safety plans and not taking responsibility for violence
- Nine out of ten children said that they trusted staff but less than half the children said they had helped with planning
- Over two-thirds of children indicated that they liked the programme a lot and none endorsed not at all
- A few children were unable to remember what they had learnt and four said they had learnt nothing or very little.

Best practice

For the most part it was not possible to determine which features of the programme resulted in the best outcomes for children. There were a range of reasons for this: sample numbers were small at most sites; overall, there was very little difference between those sites where children's needs were able to be examined at entry into the programme and at exit and follow-up; and the programmes were very similar in their objectives. However, some conclusions have been drawn based on the views of programme providers and key agencies, responses from children's and parents' interviews, and the results of previous research. Best practice will involve ensuring that:

Parents:

- are included in planning what will happen for their child and are informed about what is happening for their child
- receive information and support for their own needs.

Children:

- need to feel safe in the programme and at home
- need to understand key concepts, including: the nature of family violence; that they are not to blame; that other children have had similar experiences; and how to keep themselves safe in future
- need to be responded to in ways that are appropriate to their developmental level, to have fun and learn new skills
- should have a role in deciding what should happen for them
- need to have their special needs recognised.

Programmes should ensure that:

- the needs of children are assessed on entry
- a holistic and multi-level approach is taken that responds to a variety of needs in a variety of contexts

INNOVATIVE APPROACHES TO STOPPING FAMILY VIOLENCE

- they are accessible to parents in different circumstances and from different cultural and social backgrounds
- groups are children who can feel safe together and be open with one another
- children as well as their families have ongoing support after the programme.

Staff should:

- be skilled in building rapport, experienced in working with children and appropriately trained
- include Māori facilitators/counsellors in programmes for Māori children
- receive supervision and support.

Agencies need:

- secure and sufficient funding
- good inter-agency communication
- feedback on referrals
- evaluation of their programmes.

Summary

Most responses from parents and children indicated that many of the needs that existed before entering the programme were met during the programme and that these gains were sustained at follow-up three months later. In addition to children learning about family violence and safety issues, other behavioural and emotional needs such as worries, fears, aches and pains, feeling sad, anger, aggression and destructiveness were substantially reduced.

However, not all the children's needs were met. For some there was little change, and a few reported deteriorating behaviour. It is possible that these children displayed more acting-out behaviour, which seems to be a particular problem for the group programmes. These children may benefit from individual counselling. Continuing child and family support is also likely to be helpful for the children whose needs remain unmet at the conclusion of the programme. Research also recommends that programmes should be part of wider interventions and involve primary caregivers, and that children who have experienced severe and repeated family violence require individual treatment in addition to any group. It is likely that developing these aspects of service provision will enhance their effectiveness in responding to children.

For the present, however, it is clear that group programmes and individual counselling can help children recover from many of the adverse effects of family violence and that the changes observed are consistent with the prevention of long-term harm. It seems likely that in New Zealand, as elsewhere, providing programmes for children who have witnessed or experienced family violence may be an important part of breaking the cycle and reducing the costs of harm. However, to succeed in assisting the children who have been harmed by violence, it is likely to be important to develop a variety of programmes to meet different needs and to further research what is likely to be most effective.

2. Evaluating Women's Living Without Violence programmes

This research used both quantitative and qualitative methods to assess the performance of two programmes for women who had experienced domestic violence. Data were obtained from 40 women who had attended the programmes in 1988 and 1999; 21 women who were eligible but had not attended; programme providers and key community and government agencies that the programmes worked with in each area. Women were interviewed when they entered a programme, when it ended and approximately three months later. First and second interviews were held with the comparison group at approximately the same times as the exit and follow-up interviews for the programme women. Both groups were asked the same questions about what they wanted for themselves and about barriers to accessing programmes. In addition, the programme group were asked about expectations and gains from the programme, while the comparison group were asked about reasons for not attending a programme.

The results can be summarised under the objectives identified in the Domestic Violence regulations and other target objectives for the evaluation.

Promoting the protection of persons from domestic violence

- Approximately 90% of the women attending the programmes reported that they had learned how to keep themselves safe, and their children safe and had developed a safety plan for themselves and their children. The percentages of women who reported these gains from the programme were higher than in the comparison group.
- There was a significant reduction in the incidence of all types of abuse reported over the period of the programme compared with experiences

of abuse in the past, but this was also true for the comparison group.

- The programme women felt safer at exit compared to entry interview but this change was slight and did not reach significance. Similar slight but insignificant changes in perceived safety scores were noted for the comparison group women.
- However, many women were not completely confident that the person who had previously abused them would not re-enter their life and abuse them again and this was not under their control.

The comparison group women reported similar changes in their perceived safety over time and they also reported using similar safety strategies. This may indicate that the programmes are not the only way for women to develop safety strategies. One important strategy used by all the comparison and some of the programme sample was obtaining a protection order — for many that certainly contributed to their perceived safety. It was also seen by the women as a way of taking active steps to keep themselves safe. Significantly more of the comparison group women also reported that they were receiving support and assistance in doing this from others, including family, friends, lawyers, courts and other agencies.

Achieving specific programme goals

On measures designed to assess specific goals in regulation 28 of the Domestic Violence regulations, approximately 90% or more of the women reported gains except on a few topics that may have been relevant to only some of them. Particular topics with lower reported gains were: obtaining information about protection orders (two-thirds reported gains), programmes for children (81% reported gains) and programmes for men (26% reported obtaining this information).

Identifying best practice

The nature of the data in this study did not permit best practice standards to be drawn from comparisons across programmes. Conclusions about best practice were, therefore, drawn from the views of women, providers, key agencies and published sources.

- **Meeting the needs of women** to develop key understanding about family violence, to have the opportunity for personal growth and to learn new skills, to be provided with appropriate and relevant information to meet individual needs, to assist them to keep themselves safe and to be supported. Both programmes have met these standards

except that they are not always able to meet special needs for individual counselling and make referral to other appropriate services. A second difficulty is that the programmes have not always been able to provide the support needs of some clients after the programme has concluded.

- **Programme characteristics** of being accessible to potential clients, being culturally responsive and including quality assessment procedure. Both programmes reported some problems in these areas. Resources limited the options for programme venues and assistance with transport and child care. Recruitment of Māori facilitators was a difficulty for the programmes at the time of this study. Neither programme carried out full and comprehensive assessments and evaluations of all clients, although individual pre-programme assessment interviews were conducted for all clients at one of them. Ongoing informal assessment was carried out at the other and both programmes asked evaluative questions in the final session.
- **Staff characteristics:** staff selected as suitable, provided with appropriate training, supervised and supported. Although the programmes appeared to be meeting these standards at the time of this evaluation, both programmes reported difficulties with resources to recruit, train and retain staff.
- **Organisational characteristics** including having secure and sufficient funding that enables an effective infrastructure to be built, developing good interagency networks, and monitoring and evaluating the programme. Neither programme reported having sufficient or secure funding to carry out the range of tasks that were perceived as necessary for best practice.

Overall, the two agencies met standards well. However, their limited funding and resources contributed to difficulties or limitations in completely meeting women's needs, providing for the range of necessary organisational support and arranging for formal assessment, monitoring and evaluation procedures.

Identifying factors affecting take-up and attendance

— Take-up

The replies of the comparison group indicated that about a third did not yet feel ready to attend a programme. About a quarter said that they did not know that they could attend a programme and another quarter reported

that they felt safe. A small group reported that a major reason was that they did not know where they would be living. Attitudes and personal circumstances were also undoubtedly a major factor in the decision not to attend. Barriers to take-up included psychological barriers such as concerns over sharing their experiences with others; practical difficulties of timing, transport and so on; and fears in relation to their partner. The comparison group contrasted with the programme group in reporting having good support from family and friends: indeed, this was the most important difference between the two groups.

— Strategies for effectively informing women

Many women in the comparison group endorsed the possible option of being sent a letter when the order was made. At least a quarter said they would also appreciate a phone call and nearly one in four said they would like a second letter when an order was made final.

— Factors leading to drop-out

Nearly two-thirds of the women were not able to attend all the sessions. Common reasons were illness, work, the needs of children and lack of transport. A small proportion of women mentioned the attitude of their partner or fear of their partner finding out about their attendance. The small number who completed fewer than five sessions dropped out because they did not feel comfortable in the programme. They were embarrassed at having to talk about what had happened to them, did not feel comfortable with the facilitators, found what was discussed upsetting, did not feel they were getting enough out of the programme, felt under pressure to leave their partner, or that the programme was not culturally suitable. These findings indicate the need for a variety of options for women so that they can choose a programme that is likely to best suit them in terms of theory and process.

— Examining the extent to which programmes met the needs and values of Māori clients

Both programmes aimed to respond to the needs of Māori women through building facilitators' knowledge and understanding of tikanga Māori and working with other Māori groups in the community. They included features that were designed to make the programme comfortable for Māori women such as using mihi, karakia and concepts of whānau, and had attempted (unsuccessfully) to recruit and train Māori facilitators. However, the replies of Māori women in the sample suggested that these

actions were appreciated but were not sufficient to make the programme consistent with a Māori kaupapa. Some of these women would have preferred a Māori programme where they could deal with issues in a Māori way. They saw renewing cultural links and affirming cultural identity as part of the healing process.

These findings notwithstanding, the data also indicated that many Māori women found that being in a programme where there was a cultural mix enabled them to recognise that domestic violence occurs in all cultures. The data also indicated that, in general, they shared the same goals for themselves and made the same gains as the Pākehā women. These are important findings because they indicate that any programme that is designed for Māori women should also provide for the same needs as those described here: stopping the violence is a primary goal for the women.

Conclusions

Caveats need to be made about these results because of the relatively small samples and the differences between the comparison and programme samples. Nevertheless, these findings provide evidence that the programmes help women understand how to keep themselves and their children safe and enable them to develop safety plans. Reasons for not attending programmes reflect practical difficulties, the attitudes of the women towards the programmes and a lack of information about them. The views of Māori women and community agencies indicate the need for the development of kaupapa Māori programmes for Māori women who have been victims of family violence. The programmes reviewed here are for the most part meeting the best practice standards suggested by New Zealand practice, opinions and overseas research. However, limits on resources and funding make it difficult for all the standards to be met, especially with respect to the special needs of individual women and requirements for ongoing support.

3. The evaluation of community-based Stopping Violence programmes for men

This study assessed four different programmes using both quantitative and qualitative methods. Data was obtained from 83 men entering the programmes, 41 of their women partners, programme providers and key community stakeholders. Questions for the men and their partners assessed the extent of the men's violent and abusive behaviour prior to, during and after the programme. Also included were questions about background

factors in the lives of the men, current lifestyle, mental health, alcohol use, attitudes to women and help-seeking behaviour. The main findings of this study can be summarised as follows.

- Group programmes of approximately 16 weeks of 2-3 hours duration can be effective in working with men to change abusive and violent behaviour. The results reported fully in the 2000 report listed at the end of this chapter indicate that the change measured in the men's behaviour tended to be across the board: in emotional abuse, controlling behaviour, threatening behaviour, property damage, physical abuse, sexual abuse and very serious violence. Those who had the highest scores before tended to change more than those who had lower scores, perhaps because they had more room to change. This result is interesting because it is sometimes suggested that the most abusive men are the least likely to change.
- Many men are reluctant to engage in programmes, and are even less likely to engage in evaluation research. It is possible that those who were prepared to participate in the evaluation were more committed to change in the first place and were therefore more likely to sustain their involvement in the programmes. Comparing men who entered the programmes with those who did not, and men who participated in research with those who did not, would provide important additional information on the groups for whom programmes are likely to be most effective.
- Many of the Māori men indicated that they would value programmes that included tikanga concepts and Māori processes.
- Some beliefs in the family violence field are not supported by this research. Theories that men downplayed their violence prior to undertaking the programme but recognised its true extent once they had been on it do not receive support in this study. Men described their behaviour reasonably fully at entry to a programme, indicating both that they knew the extent of their abusive behaviours and were willing to identify them. Also, the belief that men's emotionally abusive behaviour gets worse as their physical abuse reduces was not, on the whole, supported in this research. The results indicated that levels of abuse of all kinds showed an overall decline between entry and exit and diminished even further at follow-up.
- The Power and Control Scale developed for this research has proved to be a reliable and valid measure of men's abusive and violent

behaviour. It provides information on violence of different kinds, can be used over different time periods and can be answered by both men and their partners. It could also be a useful assessment tool for programmes and provide a way in which they could routinely measure their effectiveness.

- This evaluation has shown how difficult it is to obtain useful data on programmes through independent evaluation. An alternative strategy for informing policy and practice is to set up a standard accountability structure for all programmes contracted by government agencies, based on an examination of assessment information on clients at entry, change information on clients at exit, facilitators' assessments of clients at exit and, where possible, women's views of the changes in their men's behaviour.

The results reported here are very positive but the numbers in the samples are quite small, especially in the women's sample, and the follow-up period was short. Ongoing research that builds on these findings is needed to examine the effect of different approaches to programme delivery and to draw firm conclusions about why some men respond more than others.

Dilemmas

So far this chapter has traversed dilemmas that surround evaluation research. In summary, these include:

- Inevitable differences in perspective of programme providers and evaluators can hinder an effective working relationship. The commitment to an approach that is consistent with effective service does not easily marry with the scepticism that must accompany independent evaluation. The shared goal of determining what works best for clients is necessary to effective collaboration.
- Providers' need to ensure they provide a safe environment can be seen as conflicting with sharing information with evaluators. However, evaluators must also develop effective protocols and procedures to ensure client safety and privacy. A partnership approach to these goals can lead to effective relationships in which both parties are able to trust one another. It is important that providers see evaluators as allies.
- The breadth and complexity of evaluation goals is often inconsistent with practical design frameworks, relatively short timeframes and limited funding. It is important that both funders and researchers work

together to develop frameworks for evaluation that are practical and achievable. It is important to collect both quantitative and qualitative data because these are not alternatives, rather they provide different kinds of information. It is also important that, despite the inevitable limitations associated with any particular evaluation, good planning can usually identified useful findings.

- Problems in collecting data can be greater in family violence research than in many other arenas. Those who participate in programmes are often at a point where their lives are in disarray, their emotions distressed and their capacity to manage their lives damaged by the violent events of the recent past.
- Support for the evaluation team is important. Funding, timeframes and provider relationships are two critical issues that often limit what can be achieved by an evaluation team. Other factors that can limit or enhance the quality of an evaluation are the quality and nature of the relationship with the management or advisory group, the skill and stability of the interview team and the quality of cultural advice that can be drawn on.
- A contract research environment and a focus on relatively short-term programme evaluations is not consistent with undertaking research that focuses on prediction, developing systems for identifying and monitoring indicators of quality practice and for collecting and maintaining data using standard instruments in a variety of different programmes over time. New ways must be found to ensure that larger-scale studies with broader goals are able to be undertaken effectively.
- Special issues will be encountered with specific groups of clients. In this chapter we have explored some that can emerge when examining generic programmes providing services to Māori clients. The study of children's programmes examined some of the issues around working with siblings. The study of women's programmes discussed some of the difficulties in providing services for women from other cultures, particularly refugee women and those who are not native English speakers. Other papers in this volume discuss issues of providing services to people who have experienced violence in same-sex relationships, gangs and adolescents. Evaluating these programmes will also have special problems.

Increasing the effectiveness of evaluations

Given all these potentially limiting factors, what can be gained through evaluation? The three studies described here all suffered from some of the difficulties and dilemmas that have been identified. Yet they were all able to identify goals being achieved by the programmes, at least for some of the clients. It was also possible to identify improvements which could be made and to suggest some principles that could guide best practice.

However, most programme evaluations are unable to retain and follow up sufficient samples over long enough periods using measures indicative of the desired long-term outcomes. Yet without such data it is not possible to measure the extent of violent behaviour by the men and the impact of the programmes on the lives of the women and children. Nor has current research identified indicators that are effective predictors of future behaviour.

As I indicated at the start of this chapter, comparisons are problematic in programme evaluation. Control groups are rarely feasible. Comparison groups of those who choose not to attend are a practical, useful alternative, although non-attenders are likely to differ from the programme group in important ways. Another highly desirable comparison would be to build up data over time by using the same scales, questionnaires and checklists in different evaluation studies, to enable meta-analysis.

The challenge for the future is to develop evidence-based practice and to monitor programmes on critical indicators that will ensure we enhance the outcomes for victims of family violence and reduce its incidence in the future. The first step will be a commitment by everyone involved to increasing the amount and scope of evaluation research.

The three reports referred to in this chapter:

Maxwell, G, Anderson, T & Olsen, T (2001) *Women living without violence: An evaluation of programmes for adult protected persons under the Domestic Violence Act 1995*. Wellington, Ministry of Justice.

McMaster, K, Maxwell, G & Anderson, T (2000) *Evaluation of community-based Stopping Violence prevention programmes*. Research Report prepared for the Department of Corrections.

Shepherd, P and Maxwell, GM (1999) *Evaluation of specialist services for child victims and witnesses of family violence*. Final report to the Children, Young Persons and their Families Agency and the Crime Prevention Unit.

The Maxwell et al study on *Achieving Effective Outcomes in Youth Justice* referred to here is now being prepared for publication in 2003. The full reference is obtainable from cjrc@vuw.ac.nz