

A Guidebook for Evaluating Community Based Drug Law Enforcement Projects

Richard Midford¹, Simon Lenton¹, Kevin Boots², John Acres¹, Wendy Loxley¹, Chris Canty³, Steve James³ and Adam Sutton³

1 National Drug Research Institute, Curtin University of Technology

2 Formerly National Drug Research Institute, Curtin University of Technology

3 Department of Criminology, University of Melbourne

This research was funded by the National Community Based Approaches to Drug Law Enforcement Board of Control

May, 2000

TABLE OF CONTENTS

TABLE OF CONTENTS.....	I
Tables.....	ii
Figures.....	ii
ACKNOWLEDGEMENTS.....	iii
CHAPTER 1: EVALUATING COMMUNITY BASED DRUG LAW ENFORCEMENT PROJECTS: SOME PRELIMINARY CONSIDERATIONS.....	
Introduction.....	1
Why do an evaluation?.....	1
What’s different about evaluating a harm reduction approach to drug law enforcement?.....	5
CHAPTER 2:PRACTICAL ISSUES TO BE CONSIDERED IN EVALUATING COMMUNITY BASED DRUG LAW ENFORCEMENT PROJECTS	
Who should be involved and how?.....	9
What problems are likely to come up and how can these can be overcome?.....	12
CHAPTER 3: EVALUATION PARADIGMS AND APPROACHES	
INTRODUCTION	19
The relative merits of the scientific and naturalistic paradigms	19
The relationship between paradigm assumptions and evaluation methods	21
The relationship between assumptions and evaluation techniques and tools	23
Conclusion	23
CHAPTER 4: TYPES OF EVALUATION.....	
Formative evaluation	25
Process evaluation.....	26
Participant evaluation.....	28
Impact evaluation.....	28
Outcome evaluation	29
Selecting the appropriate type of evaluation.....	29
Overview.....	30
CHAPTER 5: TECHNIQUES FOR DATA COLLECTION AND ANALYSIS.....	
Introduction.....	31

Measurement and Evaluation	31
Sampling.....	32
Sample Size	34
Obtrusive or Overt Measures	36
The Limitations of Obtrusive or Overt Methods.....	40
Unobtrusive or indirect measures.....	40
Alternative Categorisations of Measurement Methods.....	42
CHAPTER 6: THE PROCESS OF EVALUATION.....	45
Initiation	45
Detailed planning	47
Data collection and management	50
Data analysis	51
Reporting.....	51
CHAPTER 7: WRITING THE REPORT AND MAXIMISING ITS INFLUENCE ON DECISION MAKING.....	53
Introduction.....	53
What is an evaluation report?.....	53
The audience	53
Report format	54
Disseminating the report and maximising its input to decision making	57
References	59

TABLES

Table 1: Report format	55
------------------------------	----

FIGURES

Figure 1. Example of a Data Description Matrix.....	16
Figure 2 Sample Size Chart	36

ACKNOWLEDGEMENTS

The Authors would like to thank Mr Any Brock for his helpful comments in regard to the NCBADLE material included in this evaluation guide. The authors acknowledge that some of the material contained in this evaluation guide has been derived from a Health Promotion Program Evaluation study guide, written for Edith Cowan University, by R. Midford, N. McBride, and K. Boots.

CHAPTER 1: EVALUATING COMMUNITY BASED DRUG LAW ENFORCEMENT PROJECTS: SOME PRELIMINARY CONSIDERATIONS

INTRODUCTION

Who should read this guidebook?

This guidebook has been written for those undertaking evaluations, such as that associated with the National Community Based Approach to Drug Law Enforcement (NCBADLE) Community Based Drug Law Enforcement project, which involve a police/health interface. It has been prepared for those who already have some expertise and experience in conducting research and evaluation, rather than the complete novice.

What should you get from reading the guidebook?

In writing this book we have tried to combine the conceptual and the practical in a way that will guide how to set up and conduct useful evaluations of innovative community based drug law enforcement programs. We hope that readers will get a frame of reference for the evaluation process, a practical introduction to evaluation options, the tools of evaluation and the process of applying them. In addition, we hope to convey from our experience, the particular issues for evaluation of programs, which focus on innovative law enforcement responses to illicit drug use.

What readers won't get is a prescriptive, step by step 'Evaluation Cookbook' nor an omnibus, 'Everything you wanted to know about evaluation (but were afraid to ask)'. Where possible, however, we have attempted to complement our overview and experience with references to other useful sources.

WHY DO AN EVALUATION?

While the intent of law enforcement giving greater emphasis to community harm would have few detractors, programs have not always achieved their aims and in some cases, have had completely unintended consequences. Some initiatives in the past have fallen into the trap of relying on being self-evidently 'worthwhile', without undergoing a formal evaluation. Undertaking a community policing activity, simply because its intention is good, is difficult to justify in terms of effective service

provision, let alone scientific rigour. In order to avoid this mistake, community policing interventions need to demonstrate and document effectiveness in terms of their stated aims. Evaluation is the means by which this can be done. Building evaluation into new interventions allows you to determine how to develop an intervention best suited to the circumstances; how the intervention worked in practice, who benefited and how; and where modification could be made to improve future practice.

Evaluation as good professional practice

Evaluation is an important part of any social intervention program. Evaluations should be custom designed and an integral component of the undertaking from the beginning. As Hawe, Degeling and Hall (1992) indicate, evaluation is part of a commitment to good practice, because it allows you to monitor the progress of your program and determine if it is fulfilling your intentions.

It is easy to get overwhelmed by the demands and conflicts associated with evaluation. The program funders want to know what has been achieved. You need to report what you are doing in the name of administrative accountability. There is pressure to focus resources on program delivery rather than evaluation. There are so many different types of evaluation, how do you decide what best suits your program? In dealing with all of this and making decisions about how to conduct your evaluation, remember that evaluation is not an end in itself. Most important, according to Pirie (1990), is asking the right questions and presenting the results in a way that usefully informs future decision making.

What is evaluation?

Casswell and Duignan (1989) consider that evaluation can be simply described as asking the question of a project:

Is this the best way of doing this? (Casswell and Duignan, 1989, page 7)

However, the 'best way' depends on a number of factors, including the perspective of the evaluator.

Hawe, Degeling and Hall (1992) use Suchman's definition of evaluation:

evaluation is the process by which we judge the worth or value of something (Hawe, Degeling and Hall, 1992, page 6)

They consider that evaluation is a judgement of something, which in turn depends on what is considered important. This affects various aspects of how an evaluation is conducted, including what data are gathered and how results are interpreted. Evaluation may seem to be an objective process, but in reality it is not. As somebody who may be involved in undertaking or interpreting evaluation of community policing initiatives in relation to drug use, it is important to be able to critically evaluate an evaluation in terms of its purpose: what questions were addressed and why; whose interests were served; what methods were used; and how were results presented.

Green (1990) attempts to address the issue of judgement subjectivity in his definition of evaluation. He considers that evaluation involves comparing the object of interest with the standard of acceptability. Without a comparison against some benchmark, even an imaginary ideal or a subjective preference, there is no evaluation. This definition is sufficiently broad to encompass evaluation of both processes and outcomes and does not suggest that evaluation is objective, merely that some standard of reference is required, however this is derived.

What do you get from evaluation ?

Hawe, Degeling and Hall (1992) take a minimalist position on evaluation in their assertion that it should be done to at least ensure that an intervention is not making the problem worse. However, in terms of best practice, it is important to go beyond this and find out what an intervention is achieving. These authors also see evaluation as providing rewarding feedback on progress to those with a stake in a project and as a way of demonstrating the value of their work. In a similar vein, McDermott, Pyett and Hamilton (1991) consider that evaluation creates an understanding of the various meanings that an intervention has for all those involved.

Pirie (1990) states that evaluations should be conducted in a way that makes them useful when decisions are to be made about programs. However, she notes that it is important to realise that evaluation results are not the only inputs into decisions. Funding, political considerations, staffing, changed priorities etc. all play a part in decisions about a community policing program. The particular contribution of

evaluation information to the decision making process is its focus on implementation and achievement; what the program is doing or has done.

Rossi and Freeman (1985) consider that good program evaluation is capable of providing answers to five important questions.

- Is the intervention reaching the appropriate target population?
- Is it being implemented in the ways specified?
- Is it effective?
- How much does it cost?
- What are its costs relative to its effectiveness?

Windsor et al (1994) however, consider that the reasons to undertake evaluation are broader. These authors cite ten major purposes of evaluation.

- Determining the attainment of program objectives.
- Ascertaining the strengths and weaknesses of program elements for planning purposes.
- Monitoring performance standards and establishing quality control mechanisms.
- Providing public or fiscal accountability.
- Providing feedback to staff to improve planning, implementation and evaluation skills.
- Fulfilling grant or contract requirements
- Promoting public support and community awareness
- Determining the generalisability of the overall program or components to other populations and settings.
- Contributing to scientific knowledge.
- Creating hypotheses for future study.

The last two purposes are worth elaborating and emphasising, because many social intervention program evaluations focus on the results of that particular program. Such program specific evaluations overlook the wider ramifications of their findings, as they are not immediately relevant. However, community policing, particularly in terms of reducing drug related harm, is a relatively new field and evaluation should contribute

to theory building. It is important to emphasise that irrespective of the success or failure of the program being evaluated, the knowledge gained from conducting an evaluation will increase understanding of community policing overall.

WHAT'S DIFFERENT ABOUT EVALUATING A HARM REDUCTION APPROACH TO DRUG LAW ENFORCEMENT?

Law enforcement has a crucial role to play in reducing drug-related harm in the community. Although the application of a formalised harm reduction approach to law enforcement is only beginning to be embraced by operational police, the approach has a great deal of promise. While police could be said to have a harm reduction role simply by limiting the supply of illicit drugs, more applications of harm reduction principles to law enforcement are being developed and trialed. However, there are some challenges to be overcome in this regard.

As the application of a harm reduction approach to drug law enforcement is a new area, its evaluation is likely to be influenced by a number of contextual factors. Some of these contextual factors are addressed below. Many of them are also addressed as part of a specific issue elsewhere in the manual.

- **Meanings and definitions.** Different terms mean different things to different stakeholders. Perhaps the best examples of this are the terms 'harm reduction' or 'harm minimisation' (see Lenton & Single, 1998). While one cannot assume consistency of meaning at the best of times, the terms may be used very differently in law enforcement and health circles
- **Core underpinning philosophy.** According to Lough (1998), the application of a harm reduction approach to alcohol and tobacco poses no real problems for law enforcement, as there are usually clear legislative guidelines, which distinguish legal practices from those, which are not. However, when it comes to illicit drugs, law enforcers are asked to exercise discretion in the name of 'harm minimisation'. This, he explains, poses difficulties for many police who have been inculcated in a 'black and white' approach to law enforcement. They may also be justifiably wary that exercise of well intentioned discretion could be interpreted as corrupt practice. Lough identifies the organisational constraints on police, societal expectations of police and the attitudes, values and behaviours that police acquire as a result of their job, as the main

factors that impede the adoption of a harm reduction approach to law enforcement. The organisational constraints stem primarily from a military model of law enforcement with its rigid lines of authority and bureaucratic ethos, characterised by strict rules and regulations. All of these can limit the use of discretion. Dixon & Coffin (1999) note that a harm reduction approach assumes a problem solving stance, rather than simply assuming the priority of law enforcement in dealing with quality of life problems caused by drug markets. This requires a radical departure from how law enforcement has traditionally been viewed by the police themselves and the broader community.

- **Goals targets and boundaries.** The evaluators need to establish whether the goals of the project have been clearly stated and whether the different stakeholders share understanding of what these mean. For example does 'reducing harm from drug use' include evaluating the harms attributed to the enforcement of the law itself (eg criminalisation of minor drug offenders) or does it simply mean more effective intelligence to assist drug arrests? In essence, evaluators need to establish what's on the table, and what's not, right at the beginning of any harm reduction policing project.
- **The official position.** In a similar vein, evaluators need to separate official organisational support for a harm reduction initiative from the level of support offered in the field by the police undertaking the initiative. Any dissonance can then be dealt with by educating the responsible personnel and/or modifying the aims of the project in way that takes into consideration legitimate concerns of those responsible for implementation.
- **History.** Every initiative has its own developmental history and in turn is part of a broader historical context. The application of harm reduction based community policing initiatives is no different in this regard (see Sutton and James, 1996).
- **Problems at the interface.** To a large extent, law enforcement and health have different conceptual backgrounds, data sources etc. The two messages, don't use drugs; but if you do, do it as safely as you can, are difficult for some in law enforcement to reconcile (see Dixon and Coffin, 1999; Lough, 1998).

- **Fundamental change.** What Sutton and James (1996) have proposed is more than simply some training for police in harm reduction and attempts at attitude change – they are proposing fundamental structural and procedural change within drug law enforcement, so as to institutionalise harm reduction as part of normal policing practice. Conducting evaluation of demonstration projects in order to identify if such fundamental change is occurring has its own difficulties. These are explored elsewhere in this manual.
- **Illicit drug use as hidden behaviour.** Due to its very illegality, most illicit drug use is ‘hidden’ from outside scrutiny and therefore it is difficult to get locally relevant data on prevalence of use etc. It may be particularly difficult for police to access certain aspects of drug accessing and using behaviour, because of their law enforcement role.

CHAPTER 2: PRACTICAL ISSUES TO BE CONSIDERED IN EVALUATING COMMUNITY BASED DRUG LAW ENFORCEMENT PROJECTS

WHO SHOULD BE INVOLVED AND HOW?

Senior Police and Politicians Police services are hierarchical militaristic systems. Therefore, whilst having support and endorsement of senior police does not guarantee success of the project, it is a necessary pre-requisite.

In an area encompassing the hot topics of drugs and law enforcement, which generate a high degree of interest among the general public, it is likely that the government of the day will want to manage the public perception of initiatives, for which they have responsibility. As such there is likely to be a fair amount of political 'interest' in the evaluation of such projects and a wish on the part of the relevant bureaucrats to closely monitor both the project and its evaluation

Local police. Local police should be involved in any such evaluation. It is within this level of the police hierarchy that the contact occurs with drug users and other relevant stakeholders from the community. Accordingly, local police will often have a good practical understanding of the immediate community impact of drug use. However, as with any organisation, particularly those based on hierarchical models with top down formal pathways of communication, expressions of support (or lack of it) for a position may reflect the 'official' organisational position rather than the person's individual view. Evaluators should not be surprised at this.

Locating evaluation staff within the police service and having them mixing with local police on a day-to-day basis may assist them identify any dissonance between the formal and informal support for the project. This is important, because these differences might impact on the way any project is implemented and its success or otherwise.

Generalist vs specialist police. Many police forces have specialist units (eg. drug squads, major crime divisions). which have a designated role to address drug-related crime. While the specialist units may have a focus (eg. large suppliers) and sphere of operation, which is different to that of generalist police (lower level dealers and users),

in various areas there may be considerable overlap. Evaluators will need to seek advice on the extent to which specialist units are included in the evaluation. At a minimum, it may be prudent to have a contact person within specialist units, who is informed about the project and its evaluation, to assist in dealing with problems that arise and facilitate data access. It may also be prudent to discuss with local police what role or place specialist units have in your area of interest.

Local alcohol and other drug service providers. There will be some specialist alcohol and drug agencies (eg. treatment services, needle exchange providers), which will be located in the geographical catchment area of the study. Often these will have a sound, but again, somewhat compartmentalised understanding of local drug issues. However, there will be other agencies, which while not physically located in the geographical catchment, may provide services to clients who reside within it. The evaluators will need to make decisions about which of all these agencies they will attempt to involve in the evaluation. Agencies will usually have their own agendas, which may have little or nothing to do with the aims of the evaluation project. Evaluators need to be alert to what these might be and how they may impact on the process of the evaluation and its findings.

Local health agency suppliers. There will be other more general health agencies that also provide drug related services. These agencies may need to be involved in the evaluation. Such agencies will include hospitals, health centres, general practitioners, pharmacies, community welfare offices, community corrections agencies, etc. Comments made in the above paragraphs also apply to these service providers.

User groups and representatives. User groups may be located in the catchment area of a study or have services which operate from that geographical area (eg. mobile drug outreach van). Additionally, although such groups may not have services in the area, they may provide a useful link to recruit local drug users, who can provide useful information and comment about drug matters at the local community level. We strongly believe that wherever possible (and this should be on almost every occasion) efforts should be made to involve members of the target group in the evaluation. However, it should be noted, that especially in the case of marginalised and hidden groups, the views expressed by those who consider themselves to be the 'spokespeople' for the group do not necessarily represent the breadth of views about

an issue that are present in the group, nor the views of those who are most marginalised or most hidden. Often agencies such as user groups are aware of these issues and may be able to help to overcome some of these pitfalls – such as by identifying other ways of contacting and including the views of these difficult to reach members of the target group.

Local councils and other arms of government. As the level of government, which is closest to the people, local government will likely be providing services and fielding complaints from ratepayers that are of particular relevance to the evaluation of drug use and law enforcement. Such areas would include youth services, sanitation and waste disposal, provision of safe recreation areas etc..

State and federal government policies and programs may have a direct impact at the local level and consideration should be given to the extent to which they should be involved in the evaluation. Clearly policies of all levels of government are likely to have an impact and at the very least, evaluators should be keeping track of changes in policies and long standing programs, which could account for differences on key indicators at the local level.

Community action groups. There will likely be a number of less formal community groups, which will see themselves as stakeholders in the issue of drugs and law enforcement. Some of these will be groups, which were initially established by government (eg. Safer WA Committees, Local Drug Action Groups or LDAGs) whereas others will have more grass roots origins (eg. parents support groups, traders groups). The evaluators will need to decide which of these groups should be involved in the evaluation and at what level. This will differ on a case-by-case basis. As with the involvement of any stakeholder groups it is important to identify whether the group tends to advocate a particular point of view, which might affect their perception of the initiative being evaluated.

Community service organisations. Community service organisations, which have a remit far beyond drugs and law enforcement, such as Lions, Rotary, etc., may be worth involving in the evaluation, particularly if they are involved in projects that are related to drug use.

WHAT PROBLEMS ARE LIKELY TO COME UP AND HOW CAN THESE BE OVERCOME?

The problem of evaluation apprehension. It is invariably the case that 'evaluation' will invoke apprehension and concern on the part of the individual, group or organisation that is being evaluated. This is a natural human reaction to being scrutinised and should be of no surprise to the evaluator. Often this concern will be quite rational. People will have direct and indirect experiences of previous evaluations that have adversely affected the service being evaluated. Many people find it hard to distinguish appraisal of a service from judgements made about themselves as workers in that service. This is more likely to be the case where people are passionate and committed to the field in which they work.

This apprehension can be reduced, if not overcome, by the evaluators in a number of ways:

- Being clear that you do not have pre-conceived views and other agendas in carrying out the evaluation (If you do, you must be able to put these aside or else perhaps you ought not be doing the evaluation!)
- Ensure that representatives of the agency being evaluated will be represented on an evaluation project 'reference committee' or 'consultation group'. Use these persons formally and others informally to anticipate and help develop a response to any problems which might arise
- Reassure agency staff that you are looking at process as well as outcome and looking at the strengths, as well as the weakness of the program under study, and that the information should be useful for those working in the agency or program
- Ensure that representatives of the agency know that they will get an opportunity to comment on drafts of the evaluation report. Always do this – it is much better to hear and respond to their concerns in private rather than in a public brawl. And remember, it might just be that you got it wrong!
- Try to anticipate events or new information that is likely to raise anxiety and keep agency representatives informed of this.

- Understand that it may take some time for evaluation anxieties to diminish, and that they may never dissipate entirely.
- Having multiple sources of information on a particular evaluation issue (triangulation) can re-assure agency staff that conclusions will not be made on biased or limited evidence.
- Treat those people who provide you with information as “experts”.

Evaluation apprehension may be particularly problematic in militaristic systems such as the police, because of laid down operating procedures, strongly hierarchical management structures and emphasis on individual accountability. Strategies to deal with this include: stressing the confidentiality of information and sources, using multiple informants or data sources, and asking similar question within an interview from a slightly different perspective and monitoring responses for incongruence.

Challenges to the independence of the evaluation team. There is always the possibility that the evaluation may come up with findings and conclusions that the agencies involved, including the funder, strongly disagree with, or find unacceptable. This may result in problems when it comes to editing of the evaluation report and rights to release findings publicly, in the media, or in academic publications.

The independence of the evaluation team lends credibility to the validity of the evaluation report and can be a bonus to the agency or service being evaluated where they can say they were favourably reviewed by an independent evaluator. However, the independence of the evaluation team needs to be established by:

- Having the team responsible to an outside organisation, rather than under the auspices of the project being evaluated, or its parent organisation
- Having a clear written legal contract agreed between the evaluators and the relevant agencies (eg. funders), which articulates such things as confidentiality, release of the report, other publication rights, ownership of intellectual property, and a dispute settlement procedures.

Changing membership of committees and agency representatives. This is a problem for the evaluation as well as the projects themselves, especially where longitudinal methods are employed with multiple data collections over the time span of the project. Where individual responses are important, the evaluator must ensure that there are enough 'subjects' in the early data collection to allow a realistic drop-out rate. Where key informants are being used, it may be appropriate, where possible, to ask potential individuals whether they believe they will be in the same role for the duration of the project. Where group data are employed the samples can be treated as *independent* for the purpose of analysis. Regardless of precautions, it is not always possible to avoid the loss of informants. Where this cannot be avoided a decision needs to be made whether your data source is the individual or the position the individual holds, in which case the next person in that role may be a viable alternative.

When agreement and support is not that. There is a generally accepted maxim in the helping professions that where there is a discrepancy between what people say, and what people do, believe what they do. People may express support for a particular position or initiative for a whole lot of reasons. However, this may not reflect actual support. Evaluators could do well to keep this in mind and rely on multiple sources of information (triangulation) before drawing conclusions or accepting statements at face value.

Being one removed – getting a picture through subjective opinion. As an external evaluator one is often at least one step removed from the data, which is based on the subjective self report and opinions of others (key informants, project officers, and others). In order to deal with this inherent subjectivity, one can use 'triangulation': verifying self reports by at least two other independent sources of data (eg. existing data sources, new survey data, minutes of meetings, views of other observers, etc.). Clearly, it is also important that where data is opinion, it is identified as such in any report of results.

Managing the blooming, buzzing confusion. A longitudinal evaluation such as that conducted in the NCBADLE project is bound to generate a large volume of quantitative and qualitative data. While many will be familiar with strategies for managing large amounts of quantitative data, the many and varied sources of qualitative data need to be organised and analysed if they are to be useful. A number

of texts provide examples of ways to code enter and analyse such data (eg. Barnard, 1995; Strauss & Corbin). Data sources might include minutes of meetings, notes or transcripts of interviews with key informants, logs or journals, correspondence, media reports, organisational annual reports and policy documents, etc.

One approach, which we found useful was to generate a list of themes identified from a variety of sources including the explicit goals of the project, a review of the literature, and preliminary analysis of the data. This is illustrated in Figure 1. We used the following themes: attitudes to harm reduction, attitudes to drug users, perceptions of harm. All sources of data were then coded for these themes and decisions made about adding new themes as they emerged. Each of the sources was also coded as to it's type (interview notes, minutes of meeting, etc) and importantly, time in weeks, which allowed for a mapping of the trends as they emerged and changed over time. By placing a summary of each data document, type, date, and coded themes in a large table in a word processor (see Figure 1 for an example) we were able to get a map of all the data, which greatly facilitated data location, organisation and analysis. By sorting on each column (Type, date, coded themes) we were able to quickly pull together all data sources, which related to a particular time, or theme and then consider these together as part of the analysis. It was also possible for the summaries of results of analysis of quantitative evaluation components (eg. reports on surveys and questionnaire studies to be included as data sources so that consideration of these could be done where they related to particular themes or phases of the project.

Informant	Time	Type of document	Perception of harms	Attitude to users	Illicit drugs	Attitude to harm reduction	Other
01207	Jul 97	Statistical			Street heroin at 80%	1	
00313	Jul 97	Interview	Death is the biggest harm	2			
00312	Jul 97	Survey			60% inject 20% oral 20% smoke	6	
01234	Aug 98	Focus group			Heroin is the worst drug	1	Promoting HR is of benefit to the police 5
01231	Aug 98	Telephone conversation		Users are a waste of time	1		Wary of youth agencies 1
00331	Aug 98	Interview	Increasing injecting	4			
02100	Aug 98	Interview			Heroin causing most harm	1	
01201	Aug 98	Interview			Cannabis causing most harm	2	
01233	Sep 98	Survey	Injecting causes 75% of harms	3			

Figure 1. Example of a Data Description Matrix

Problem of data collection at a distance. Two of the NCBADLE evaluation sites were in different cities to where the evaluation teams were based. This posed a number of problems, one of which was that the utility of site visits may have been greatly affected by unexpected events, such as late postponement of meetings, poor

attendance and variable operational priorities. On a day to day basis there was a less intimate connection with project implementation and consequently more reliance on secondary sources of data

Acceptance of the evaluation team. At times it is not always possible or desirable to recruit evaluation staff from within the population you are investigating. The use of external evaluation sometimes causes resentment within the target population. People being evaluated often assume that the problems, issues and solutions confronting them are unique to their setting. They have difficulty accepting that an outside person or agency could realistically understand the situation. This can result in resentment, refusal to be involved in the evaluation and confounded data. The same strategies for reducing apprehension are applicable here.

Getting access to police and agency personnel. As police personnel work shifts and both police and other service agencies have clear operational priorities, evaluation officers need to be flexible about their working hours in order to get access to personnel for interviews, focus groups, meetings, etc.

Relationship with project officers and evaluation team. The evaluation team need to manage the boundary between themselves and the project officers. Whilst the project can be facilitated by a close working relationship, the evaluation team needs to receive the opinions of the project officer as just one source of data. Similarly, the evaluation report needs to be commented on by project officers and other stakeholders prior to being finalised, however, the independence of the report needs to be guaranteed.

Everything takes longer than you expect. This is especially the case when conducting an evaluation, which depends to a large extent on others, who have different priorities and responsibilities. This calls for realistic appraisal of time lines, and using contacts to reinforce priorities and keeping to deadlines.

CHAPTER 3: EVALUATION PARADIGMS AND APPROACHES

INTRODUCTION

Before undertaking an evaluation of any sort of social intervention it is important to consider whether the methods you intend to use are appropriate in terms of the intervention setting, previous research that has been undertaken on the issue and the information you are seeking. As part of this you will find it is useful to explicitly identify and understand the philosophical assumptions and inquiry paradigms, or models, that underpin the range of methodological approaches available.

Guba and Lincoln (1981) identify the scientific and the naturalistic paradigms as the most widely used in evaluation. Each paradigm offers a different choice of conceptual model and each tends to be associated with different evaluation methods. Similarly, each has different degrees of association with certain tools and techniques used to obtain the evaluation data. Thus, an understanding of these two paradigms is important to you if you are interested in evaluating the social consequences of a particular policing approach. Your choice of conceptual framework will influence the evaluation methodology and structure the sort of evaluation you can undertake

THE RELATIVE MERITS OF THE SCIENTIFIC AND NATURALISTIC PARADIGMS

The scientific and the naturalistic paradigms respectively derive from positivist and phenomenological thought. The first, positivism, seeks facts and causes of social phenomena, independent of the subjective states of individuals. The second, phenomenological, is committed to understanding social phenomena from a participant's perspective.

Historically, these paradigms have emerged from sympathetic disciplines. That is, the science disciplines, particularly the 'hard' sciences, have sought facts and causes to explain the natural world around them, while the sociological and philosophical disciplines have sought understanding of human thought and practices. Thus traditionally, each has been used to find answers to different kinds of problems. You, as an evaluator of community drug policing initiatives need to make a choice as to the paradigm or combination of paradigms that best suits the sort of questions you are

trying to answer, as the assumptions inherent in each model will influence what you can find out.

The first assumption identified by Guba and Lincoln (1981) is an assumption about what constitutes the reality of events. The scientific paradigm sees events as interconnected sequential processes that can be studied by defining variables and isolating segments for intense investigation. This is done primarily for purposes of prediction and control (Guba and Lincoln, 1981). Naturalistic reality however, does not view events as singular or fragmented, but assumes multiple meanings which are interrelated. This is an ecological view, where ecology is defined as being:

Based on the premise that living organisms must be studied in relation to the other organisms with which they coexist and in relation to the non organic setting which they occupy... Membership in some sort of shared community is part of the ecological framework (Lehmann, 1975, page 487).

The second assumption is that related to the inquirer/subject relationship. The scientific paradigm assumes that there is no (or should be no) relationship between inquirer and subject. Proponents of this paradigm see themselves as 'objective' rather than 'subjective' and encourage the inquirer to maintain significant distance from the subject. In contrast, the naturalistic paradigm demands inquirer interaction with the subject, recognising that all phenomena are characterised by interactivity and thus no purely 'objective' stance ever exists (Guba and Lincoln, 1981). Advocates of this approach consider that recognising and taking into account interactions that occur is more honest and allows the inquirer to get closer to the truth.

Assumptions about the nature of meaning are also different for each of these two paradigms. The aim of scientific inquiry is to produce generalisations; universal truths applicable to different contexts. Naturalistic inquiry on the other hand eschews generalisations in favour of 'rich descriptions' and 'working hypotheses'. The unique qualities of different contexts are emphasised rather than the commonalities (Guba and Lincoln, 1981).

THE RELATIONSHIP BETWEEN PARADIGM ASSUMPTIONS AND EVALUATION METHODS

The corollary of considering the assumptions underlying your evaluation paradigm is the need to match your global conceptualisation with appropriate choices of evaluation methodology. This is important despite the fact that the assumptions that underpin the paradigms are barely visible in the day to day conduct of evaluators. These paradigm assumptions are however reflected in certain derivative elements. Moreover, each of these paradigms has acquired certain approaches that cannot be deduced from the basic assumptions themselves or justified by reference to them. These approaches have nevertheless become strongly associated with the paradigms (Guba and Lincoln, 1981).

The evaluation method associated with the scientific paradigm is experimental and generally seeks quantitative data, while the method associated with the naturalistic paradigm is simply named naturalistic and generally seeks qualitative data.

The experimental method, based on positivist assumptions is a very reputable and popular research and evaluation approach. For example, Nutbeam declares that:

basic experimental design, and particularly a randomised control design, are well established as ideal methods for evaluation (Nutbeam, 1990, page 85)

Limitations however are recognised.

The very nature of community based interventions denies the experimental control of many variables. Communities are complex and changing systems...(where)....the freedom to select areas randomly for intervention is also limited.....(and)...the causal chain in a community system is longer and harder to trace than a clinical research study on volunteers- the classic application of a randomised control study design. The most widely adopted solution to this problem has been the development of the quasi-experimental study design. (Nutbeam, 1990, pages 85-86)

The quasi-experimental design does not involve random assignment into either experimental or control conditions. This makes it a more suitable method for large scale police harm reduction trials, where location will usually be selected prior to

involvement of an evaluator, because of the particular combination of geographic and social circumstances that foster problematic drug use. Naturalistic methods differ in that the researcher does not manipulate the research setting, but rather seeks to understand naturally occurring phenomena in their naturally occurring states. Evaluators, coming from a naturalistic perspective, focus on capturing processes and exploring important variations in experience and outcomes (Patton, 1989). The evaluation context and the perspective of the critical reference group are important in qualitative evaluation.

Although scientific and naturalistic paradigms are considered separate entities by some evaluators they are not mutually exclusive. A good quality, large-scale evaluation will include both naturalistic and scientific elements. Each paradigm is associated with one of the four generally recognised types of evaluations; formative, process, impact and outcome evaluation. Formative evaluation is carried out in the initial stages of an intervention to identify how it can be best implemented. Process evaluation measures the activities of the program, program quality and who it is reaching. Impact evaluation measures the immediate effect of the program. Outcome evaluation measures the long-term effect of the program.

Since a general characteristic of the scientific paradigm is its purpose of verification it is most often associated with outcome and impact evaluation, while naturalistic enquiry, with its purpose of discovery, is more often associated with formative and process evaluation. You should also consider the maturity and breadth of your research project, when selecting the type of evaluation likely to provide the most meaningful information. In a new area of research, such as the effects of policing strategies to minimise drug related, harm, little is known about the salient variables. Accordingly, a naturalistic approach is more likely to be useful in early research and evaluation programs. The approach is more open-ended and consequently it is useful in identifying critical variables. As an area of research matures, more is known about the salient variables and these can then be investigated in more depth. This is when the scientific approach offers considerable advantages. Particular variables of known importance can be isolated and then manipulated to look at the effects. Comparisons can also be made against standard practice. A new approach, such as police issuing a formal caution for certain types of drug possession can be compared with existing

practice such as arrest. Finally, because scientific evaluation offers a more replicable approach, a comparison of known efficacious interventions in different settings can be undertaken. For example, does cautioning work as well in the country as it was found to work in the city. The scope of your research project should similarly be considered when selecting your evaluation approach. At one extreme, a small evaluation of a pilot intervention involving a large number of critical variables would not be well suited to the scientific approach, because resources will limit coverage and make the evaluation incomplete. A naturalistic approach would probably be more useful, because it can be more easily done with minimal resources and is likely to generate some level of global understanding. In another scenario, where a well resourced evaluation is being conducted to answer a particular question, such as 'does police attendance at incidents of drug overdose decrease reporting of such incidents on the 000 number?', the scientific approach is likely to provide more meaningful information.

THE RELATIONSHIP BETWEEN ASSUMPTIONS AND EVALUATION TECHNIQUES AND TOOLS

Two different sets of techniques and tools are associated with each paradigm. Those associated with the naturalistic paradigm tend to qualitative and subjective. They include group discussions, one to one discussions or interviews, observation, using personal documents and written questions and answers (Wadsworth, 1991; Burgess, 1984). Techniques and tools associated with the scientific paradigm tend to be quantitative and objective. They include the specification of hypotheses, random assignment of participants, use of pre-intervention/ post-intervention measures, use of non intervention control groups, use of existing 'official' data, use of standardised surveys and sampling and statistical analysis of data (Windsor et al, 1994).

CONCLUSION

In order to undertake an effective and appropriate evaluation of any social intervention, the evaluator must first take a global view of the task to be undertaken. The evaluation techniques and tools discussed in subsequent chapters, even when correctly implemented, may be useless if proper consideration of the global environment within which the evaluation is framed has not been considered. In this manual we try not to be overly prescriptive in regard particular evaluation approaches, because each situation is different and we want you to grasp what is involved in

evaluation design and be able to tailor an evaluation for the particular issues you are attempting to understand. However, you will need to recognise that some evaluations are best undertaken using a scientific methodology, others a naturalistic methodology, and others an eclectic approach. It is important for you to understand the implications of the two main evaluation paradigms before selecting the type of evaluation you will conduct.

CHAPTER 4: TYPES OF EVALUATION

There are four common types of evaluation that can be used to assess social interventions: formative, process, impact and outcome. They are generic, in that they can apply in different settings such as schools and hospitals; they can apply to different content areas such as law enforcement and sexuality and they can apply to different target groups such as individual participants and organisations.

FORMATIVE EVALUATION

Formative evaluation occurs in the initial stages of an intervention to test the efficacy of program strategies and data collection instruments (Windsor et al, 1994). It is useful to think of formative evaluation as the pilot stage of the intervention, when understandings about the setting and targets for change are generated through detailed observation and interaction with intervention participants and key informants. Formative evaluation helps to identify how interventions can be best implemented; whether objectives are realistic and acceptable to the community and key stakeholders and realistic in the context of the setting. In addition, a period of formative evaluation also helps the evaluator to identify appropriate measures of change to be used in the process, impact and outcome evaluation.

Formative evaluation can be a time consuming process. However, it is an essential aspect of any new or modified social intervention as it ensures that the intervention has the optimal potential for success. It provides insight about the application of an intervention with the group or setting of interest and helps to ensure that any problems are dealt with prior to the application of the intervention on a grander scale and prior to undertaking any impact or outcome evaluation. Formative evaluation ensures that a program is well designed for its purpose and that the money applied to the full intervention is well spent.

Formative evaluation is similar to needs assessment, in that it involves intense interactions with the target group and uses qualitative measures of evaluation including, for example, interviews with key informants, focus groups with key participants, open-ended questions on surveys. However, it differs from needs

assessment in that it focuses on refining intervention and evaluation strategies, rather than defining the problem.

Although formative evaluation requires a period of time prior to the full intervention, it is a very valuable exercise as it helps to ensure that the eventual intervention adopts the most appropriate strategies to achieve the desired aims and adopts the most appropriate measures to evaluate what changes.

Formative evaluation

- Aims to ensure that the intervention is being implemented in the best way.

Provides an opportunity to revise intervention methodology

- Helps the evaluator to identify appropriate measures of change to be used in the process, impact and outcome evaluation of an intervention
- Is used for new interventions, when pre-existing interventions are used in new settings or when modifications are made to existing interventions
- Uses qualitative assessment techniques to gain insight into how the intervention can work best. It is not appropriate to change qualitative data into quantitative data by recording frequency or testing for significance

PROCESS EVALUATION

Process evaluation should be undertaken throughout the duration of an intervention to ensure that a program is being conducted as planned. Measures need to be put in place to ensure that each component of the intervention is actually being implemented as intended. One of the major flaws with the evaluation of health and social interventions in the past has been the absence of process evaluation (Basch et al., 1985). In these evaluations, impact measures suggested that interventions didn't work or created little change. However, when further information was obtained about why these interventions didn't work the main reasons was because they were not implemented as

intended or not implemented at all (Basch et al., 1985). Thus the impact evaluation was not measuring any changes related to the intervention. In naturalistic settings such as schools, hospitals or whole communities, the main targets for health and social interventions, it is essential that process evaluation be undertaken to ensure that programs are being implemented in the desired manner.

There is no one ideal method of gathering process information. Often several methods need to be adopted to act as a check on what is actually occurring. Triangulation is the term used when applying several different types of process measures to assess the level of implementation of an intervention. For example, in a harm minimisation training package for police, three methods of assessing whether the training was implemented as intended could involve: 1) gathering self-report data from trainers, 2) interviewing students about the training they experienced in class and 3) evaluator observation of the training.

Have, Degeling and Hall (1992) suggest that when measuring the implementation of an intervention, four main questions need to be addressed:

- Is the program reaching the target group?
- Are participants satisfied with the program?
- Are all the activities of the program being implemented?
- Are all the materials and components of the program of good quality?

Process evaluation

- Occurs throughout the duration of an intervention
- Measures adherence to intervention strategies
- Requires a series of evaluation measures to ensure that information is reliable
- Includes the measurement of frequency data concerning the amount of program activity undertaken, the number of participants involved etc.

PARTICIPANT EVALUATION

Participant evaluation is often considered a sub-set of process evaluation, but it is worth mentioning in its own right in this manual, because of the community focus of police prevention programs. Evaluation is generally done to satisfy the needs of researchers or bureaucrats and is accordingly framed in terms that have meaning for these groups. This can mean that it is of little relevance to members of the community participating in the project. Accordingly, participant evaluation involves community members shaping the evaluation of projects so that the information produced is relevant to their needs. In summary, some key factors in participant evaluation are:

- Community members are involved in the evaluation process
- Evaluation outcomes are structured to maximise community relevance
- Community knowledge informs the interpretation and use of evaluation results

IMPACT EVALUATION

Impact evaluation should occur after formative evaluation and in conjunction with process evaluation. Impact evaluation assesses the performance of the complete intervention by looking at its immediate effects. Usually impact evaluation correlates with the specific objectives of an intervention and because interventions often have several objectives, impact evaluation is generally multifactorial.

When measuring the impact or change that has occurred during an intervention, it is important to consider the type of evaluation design, the type of data collection methods and the specific variables that will be assessed. Ideally, impact evaluation will be carried out within a quasi-experimental design; will involve gathering data from intervention and control/comparison groups; will involve the measurement of several domains (e. g. knowledge, attitude, behaviour) and will have several data collection points, including the collection of baseline data, prior to implementation of the intervention.

Impact evaluation

- Assesses the immediate effects of the complete intervention

- Correlates to intervention objectives and strategies and as a result is usually multifactorial
- Usually uses quantitative data that can be analysed for statistically significant change
- Needs to be carried out during program implementation or immediately on completion

OUTCOME EVALUATION

Outcome evaluation is a second level of assessing the performance of an intervention and is usually related to the long-term goal of the intervention. Generally, outcome evaluation assesses the social problem or social behaviour related to the intervention, such as do needle exchange programs improve the health of injecting drug users. However, choice of outcomes can vary between similar interventions and between stakeholders (Hawe, Degeling & Hall, 1992). The short term funding nature of most community policing interventions and the long term nature of outcome evaluation means that assessment of outcomes may need to be done well after the completion of a project.

Outcome Evaluation

- Assesses the long term effects of the complete intervention
- Correlates to intervention goals and is usually related to a health behaviour
- Usually uses quantitative data that can be analysed for statistically significant change
- May need to be done well after program completion

SELECTING THE APPROPRIATE TYPE OF EVALUATION

The type of evaluation used in the assessment of an intervention depends on several factors including the:

- Intention of the intervention
- Scope of the intervention

- Stage of the intervention
- Needs of the stakeholders
- Funds available for evaluation
- Time available for the evaluation
- Evaluation expertise available
- Planned use made of evaluation information

In addition, there may be other determinants that are specific to the intervention or organisation that enhance or compromise the quality of the evaluation.

OVERVIEW

Depending on the nature and stage of the intervention being evaluated, there will be more or less emphasis on the different types of evaluation. In the case of a large scale demonstration project, comprehensive evaluation, involving formative, process, impact and outcome components would be ideal so that all aspects of the intervention are well understood. In contrast, a smaller, less well resourced community based intervention using a well tried intervention with proven efficacy may be adequately served by process evaluation to ensure comparable implementation.

Pirie (1990) suggests that for an evaluation to be useful the results ultimately need to be taken into consideration in the planning and implementation of subsequent interventions. This process of basing subsequent interventions on the lessons learnt through the evaluation or previous interventions ideally leads to the generation of better planned and more effective interventions.

CHAPTER 5: TECHNIQUES FOR DATA COLLECTION AND ANALYSIS

INTRODUCTION

Drug-related problems result from complex interactions between: aspects of the substance (strength, effects, route of administration, price, etc); the user (drug experience, health status, age, etc.); and the socio-cultural environment (peer norms, legal structure, media portrayal etc.). As a result, the study of drug users and their interaction with law enforcement and the general community, necessitates multiple methods of evaluation. This is made even more problematic as both the illicit drug users and law enforcement cultures are resistant to outside scrutiny. In the case of complex social interventions, such as harm reduction policing, there are no pre-determined right or wrong ways to conduct an evaluation. Accordingly, evaluators need to be familiar with a broad variety of evaluation methods to be able to select the best for the given circumstance. If for example you are evaluating harmful patterns of alcohol consumption in a country town, the skill you will need to have as an evaluator is to decide which method, or combination of methods gives you the best possible information, in the circumstances. Should you conduct a survey of the population? Should you get a representative sample to keep a diary of their consumption? Should you do a screening test of liver function? Should you talk to key informants such as the police, publicans, doctors and rehabilitation workers? Should you conduct a series of focus groups with range of people from the town? Should you get alcohol sales data from the relevant government agency?

Some of the primary approaches for researching drug use include: surveys, ethnography, psychosocial experiments, studies of social and health indicators, historical studies of policy documents, traditional epidemiology informed by social science theories, as well as biological and behavioural research (Trotter & Medina-Mora, 1997; Room, 1991)

MEASUREMENT AND EVALUATION

In Chapter 1 we presented Green's conceptualisation of evaluation as involving comparison of an object of interest with a standard of reference (Green,1990). This comparison must be based on data, which in turn must be collected by a process of

measurement. Measurement involves determining the quantity or quality of an object of interest and in this respect it is the basic tool of evaluation. In this chapter we will present a range of measurement instruments and techniques that can be used to collect data.

SAMPLING

In testing for a relationship between a set of factors it is seldom possible to examine the entire population. A sample is therefore used and measures are made to demonstrate that the relationship in the sample is also true of the population. In simple terms, findings from a sample of high school students in a community can be generalised to the high school student population of the community from which the sample was drawn. However, if the intention is to generalise the findings to all teenagers in the community, then the sample must be considered 'biased', or unrepresentative of the population to which you want to generalise, because not all teenagers go to high school.

Where possible, a sample should be selected in a random manner. There are several types of sampling techniques with random components, such as simple random, clustered, and stratified. Random selection minimises the likelihood of systematic selection bias, but cannot ensure against it. Windsor et al (1994) suggest taking additional precautions such as comparing key variables in the sample against the population which it purports to represent.

However, a problem with undertaking studies of illicit drug users is that because the behaviour is illegal and hidden, it is not possible to get a good picture of the population as a whole. While surveys of drug use among the general population such as the National Drug Strategy Household Survey 1998 (Australian Institute of Health and Welfare, 1999) can be of some assistance, they are considered to represent under-reporting of illegal drug use, and have a sample size, which is generally too small to gauge drug use patterns among local communities.

As such, it is not always possible to do strict random sampling in the field and there are a number of alternatives, which can be appropriate under certain circumstances. These include quota sampling, convenience sampling, and snowball sampling. Although these methods limit your ability to generalise beyond the sample, when

based on good ethnographic and qualitative data they are often highly credible (Bernard, 1995).

In quota, 'targeted' or 'stratified' sampling, decisions are made on the basis of prior knowledge as to whether there are subpopulations of particular interest in the target group. For example in studies of drug users, age, gender, current treatment involvement and years of drug use might be relevant. When sampling the researchers recruit potential respondents within each of these groups. Thus, 40% of the sample might have to be women, 50% to have been injecting for 10 years or more, and so on (eg. Loxley, Carruthers & Bevan, 1995) . In a recent study on the effect of a cannabis conviction on peoples' lives the researchers placed quotas on the time since arrest as they knew that this would have a large impact on the possible exposure to social consequences (Lenton, Bennett & Heale, 1999).

Convenience or 'haphazard' (Barnard, 1995) sampling techniques are often used in studies of hidden behaviours such as illicit drug use. This involves accessing whoever you can of a target group. While this has obvious hazards in terms of the non-representativeness of the resulting sample, the problems can be reduced by:

- being clear about the target group and making decisions on how best to access them;
- considering the inherent biases involved in different recruitment strategies and choosing a menu of approaches that will best recruit from the targeted population (e.g. for illicit drug users, advertise in street press, recruit from community pharmacies, drug treatment agencies, needle exchanges, police lock-ups and drug outreach services and drug user groups)
- carefully documenting the methods employed, and the limitations of them, so that (a) consumers of your evaluation can be very clear about the limitations and potential biases of the resulting sample and take this into consideration when reading the results (b) your methods can be replicated by others;
- Run your proposed recruitment and sampling strategy past experts in research with this target group, and of course drug users themselves who are usually the most informed experts about their peers.

In similar way to convenience sampling, geographical sampling techniques can be used where there are various settings and locations which are ideal for observing and collecting data about drug subcultures (Trotter & Medina-Mora, 1997). For example in their study of the street drug dealing and using scene around Smith Street in Melbourne, Fitzgerald, Broad and Dare (1999) used geographical sampling and analysis to look at the impact of the local drug using scene on drug-related harms in the area.

Another strategy useful for accessing members of hard to access populations such as illicit drug users is 'snowball' sampling. In this method you ask members of the target group with whom you have contact to name others who would be likely candidates for the research you are conducting. The technique is particularly useful when doing community studies and has been used as part of outreach interventions with drug injectors (e.g. Bernard, 1995). One disadvantage of the technique is that in a medium to large population, those individuals who are well known are more likely to be nominated than those who are less well known.

It is often useful to combine convenience and snowball sampling techniques to maximise the efficiency of recruitment (e.g. Lenton, Boys & Norcross, 1997) and optimise the representativeness of the sample.

SAMPLE SIZE

This tends to be one of the last things considered in evaluation and when it is considered, decisions tend to be driven by what can be afforded in terms of the project's funding and whether the sample size that can be afforded also seems reasonable. However, there are clear problems to such an approach. If your sample is too small, your results will have such a large range of error that they cannot necessarily be considered representative of the population from which they were drawn. Even if you do not have the resources or time to get a large sample, it is important to understand and acknowledge the difference between what you can do and what is ideal. This will allow more informed interpretation of your results.

In any large study prior to starting there should be an exact calculation of the required sample size. Statisticians can work this out on the basis of such factors as the 'confidence interval', 'error range', 'population size' and 'expected results'. However, in small evaluations, such expertise may be difficult to access. Accordingly, we have

presented a 'ready reckoner' for the calculation of sample size in Figure 2. This was developed by the South Australian Community Health Research Unit (1991) and is based on assumptions generally applicable to social research. These are: an expected response of 60% (expected results), coupled with a 95% level of confidence (95% confidence interval) that the response from the whole population would fall within a 5% range (5% error range), either side of the result actually obtained from the sample. Unfortunately if you want to be confident about generalising your results to a sizeable population such as that in a country town you will need a sample size of several hundred. If you have to use a sample of less than ideal size, gather confirmatory data using different techniques. This is a form of triangulation. For example, if you want to find out if people are opposed to a needle disposal bin being placed in the toilet block of a neighbourhood park, but only have the resources to survey 100 residents in a surrounding population of 9,000, you could confirm the survey results by monitoring news items on the issue in the community newspaper; check if any objections had been lodged with council; attend a meeting of the council and/or local progress association. If several independent data gathering techniques provide converging evidence you can be more confident of your findings.

Source: South Australian Community Health Research Unit ,1991, page 162.

Figure 2 Sample Size Chart

OBTRUSIVE OR OVERT MEASURES

Self Completion Questionnaire. This is an instrument that the respondent can complete by reading and answering without assistance. It is the most convenient and frequently used method of data collection and is particularly appropriate when the phenomenon being studied is amenable to self observation and relevant data can be elicited by simple straightforward questions. It is also very useful for collecting data on behaviours that may be considered 'private' or where there are strong confidentiality concerns such as for sexual practices, drug use and law breaking. There is a trend for large questionnaire based interview studies, such as the National Drug Household Surveys in Australia, to include self-completion sections for such private behaviours. This minimises social desirability response styles where the respondent tends to tell the interviewer what they believe will not be disapproved of. One self-completion questionnaire study of drug injectors (Lenton, Kerry. Loxley et al., in press) found higher self-reported rates of needle sharing, a socially disapproved of behaviour, than an interview study of drug users (Bevan, Loxley & Carruthers, 1996). While there were a number of differences between these samples it has been suggested that the higher rate of needle sharing in the former study may be due to the greater anonymity of the self-completion methodology. However, it has a number of disadvantages associated with lack of supervision.

Questionnaires can seek to gather data on several different aspects of an issue and accordingly use different types of questions. The National Drug Household Survey (Commonwealth Department of Health and Family Services, 1996) asked questions about drug consumption patterns, knowledge and attitudes about drugs and their effects, exposure to and participation in alcohol related crime, and awareness of various strategies for dealing with alcohol and other drugs. Each series of questions was framed quite differently and it is important to use a style of question that suit your purpose. True/false questions are useful for simple knowledge questionnaires, such as:

Alcohol is a drug.....True False.

Multiple choice, with discrete categories, is useful for quantity or frequency data, such as;

I go jogging for at least 15 minutes *Everyday*
3 or 4 Times a Week
Once or twice a Week
3 or 4 Times a Month
Less Often or Never

Likert Scaling is a useful way of investigating attitudes, as it forces respondents to select from fixed categories which represent varying degrees of agreement with a stimulus statement, such as;

Cannabis should be decriminalised *Strongly Agree*
Agree
Undecided
Disagree
Strongly Disagree

Self Completion Diaries and Logs. Verbrugge in Windsor et al (1994) considered that diaries and logs produced higher frequencies for most phenomena than other methods. Diaries and logs appear to be particularly appropriate for reporting low salience phenomena, because reliance on memory is minimised. In addition, diaries could be used as a source of data for qualitative analysis. There are however problems with this method. A considerable amount of checking is required to collect quality data. Because of the recording process respondents become more aware of their own behaviour, and are likely to become over concerned about health risk symptoms. Conversely, if respondents are required to keep diaries and logs for an extended period of time, they become disinterested and frequency of events recorded can decrease by 5 to 25 percent.

Face to Face Interviewing. Windsor et al (1994) consider that in certain circumstances there is no substitute for having an interviewer conduct a survey. The face to face interview is preferable when:

- The content area is not well defined
- The questions are complex and may require clarification

- The respondents have literacy problems
- Respondents may need encouragement to participate

The primary strength of a face to face interview lies in the use of skilled interviewers who are familiar with the respondents and the issues being researched. For example, if information is required from a traditionally oriented Aboriginal community, trained local interviewers will probably be the best way to access the community.

The face to face interview is however susceptible to a number of biases. Respondents are likely to anticipate what the interviewer expects of them and act accordingly. The probing of particular matters may in itself change the way the respondents think about them. The interviewer may not obtain accurate information on highly personal matters such as sexual behaviour. You also need to bear in mind that the face to face interview is an expensive method of data collection and should be used sparingly.

Telephone Interviewing. Telephone interviewing is an attractive alternative to face to face interviewing because information is cheaper to collect and there can be greater control over the process. However it is susceptible to additional biases, particularly population sampling. Some deprived populations and remote populations do not have telephones where they live. An increasing number of people have chosen to have unlisted telephone numbers, and even those with listed numbers may have moved. Telephone interview techniques are slightly different from the face to face situation and while the pace can be faster the length of the interview has to be shorter. Market research companies are ideally set up to conduct telephone surveys. Although the cost may seem prohibitive, the efficiencies they provide often make them a better option than a 'do it yourself' approach.

Direct Observation. Windsor et al (1994) consider that observational methods are most useful for collecting behavioural and capability data. Data can be obtained directly by observers, or by video or audio tape recorders. Observation may be concerned with simply recording the frequency of certain phenomena or at a more complex level with the relationship between events. It is an expensive approach which gathers a limited amount of data over a given period of time but has the capacity to provide a more contextualised picture of the phenomena being investigated (eg.

Carruthers, 1997). The introduction of observation may in itself influence behaviour and this should be factored in when considering use of this method.

Participant Observation. This is a variation on direct observation and is considered to be a combination of observation, participation and discussion (South Australian Community Health Research Unit, 1991). The approach recognises that there is no such thing as purely objective observation in social research. Consequently it is important for investigators to be aware of the perspective they bring to the evaluation and the effect their presence has on the data they gather. An action research (or advocacy) model is usefully served by Participant Observation, because it involves close interaction with the community and findings can be quickly feedback to community decision making bodies to inform project choices and direction. The particular value of Participant Observation is that it involves a considerable amount of interaction between researcher and the community. However care should be taken to maintain a critical research perspective and not accept community perspectives at face value.

Group Techniques. This section examines focus group, and nominal group techniques. Group discussions can be an economical method of getting a lot of information and may be a useful way of getting some insight into the interpersonal dynamics of a community, but the limitations of the approach need to be understood. Participants, particularly the police are likely to have pre-existing relationships and these will affect the group process. Similarly participants may want to project a certain image of themselves and this will influence how they participate in the discussion.

The focus group involves bringing together up to ten people to talk about an issue with the researcher monitoring and guiding the discussion. It can be a quick way of getting an overview of community concerns and a particular advantage of the technique is that it enables the participants to interact, which in turn enables ideas to be elaborated, refined and clarified. The main disadvantage of focus groups is that the information obtained, is only what participants are prepared to provide. This may also be distorted by the group process, where for example, only some dominant participants talk or certain points are exaggerated in order to create a particular impression.

The nominal group technique is a process by which a group of people can clarify their priorities on a selected issue. It involves participants individually recording their ideas on cards. These are then sorted to create topic or issue categories. Participants are then allowed to vote on the topics or issues they consider most important which are then prioritised accordingly. The main advantage of this process is that it allows equal participation by all group members not just the most vocal, however it allows for minimal elaboration of the issues raised and qualitative assessment as to the importance of issues cannot be undertaken.

THE LIMITATIONS OF OBTRUSIVE OR OVERT METHODS

The great number of 'obtrusive' or overt data collection methods can tap a broad variety of data, but each is subject to some form of bias, particularly the subjective bias of the respondents. Accordingly, investigators should use a number of methods, which are subject to different sources of error. In particular, a component of empirical data should be incorporated into an evaluation where possible. For example a program designed to reduce alcohol related injury may conduct a survey to see if people in the community perceive that it has achieved its aims, but it may also measure alcohol consumption to see if this has gone down and examine hospital emergency admission data to see if alcohol related injuries have been reduced. If the data from these sources were consistent, then the evaluation conclusions would be more robust and more difficult to refute.

UNOBTRUSIVE OR INDIRECT MEASURES

Abstraction of Existing Records. In many cases, data which is relevant to a particular community policing program is already being collected by various agencies. Having access to this routinely collected information can save a lot of time, and if collected consistently may provide useful trend information. However a disadvantage is that it may not necessarily align well with the project objectives, and may under-record statistics on data the agencies were not established to measure as core business. Additionally, it may take time to get approval to access relevant data bases. The success of analysis of existing records depends to a large extent on the co-operation of the agencies who hold the data. Thus in terms of future projects, it is important not to put individuals and agencies off-side. Access to data requires both high level political support to make the data available and cooperation and trust of on-the-ground staff in

the agency, who can make it happen and smooth out any problems. In order to do this agencies need to know that confidentiality is assured. Agencies will also want to know there is some benefit, or at least minimal cost to them. In addition, you will want to be sure that quality control is high in that the data are as comprehensive and accurate as possible. These issues are discussed in more detail elsewhere (Jean-Francois, 1997). In practice they are often addressed directly and explicitly early in negotiations between agency and evaluator in a 'Memorandum of Understanding' (MOU). Any inherent bias in the data should also be taken into consideration when interpreting meaning.

Relevant existing sources of data could include:

- Consumption data, (for example illicit drug use data from the national household survey data broken down by collector district, alcohol consumption estimates based on sales of alcohol by geographic region)
- Needle and syringe distribution data (from relevant state drug agency)
- Police data: arrests, charges, infringement notices, cautions, statistics on price and purity of seized drugs, number of clandestine laboratories raided, etc.
- Court , Community Corrections and Imprisonment Data
- Drug treatment agency data (enquires, admissions, etc.) which could be collected from individual agencies or more efficiently from peak bodies and statewide data sets (such as that maintained by Next Step Alcohol and Drug Services and the WA Network of Alcohol and Other Drug Agencies in WA)
- Ambulance attendance data (e.g. of overdoses)
- Hospital admission and Accident and Emergency Occasions of Service data
- Information from local councils (for example data on discarded needles, etc).

Collect new data. Wadsworth (1991) suggests that in some evaluations it may be useful to collect new data, which is better aligned with the purpose of the project. This will be very useful in its own right, but there can be additional benefit if the local

community is involved in gathering this data. It is likely that the process itself will create greater community understanding and commitment, but it is arguable whether such a method can still be classified as unobtrusive or indirect.

Media reports. A source of new data, which might be useful are media reports (newspapers, radio and television) about drug and crime issues. Monitoring articles and letters to the editor in the local paper can give a picture of trends in the immediate community and a picture of community perceptions of the issues involved. Monitoring broader media (eg. Statewide), can alert the evaluator to more macro factors which can have local impact). Thus, media coverage of related issues could be both a dependent and an independent variable. Media monitoring companies exist which, for a fee, will provide copies of articles and transcripts or summaries of electronic media pieces on various topics and within specified media outlets. A cheaper way of doing it is to keep a scrap-book of media articles, noting the source, date and page numbers. This can provide a useful tool as part of an evaluation package.

ALTERNATIVE CATEGORISATIONS OF MEASUREMENT METHODS

As a way of providing some logical order to the considerable variety of measurement methods presented in this Chapter we have categorised them as either obtrusive or unobtrusive, however the same methods can also be classified in other ways.

Qualitative and Quantitative Measurement. These terms are used extensively in evaluation to describe the nature of the data being collected. Qualitative data is descriptive data derived by such means as focus groups, in depth interviews with key informants and members of hard to reach groups such as drug users and law enforcement officers, answers to open ended questions on surveys, and textual analysis of themes etc in other sources (usually but not exclusively written documents) such as minutes of meetings, logs or journals, correspondence, media reports, etc.. The information gathered by qualitative means can provide detailed information about people's thoughts and feelings regarding their drug use behaviour, the intervention, or related contextual factors. Drug use and its consequences varies according to individual, social and contextual factors and qualitative measures are particularly sensitive to variations in these conditions (Trotter & Medina-Mora, 1997). Qualitative data is most commonly gathered during formative and process evaluation to provided

feedback to project staff about modifications to intervention strategies and to provide insight about an issue or perspective. As such, it is generally considered a naturalistic method of investigation. Additionally, qualitative methods can be used to conduct on-going monitoring, and to generate new interventions. Qualitative data should not be used to indicate frequency of responses or as the basis for statistical analysis, although studies can often employ both qualitative and quantitative components (e.g. Lenton, Boys & Norcross, 1997; Lenton & Davidson, 1999).

Computer programs exist to assist in the management, analysis and interpretation of qualitative data. Two such programs commonly used in Australia are Ethnograph (ref) and Nud*ist (Qualitative Solutions and Research Pty. Ltd., 1997). In essence, these programs assist in making sense of large amounts of non-numerical data.

One advantage of good qualitative data is that it can be understood by most people as it is often presented through stories, examples, descriptions and discussion of processes. Such methods of presentation are a familiar part of every day life experiences (Trotter & Medina-Mora, 1997). Although to many, the analysis of qualitative data may not seem as daunting as applying statistics to quantitative data (see below), qualitative analysis is not 'easy', nor is it something that should be done without adequate training and understanding of the theoretical underpinnings of the various approaches and methods. Those wishing to employ qualitative methods as components of evaluation need to be aware of the boundaries of their competency and be in a position to enlist others with the necessary expertise to assist, or conduct such analysis.

Quantitative data are basically data, which can be counted, or turned into numbers. It is generated by a variety of methods and can take a number of forms including: categorical data such as Yes/No, rank data, such as which drug information brochure was best, second best, third best etc., or numerical data, such as the number of questions answered correctly. The most common method of collecting quantitative data is the survey, but many existing data sets collected by relevant organisations are often quantitative in form. The information gathered through quantitative means provides information about the number of people or places that demonstrated a certain behaviour, i.e., the number of communities that have a neighbourhood watch scheme. Quantitative data are most commonly used during impact and outcome evaluation in

the context of an experimental or quasi-experimental research design. Quantitative data are generally used to indicate frequency of an event or behaviour, in order to determine whether or not the intervention has generated statistically significant change.

Quantitative data are often seen as having more external validity than qualitative data, as the tools of sampling allow one to generalise from the sample to the population with greater certainty. However, such data have their own limitations. Quantitative methods can detach the investigator from the people being studied and the context in which the behaviours occur, grouping data often hides important individual differences, which can be important and useful, and quantitative methods often depend on numerous assumptions being met, such as sample size and the statistical normality of the sample.

Most people will be aware that computerised statistical packages exist for the analysis of quantitative data. Names that people may have heard of include SPSS (SPSS Inc., 1994) SAS (SAS Institute, 1988) and BMDP (BMDP Statistical Software Inc., 1988). Analysis of quantitative data can also be done 'by hand' with a calculator, but in most cases specialised packages are used as they save time and money. With a little bit of training and reading of the relevant manuals, most informed individuals can get data in and results out of such packages. However, training in statistics is required to ensure that the assumptions of the statistical tests employed have not been violated and the results correctly interpreted. Those conducting an evaluation using quantitative methods once again need to be aware of the boundaries of their competency and be in a position to cajole or contract in others with the necessary expertise to assist, or conduct, such analysis.

CHAPTER 6: THE PROCESS OF EVALUATION

INTRODUCTION

The purpose of this chapter is to provide a step-by-step overview of the kinds of tasks that will need to be completed in undertaking an evaluation of a project such as NCBADLE. It has been prepared for those who already have some expertise and experience in conducting research and evaluation, rather than the complete novice. Those with little experience should consider contracting in people with the necessary skills. This will in most cases have budget implications. Reference sources which have proved useful in preparing this material include Rootman and Moser (1984), Smart & Sloboda (1997), and Jean-Francois, Medina-Mora and Saxena (1997).

Many consultants, particularly those, who are university based, will usually negotiate rights to use and publish data, which is collected as a part of an evaluation project. This can bring credibility to the evaluation due to its independence and public presentation of findings for external scrutiny, but with it may come agency concerns about confidentiality and political sensitivities. It is important that at the outset all these issues are discussed and decisions made on how they will be addressed. Such agreements are then usually articulated in a formal contract or memorandum of understanding between the parties involved which is signed prior to the commencement of the project. Such an agreement, once signed, should be regularly referred to and considered in any other agreements which need to be made (eg. between the evaluators and agencies providing data for the project).

INITIATION

How the evaluation project unfolds will be greatly influenced by how it was initiated. In some cases the evaluation plan and the evaluation team will be already determined by the existing project parameters or staff available. In other cases the evaluation will be tendered out with the agency or 'purchaser' specifying broad parameters for the evaluation, but relying on the expertise of the successful tenderers or 'providers' to come up with an evaluation plan that will meet these parameters and budget. Usually there is more room for flexibility in the way the evaluation is carried out in the latter situation. It may even be possible to negotiate to modify the parameters articulated by the purchaser, if the case for this is well articulated.

In either case, thought needs to be given to the make-up of the evaluation team.

Useful questions include:

- Is there enough expertise among the members to cover all likely eventualities?
- If not, who else might be involved?
- Would they be part of the core team or provide a specific function but not be more intensively involved?
- Will there be an associated cost?
- What are likely to be the contributions (in time, content, etc) of each of the team members?
- If we need to recruit project staff how will we go about this?

The submission of an evaluation proposal requires a fair amount of work, but it is a necessary component of submitting an external tender, and is even becoming more common in internal tenders. An evaluation proposal usually involves:

- Description of the evaluation team
- Preparation of a background literature review
- Articulating a proposed evaluation design and methods
- Description of proposed outcomes of the evaluation
- Preparation of a budget and its justification
- Proposed time-line for the major evaluation tasks
- Discussion of ethical issues
- Copies of letters of support from important agencies (eg. those holding relevant data).

DETAILED PLANNING

Aims & Goals

The starting point for the detailed planning of most evaluations is to review the objectives of the evaluation specified in the tender brief. In a nutshell, what are the key questions that the purchaser wants answered? This will assist in identifying the information that will need to be collected and guide the selection and design of data collection approaches. There will also often be questions that the evaluation team have an interest in answering, which the purchaser may not have articulated. In most cases a rationale can be provided for why these questions are also worth attempting to answer for the benefit the evaluation team as well as those requesting the evaluation. Often the purchaser expects the evaluator to add value to the evaluation by extending its depth and scope beyond the basic parameters, on the assumption that it is within budget, timeline and other requirements.

As an example, one of the key changes negotiated by the tender team to the NCBADLE evaluation tender specifications was to make the evaluation much more process oriented than the strict impact and outcome evaluation model specified in the original tender. The evaluation team, believed that the success or otherwise of the four demonstration projects in the different localities would be determined in part by the process, which unfolded. As 'community' projects they would appropriately be influenced by the views of the local community, and how the intersectoral partnerships were managed. Thus, the focus of the evaluation was broadened beyond comparing the outcomes of the four interventions to looking at 'how' they were implemented. It is understanding this process, which could be useful in explaining why some projects were more able to meet their stated objectives than others. The NCBADLE Board of control accepted this argument and the evaluation, which eventuated, reflects this change of emphasis.

Review of existing documents

An obvious early task, but one that can prove costly if overlooked, is to carefully review all existing documents such as tender specifications, project proposals, evaluation contracts etc. This can serve to re-orient you to the task ahead and help to identify likely snags and problems, which might need early attention.

Potential audience for the evaluation

In planning the evaluation one needs to take into account the likely audience of the evaluation report. As well as the purchaser there are likely to be a number of other potential 'consumers' of the evaluation product. Who might they be? What might they be looking for? Which of them might you consider in the design of the project, and which not? The value of the evaluation report to the purchaser in terms of its 'impact' will often be improved if it seen as relevant to a range of potential consumers.

Identifying key stakeholders

Beginning early to identify key stakeholding individuals and agencies will greatly aid the process of evaluation. The evaluation team will probably want to interview some of these people to assist in project design as well as implementation. In the NCBADLE project these included Project Co-ordinators, action team and reference group members, local law enforcement personnel, and others.

Setting up steering groups and reference groups etc.

Invite appropriate stakeholders and others to join project steering groups and/or reference groups. These will be invaluable in identifying potential problems and the solutions to them and may help facilitate access to agency data, etc. Plan to hold initial meetings early in the project's life to get this process started.

Identify data sources

We have noted elsewhere that the best strategy to maximise the validity of data in such an evaluation is to use multiple convergent methods or 'triangulation'. This is often difficult in practice due to limitations on time and other resources. Decisions need to be made on what existing data sources will be integrated, what new data will need to be collected (and how this will be done) in order to fulfil the goals of the project.

In the NCBADLE project existing data sources included:

- The WA Police (WAPOL) database
- The WA Drug Abuse Strategy Office (WADASO) "Needle and Syringe Distribution, WA 1994 to 1997" data
- The Geraldton Public Health Unit's needle and syringe distribution database

- Data routinely collected by the Australian Bureau of Statistics.

New data collected included:

- Interviews with key informants at the inception of the project, the middle of the project and at its conclusion
- Police focus groups early and late in the project
- Observation and collection of minutes of DAT, Safer WA and NCBADLE Steering Group meetings
- Systematic review of the contents of local newspaper articles in the project sites

Ethics committee approvals

Most evaluation proposals which involve collection of data from human subjects will be required approval by an appropriate Institutional Ethics Committee (IEC). These are usually attached to tertiary institutions and hospitals. Some other organisations will also have their own ethics committees. It is critically important for the evaluators to have demonstrated that what they are proposing to do has been judged to not adversely affect the individuals whose responses provide data for the evaluation. Even if organisations or the purchaser do not require this, it is good practice to get appropriate ethics clearance for all evaluations. It is also worth noting that many professional journals will not accept papers based on research, which have not obtained such clearance.

In some cases where evaluation involves analysis of existing data bases which do not contain people's names (not 'name identified') it is not required that IEC approvals are obtained as it is considered that no adverse consequences could be experienced by individuals. However, it is possible that as a result of a research project or evaluation detrimental consequences can occur to a group or 'class' of people (eg. drug injectors, or dance party attenders), even if this does not adversely impact on individual research participants. For all the above reasons we believe most evaluations, including those such as NCBADLE, should pass formal submission to Institutional Ethics Committees. Note that this can take some time, but most IECs can advise on the likely duration of this process.

Other approvals

Most evaluations require other approvals to be gained apart from clearance by an ethics committees. Access to existing data bases, agency staff and clients needs to be negotiated with responsible bodies, and takes time.

Review budget, design, timelines

As the project develops you will need to monitor a number of critical factors. Are the project budget, evaluation design, project timelines still feasible and workable? How do they need to be modified or developed? Do any changes need to be negotiated with the funding organisation or other stakeholding agencies.

DATA COLLECTION AND MANAGEMENT

The real job of data collection will involve a mix of collecting new data and analysing data from existing data sets. If the preparation stages of the study have been done well, then the data collection is more likely to run smoothly. However, invariably there are problems and delays and for this reason it is important to allow ample time for this phase of evaluation.

You may find it helpful to carefully review a sample of early returns of data (eg. completed questionnaires, transcripts of focus groups, or records from an existing data source) to check that there have not been any errors or omissions. It might be that these can be rectified at this stage, rather than waiting until you get the whole data collection completed and find an error which compromises all that material.

The various data collections for the project need to be monitored closely. Each will have its own timelines, and management requirements. There is always a danger that important steps will be omitted or forgotten. Keeping a timeline, which summarises each project, can help to monitor the varied data collections. Keeping a daily log or diary of all contacts and events can also help to organise the data collection.

All data needs to be coded, cleaned, checked and if computer analysis is being used (either quantitative or qualitative) it needs to be entered into a computer. This can be a time consuming business. It may be prudent to contract out some of these tasks (eg. data punching) to businesses set up for these tasks.

Master copies of both electronic and paper versions of data collection instruments (eg. questionnaires or interview guides) coding formats etc. should be kept in a secure and accessible place. Copies of these will need to be included in project reports.

DATA ANALYSIS

Both qualitative and quantitative data analysis tends to be an iterative process. Initially crude analyses of the data (eg. frequencies and crosstabs of quantitative data) are conducted to get an overall picture of what is there. The selection of subsequent analyses builds on the results of earlier ones as well as things set at the beginning of the project such as the aims of the evaluation. As the report of the project is written, it may become apparent that some further analyses need to be conducted.

Again it is important that care is taken to organise both paper and electronic copies of data, command files and output files to aid in the writing of the report. NCBADLE was a large, multi component study and we have found it useful to have a hard copy of the output files when writing of the report. As the report was written we highlight output data as it is used, to stop double reporting and facilitate checking of the report against the data.

REPORTING

The writing of the report is addressed in the next chapter of this guide. The only additional points to be made here is that it is important that ample time is made to allow key stakeholders (police, drug user representatives, health service providers, policy bureaucrats and others) to review drafts of the report and provide comment. This might be one role for the evaluation reference group.

In most cases you will find that key informant comments will point to important editorial changes, which should be considered carefully. Comments don't necessarily have to be taken on board, but at the very least, the authors need to hear them.

CHAPTER 7: WRITING THE REPORT AND MAXIMISING ITS INFLUENCE ON DECISION MAKING

INTRODUCTION

The evaluation report is the focus and tangible product of all the work that was carried out as part of a program evaluation. This means it should distil all the relevant information in a way that illuminates the relationship between program intervention and any consequent changes. In this way it is a permanent resource that can aid decision making about the evaluated program and serve as a guide for future decisions about similar initiatives. Producing a good evaluation report is necessary to influence decision making, but it is rarely sufficient, because decision making in law enforcement, or indeed most human service areas, is part evidence based and part political. Accordingly, it is important to recognise the social context of evaluation and identify the various key stakeholders and the interest they have in the evaluation. In this way the evaluator can draw relevant people into the evaluation process; frame the report in terms that are meaningful to stakeholders, or even write separate reports for different groups and then work with decision makers and stakeholders to implement the evaluation findings in a way that is sensitive to the needs of all concerned.

WHAT IS AN EVALUATION REPORT?

Windsor et al (1994) define an evaluation report as:

the document that ties together a problem, program, analysis of program impact, and outcome. This synthesis is done by presentation of data to illustrate, if observed, cause-and-effect relationships. (Windsor et al, 1994, page 391)

The report should be succinct, dynamic and of immediate relevance to those with an interest in the program being evaluated. The report should discuss the program background, processes and outcomes and should provide guidance for decision making about the program evaluated and other similar initiatives.

THE AUDIENCE

The organisation funding the evaluation is likely to be an important audience for your report, simply because they paid for it to be written. However, there will also be

others and it is important that you consider whom you want to inform and influence when planning your report. Windsor et al (1994) consider that separate reports should be written to suit the needs of their respective target audience, but unless you are doing a very large and well funded project this is probably unrealistic. What may be feasible, is to have particular sections of your report tailored to the needs of particular target audiences by including more information of interest to them. Probably the best way of satisfying your respective audiences is to find out what they are interested in before conducting your study and writing the report. In this way you can ensure that relevant information is gathered and presented.

REPORT FORMAT

The format for a typical social intervention program evaluation report is presented in Table One. This provides logical organisation of material in a coherent sequence so that readers can easily identify and access the information they want. In shorter reports some of these sections could be collapsed, but the information will still need to be organised in a sequence that clearly allows findings to be related back to results and the nature of the intervention.

Table 1: Report format

Title page	title of report, names of authors, name of sponsoring organisation/publisher, date of release/publication
Table of contents	add lists of tables and figures if these are numerous
Executive summary	a brief overview of the problem, the investigation methods and the main findings
Introduction or background	this should provide an introduction to and rationale for the evaluation
Goals, aims and objectives	the major evaluation questions should be clearly stated
Method	this should outline the design of the evaluation and how it was conducted
Results	present results in a way that makes them easy to understand and indicates competent analysis
Discussion and conclusions	Interpret, explain and draw conclusions from the results
Recommendations	Base these on the evaluation's findings

Executive summary. This should be at the front of the report and provide a brief but succinct overview of the evaluation conducted. Few people will take the time to read the entire evaluation report so it is important that this section is well written and presents the major findings and recommendations in manner that has maximum impact with decision makers.

Introduction or background. This section should provide an understanding of the program being evaluated and why the evaluation is being conducted. It should contain a review of the literature on similar programs and their achievements and should convey the importance of finding out whether this particular program achieved its purpose.

Goals, aims and objectives. This should only be a short section, which clearly sets out what the program is intended to achieve. At its most elaborate, the section could

enunciate a long term goal, such as a reduction in drug related harm. It could talk about the process, by which this was achieved, such as issuing a caution and compulsory referral for treatment, rather than arresting drug offenders. It could nominate measures that indicated achievement of the goal, such as a higher employment rate and better physical health in the targeted population .

Method. This section should contain two elements. A description of the evaluation design framework, such as: was the intervention sample randomly selected; was a control group used; was qualitative or quantitative data gathered. Specific details on the techniques and parameters of the data gathering processes employed should also be provided, such as: what was the sample size; what selection processes were used; how was the intervention implemented. Windsor et al (1994) state that it is important to acknowledge the limitations of the evaluation in this section. Evaluations in naturalistic settings are difficult and imperfect and it is important to let readers know the particular weaknesses of your study.

Results. The way results are analysed can be presented in this or the method section. Windsor et al (1994) suggest that the method of analysis should be in the results section, however other authors consider that this logically belongs in the method section, with only the resulting data constituting results. Certainly there is little disagreement that this section should present results in a detailed and logical manner so that they are easy to read and understand.

Discussion and conclusions. Howe, Degeling and Hall (1992) consider that this is where you comment on your results and provide interpretation and explanation of their meaning. This section will probably be the most read after the executive summary and it is important to bring out the salient findings of the evaluation. However, you should be mindful not to go beyond your data in drawing conclusions about the effect of the program as this will undermine the credibility of the evaluation.

Recommendations. If your finding logically support recommendations these should be made. Recommendations, like conclusions, should be based on the findings of the evaluation. You will generally be on safer ground if you make recommendations on programmatic changes, such as generalisation of the intervention to a broader population. Broad policy recommendations will probably be difficult to sustain on the

basis of findings from a single program and should only be made if there is a range of other supportive evidence.

DISSEMINATING THE REPORT AND MAXIMISING ITS INPUT TO DECISION MAKING

An evaluation report that does not get to the appropriate decision makers, does not get read, or whose findings do not influence practice is of little use. Consequently you need to develop strategies to get your report to key decision makers and optimise its use in making evidence based decisions. Zweig and Marvin in Windsor et al (1994) emphasise that the education of one group by another, in this case decision makers by evaluators, will not be successful unless the process respects the institutions, culture and practices of each. Accordingly, you will need to consider how you present your findings to decision makers and other stakeholders in a way that respects their points of view. This issue will probably be particularly salient in the case of groups where a shared understanding of meaning is less likely. Ethnic groups and traditional Aboriginal communities are probably good examples and in such instances it is important to understand the group's frame of reference and interpret your findings within their framework. Zweig and Marvin also recommend evaluators take into account that their findings will never influence decision making as much as the day to day pressures. Accordingly you need to determine how you can present your findings so that they relate to priorities of the decision makers.

Finally, it is important to recognise that getting your findings implemented goes beyond how well they are presented in a report. Social circumstances impact on the influence of findings and you will need to use a variety of interpersonal skills to promote yourself as a good evaluator and your findings as timely and useful in the decision making process.

REFERENCES

- Australian Institute of Health and Welfare (1999) *The 1998 National Drug Strategy Household Survey – First Results*. Canberra: Australian Institute of Health and Welfare.
- Barnard, R.H. (1995) *Research methods in social anthropology: Qualitative and quantitative approaches*. (2nd. Edn.). London: Altamira Press.
- Basch C, Sliepcevich E, Gold R, Duncan D, and Kolbe L. (1985) Avoiding type III error in health education program evaluation: A case study. *Health Education Quarterly*, 12, 4, 315-331.
- Bevan, J., Loxley, W. & Carruthers, S., 1996. *Getting on and getting off in Perth 1994: Report of the Australian study of HIV and injecting drug use (ASHIDU)* (Perth, National Centre for Research into the Prevention of Drug Abuse).
- BMDP Statistical Software Inc. (1988) *BMDP Statistical software Manual, Volumes 1 and 2*. Los Angeles: BMDP Statistical Software Inc.
- Burgess, R. G. (1984) *In the Field- An Introduction to Field Research*, Allan and Unwin: London.
- Carruthers, S. (1997) Hepatitis C and injecting drug use: identifying the risks using videos. Paper presented at the First Australasian Conference on Hepatitis C, Sydney 16 18 March.
- Casswell, S. and Duignan, P. (1989) *Evaluating health promotion: A guide for health promoters and health managers*, Auckland: School of Medicine, University of Auckland.
- Dixon, D. and Coffin, P. (1999) Zero tolerance policing of illegal drug markets. *Drug and Alcohol Review* 18(4), 477-486.
- Fitzgerald, J.L., Broad, S. & Dare, A. (1999) *regulating the street heroin market in Fitzroy/Collingwood*. Parkville: The University of Melbourne.
- Green, L. W. (1986) *Community Health*, Mayfield Press: St Louis.

- Guba, E. G. and Lincoln, Y. S. (1981) *Effective Evaluation*, Jossey-Bass: San Francisco.
- Hawe, P., Degeling, D. and Hall, J. (1992) *Evaluating Health Promotion A Health Worker's Guide*, Sydney: MacLennan and Petty.
- Jean-Francois, M. (1997) existing information sources. In: World Health Organization (1997) *Guide to drug abuse epidemiology*. Pre-publication version. Geneva: World Health Organization.
- Jean-francois, MJ., Medina-Mora, ME. and Saxena, S. (1997) Reporting and application of results. In: World Health Organization (1997) *Guide to drug abuse epidemiology*. Pre-publication version. Geneva: World Health Organization.
- Kaplan, C. , Mercera, B. , Meulders, W.A.J. , Penners, G. , Bieleman, B. (1992) The " Boule de Neige" project ; lowering the threshold for AIDS prevention among injecting drug users. *International Journal on Drug Policy* 3(4), 170-175.
- Lehmann, S. (1975) Psychology, Ecology and Community: A setting for evaluative research. In E. Struening and M Guttentag (eds), *Handbook of Evaluation Research*, Sage Publications: Beverly Hills.
- Lenton, S., Bennett, M., Heale, P. (1999) *The social impact of a minor cannabis offence under strict prohibition - the case of Western Australia*. National Centre for Research into the Prevention of Drug Abuse, Curtin University of Technology: Perth.
- Lenton, S., Boys, A., Norcross, K. (1997) Raves, drugs and experience: A study of drug use among a sample of people who attend raves in Western Australia. *Addiction*. 92(10), 1327-1337.
- Lenton, S., Davidson, P. (1999) Raves, drugs, dealing and driving: qualitative data from a Western Australian sample. *Drug and Alcohol Review*. 18(2), 153-161.
- Lenton, S., Kerry, K., Loxley, W., Tan-Quigley, A., Greig, R. (in press) Citizens who inject drugs: the 'Fitpack' study. *International Journal of Drug Policy*.

- Lenton, S., Single, E. (1998) The definition of harm reduction. *Drug and Alcohol Review* 17, 213-220.
- Lough, G. (1998) Law enforcement and harm reduction: mutually exclusive or mutually compatible. *International Journal of Drug Policy* 9(3) 169-173.
- Loxley, W., Carruthers, S. & Bevan, J., (1995). In the same vein: First report of the Australian study of HIV and injecting drug use (ASHIDU) (Perth, National Centre for Research into the Prevention of Drug Abuse, Curtin University of Technology).
- Nutbeam, D. et al, (1990) Evaluation in Health Education. A review of progress, possibilities, and problems. *Journal of Epidemiology and Community Health*, 44: 83-89.
- Patton, M. Q. (1989) Qualitative Methods in Health Care Evaluation”, in Public Health Association. *Health Care Evaluation- Report of the National Health Care Evaluation Workshop*, 1989.
- Pirie, P. L. (1990) Evaluating Health Promotion Programs, In N. Bracht, (Ed) *Health Promotion at the Community Level*, Newbury Park: Sage Publications.
- Qualitative Solutions and Research Pty. Ltd. (1997) NUDIST Release V 4.0, Melbourne: Qualitative Solutions and Research Pty. Ltd.
- Room, R. (1991) Paradigms of social research on alcohol and drug use and dependence. Presented at International Symposium on Drug Dependence : From the Molecular to the Social Level, Mexico City, January 22-25, 1991.
- Rootman, I. and Moser, J. (1984) Guidelines for investigating alcohol problems and developing appropriate responses. Geneva: World Health Organisation.
- Smart, RG. and Sloboda, Z. (1997) Defining the problem (initial assessment of the situation). In: World Health Organization (1997) *Guide to drug abuse epidemiology*. Pre-publication version. Geneva: World Health Organization.

Rossi, P. H. and Freeman, H. E. (1985) *Evaluation: A Systematic Approach*, Third Edition, Beverly Hills: Sage Publications.

SAS Institute (1988) *SAS procedures guide : release 6.03*. Cary, N.C. : SAS Institute Inc.

South Australian Community Health Research Unit (1991) *Planning Healthy Communities*, Bedford Park: Author.

SPSS Inc. (1994) *SPSS 6.1 for Windows Base System User's Guide* , Chicago: SPSS Inc.

Strauss, A. & Corbin, J. (1990). *Basics of Qualitative Research*. Newbury Park: Sage.

Sutton, A. & James, S. (1996) *Evaluation of Australian Drug Anti-Trafficking Law Enforcement*. Adelaide, National Police Research Unit.

Trotter, RT. & Medina-Mora, ME. Qualitative methods In: World Health Organization (1997) *Guide to drug abuse epidemiology*. Pre-publication version. Geneva: World Health Organization.

Wadsworth, Y. (1991) *Everyday Evaluation On The Run*, Action Research Issues Association: Melbourne, 1991.

Windsor, R., Baranowski, T., Clark, N. and Cutter, G. (1994) *Evaluation of Health Promotion, Health Education and Disease Prevention Programs*, Mountain View, California: Mayfield Publishing Company.