199361

Final report 'Social Research – Getting it Right for Practitioners and Policy Makers'

Gloria Laycock

U.S. v UK distinction
– a problem boundence

ACCEPTED AS FINAL REPORT

Approved By: _____

Date: ___2/6/03___

CONTENTS

Chapter 1: Against this background ...

Chapter 2: What's the Problem?

Chapter 3: Getting it right

List of Figures

List of Boxes

List of Tables

Appendices

Chapter 1: Against this background …

1. The research proposal

The proposal on which this work is based is set out fully in Appendix 1. It has been modified during the course of its execution but remains focused on the ways in which three key communities can work together more effectively – researchers, practitioners and policy advisers. The practitioners in this case are the police, who serve as a useful example. But I will argue that the conclusions of the project are relevant to other parts of the criminal justice system and to the development of policy and practice in other areas of governmental concern such as health and education.

The issue of better linking research to policy and practice appears to arise in most advanced democracies. Desautels (1997) reports from Canada, for example, that in searching for existing program evaluations to support the work of the Canadian Auditor General, 'in too many cases we are finding that no information – or no useful information – on results is available' (page 74), and '… we were disappointed in the inability of evaluators to demonstrate the value added by their activities. … the case for evaluation still needs to be made...' (page 77). Desautels' comments raise the possibility that research and evaluation exercises are failing to influence policy and practice not because of a lack of interest on the part of the policy adviser or practitioner, but because the evaluations themselves are not delivering what these people need.

In similar vein, Kinsinger (1999) asked the rhetorical question in his paper presented to the American Society of Criminology – "How can it be that we have spent millions of dollars on research but we don't know how to reduce sex offending in the community?" His answer was that this is what happens when there is no focus on results – when researchers are allowed to run their own agenda and are not answerable to the funders for the production of a useable product.

There is nothing new in these complaints. Evaluations and research exercises more generally have been criticized for their lack of timeliness, irrelevance, lack of internal or

external validity and incomprehensible writing style for over 30 years (summarized in Chen, 1990). What *is* new is the growing demand for an evidence-base to policy development and practical action. This will put increasing pressure on researchers, in the broadest sense, to provide advice in a style and to a timeframe that meets the needs of those with the responsibility to deliver results, in an environment that is becoming increasingly outcome focused. One of the purposes of this study is to identify some of the reasons for the lack of progress in addressing this agenda over several decades, but more importantly to make some suggestions as to what might be done to improve the situation.

Reading and discussions with researchers and practitioners during the course of this project have confirmed that the time seems right for such a review. Indeed there is some support for the notion that any changes will probably need to be quite radical in addressing the way in which the research community relates to practitioners, and *vice versa*. It also seems unlikely that significant changes could be delivered in the short term without major structural alterations to the way in which business is conducted. Some of these changes are within the authority of NIJ or the Justice Department in the United States (the Home Office in the UK) to deliver but most are not. It seems likely, therefore, that we need to take a longer-term view if real and sustained progress is to be made in bringing policy, practice and research closer together. This means a change in what I have called the 'deep structure' of the organizations: Not tinkering around the edges but making fundamental changes to the ways in which the relevant agencies operate – a cultural shift in the business of research, policy and practice. Some of these changes are already underway, being driven by the increasing pressure for the delivery of 'outcomes' and hard targets in the public sector.

2. Something different

This is not a 'normal' fellowship report. There is very little hard data and an excess of opinion. It seems proper, therefore, to justify this (some may find it unforgivable nevertheless) by setting out a little of my personal background and how I came to the position that will unfold in the remainder of this report.

I was trained in psychology and mathematics at University College, London, during the late 1960s. Although I did not realize it at the time I got an excellent background in experimental psychology, which left me vulnerable to the label 'positivist'. I drifted into criminology, where this label was thrown about over the next decade, as an insult. I took it then as a complement.

My PhD was completed on a part time basis from Wormwood Scrubs Prison in London. At that time part of this large prison served as a major holding institution for young offenders on their way to 'borstal' – essentially prisons for those aged between 17 and 21. The trainees, as they were called, were assessed in the prison and then sent to serve a semi-determinate sentence in one of a range of establishments around the country. A number of them absconded, and that was the subject of my PhD thesis – how to reduce absconding from borstal.

I felt my PhD thesis had some useful messages for the governors and staff of these institutions. But it was obvious even to me that they were unlikely to read a thesis lodged in a London College. It was at the suggestion of Ron Clarke (then at the Home Office Research Unit and now University Professor at Rutgers School of Criminal Justice), that I rewrote the thesis for publication in the Home Office Research Study Series (HORS). It was eventually published as HORS 41 (Laycock, 1977). I learned a great deal about how to write, or rewrite, during that process, not because there was any formal training or guidance offered, but because a colleague in the next office had studied English at college; he was a good teacher and generous of his time in working through the many iterations of this report. I learned from this that a writing style appropriate to a PhD is very different from that for publication in a series intended for both practitioners and academics. I was also disappointed at the lack of formal training and support for junior staff in the Prison Psychological Service of the day.

When the report was finally published a number of copies were sent to the prison service Headquarters and stored in their publications cupboard. It was several years later, when

wandering down the Headquarters corridor that I overheard a senior member of the staff asking her secretary 'what are we going to do with that lot?' – referring to my reports which were accumulating dust. I realized that the report, on which I had labored long and hard, and which was the first such research report to be published by a prison psychologist (a point of some pride at the time), had had close to no impact. Indeed it was seen as something of a nuisance! I also realized, very late in the day for this piece of work, that research does not sell itself. It needs a dissemination strategy and a marketing program if it is to maximize its effect on practice.

In 1978 I moved from the Prison Service to the Home Office Research Unit where I worked on a study of maximum-security prisoners. Although I was largely unsupervised throughout the study, this was not a problem until it came to writing up the results. The message in the report was a difficult one for policy makers to handle. It basically said that their security classification was wrong at best and dangerous at worst. But is said this badly. It was written from a 'pure logic' point of view, which paid no attention to the political realities of the system, the pressures on both policy makers and practitioners and their initial assumptions and view of the world. For example, both practitioners and policy makers believed that they could assess 'dangerousness', which was the cornerstone of the security policy. The fact that there was a wealth of research and philosophical evidence pointing to the difficulties of assessing an individual's dangerousness was not a welcome message. Again, I learned the hard way not only of the importance of writing well, but of doing so in a way sensitive to the world-view of the recipients. This is even more important if the conclusions of the research are critical of the *status quo* and in a politically charged atmosphere.

Following this piece of work I moved to join a team of civil servants who were carrying out a 'Rayner Scrutiny' on overcrowding in prisons. These reviews were established by the then Prime Minister (Margaret Thatcher) under the oversight of Lord Rayner who was at the time on the board of Marks and Spencer. The reviews were not popular with the established civil service who saw, inherent in them, a criticism of their work. The review was carried out as an audit, which is different in style from research as will be

discussed more fully in the final section of this chapter, but the results were not well received, not in my view, because the work was badly conducted, but because it was delivered in a hostile context to an ill-prepared audience.

In 1983 I moved to the Home Office Crime Prevention Unit and eventually became responsible for the research and development program there. In that position, again with the support of Ron Clarke (by this time Head of the Home Office Research and Planning Unit), I was able to establish a new research series – Crime Prevention Unit Papers – which were intended for practitioners, primarily the police, although they were also expected to be academically acceptable. They were shorter than the HORS and either kept statistical and other experimental design features to a minimum or placed them in an appendix. Over the following years their presentational style evolved and their publication was accompanied by a one-page summary and a list of action points for the police.

An early report in this series described an operation identification program in North Wales (Laycock, 1985). It showed that although crime had reduced in the experimental area there was good reason to believe that it was little to do with identifying the ownership of property. This was confirmed in a later report (Laycock, 1992), which showed that the effect was heavily influenced by publicity. Despite this, the research is consistently quoted as saying that operation identification reduces crime (Eck, 1997; Welsh and Farrington, 1999). A notable exception is a full account of the research in a book by Pawson and Tilley (1997) to which reference will be made in Chapter 3. At this point it is fair to say that depression set in. If well-respected academics do not read reports then what hope have we? What seems to be taken from much of the voluminous papers that are written is a sound bite based on the title – "broken windows", "zero tolerance", "random patrol". The subtleties, often spelled out at length by the various authors, are lost. The report on operation identification *did* show a reduction in crime and the title of the report was "Property Marking: a deterrent to domestic burglary?" - it seems that not many people noticed the question mark!

This point was further reinforced during the1980s with the publication of a highly influential report on burglary reduction (Forrester, Chatterton and Pease, 1988). This report showed that if burglary victims on a high crime estate were protected *by whatever means sensible*, then repeat victimization, which was a feature of these estates, could be reduced, and burglary on the whole estate would also reduce. Attempts to replicate this study met with mixed success as Tilley (1993) showed. He also discussed the reasons why and I took from this some important lessons on how difficult it is to ensure that replications are carried out appropriately. This issue is also discussed more fully in Chapter 3.

In 1992 I established and headed the Home Office Police Research Group. This group was responsible for carrying out, or funding, social science research on policing in England and Wales. Its aim was '*to increase the influence of research on police policy and practice*'. From 1992 until 1999, when I left the Home Office for my fellowship at the National Institute of Justice, I tried, in meeting that aim, to put into practice many of the lessons I had learned from my career to that time. There is some irony in the fact that what follows in this report offends my positivist instincts but I find from the reading that I have been able to do as part of this fellowship, that I am in good company (Campbell, 1982; Cook, 1997; Christie, 1997).

In addition to my personal experience, as outlined above, I have benefited from discussion with a wide range of individual academics and police officers during the course of the last 12 months. A list of those who so generously gave their time and shared their experiences is given in Appendix 2. This report concentrates on the United States and the United Kingdom, although in the course of the study I also visited Australia and Israel where there are similar concerns.

3. Policing in the USA and UK

In order to focus the discussion (and because it is an area with which I am most familiar) the 'practitioner' group has been restricted to the police. In this section some of the key

characteristics of the police in the USA and in the UK are set out as they relate to issues that will emerge as significant later in the report. Of course there are huge differences in history and culture between these two sets of agencies, but there are also some shared bottom lines – how to reduce crime, to police with courtesy and sensitivity, to avoid corruption, and so on. And in neither country has research been fully integrated into the *modus operandi* of the services, although both jurisdictions recognize the increasing need to get better value from research and to see the lessons from investment being carried into practice.

Policing in the United States is a massive enterprise compared to most other countries, but then the United States is a massive country. There are almost 19,000 agencies at local, State and national level. The local police departments range in size from cities like New York, with over 38,000 uniformed officers and about 9,000 civilians to small, rural agencies with virtually one man and his dog – part-time. They share a number of common features – they are all armed; the criteria for selection, at least in the larger agencies, are broadly similar, and 7in 8 departments require officers to complete in-service training with an average annual requirement of 29 hours. More than 99% of local police departments are responsible for responding to calls for service, providing routine patrol, and enforcing traffic laws. About 88% have primary drug enforcement responsibilities (BJS, 1999). In addition to the local police there are federal agents at national level, sheriffs departments, agencies with special geographic responsibilities such as university campus police, transportation systems, parks and recreation facilities and airports. There are also State law enforcement agencies and the highway patrol.

The discussion in this report relates to those police agencies that serve major conurbations and deal with the bulk of crime and calls for service. The majority of police chiefs who head these agencies are appointed by or work closely with, the local mayor. As such they are particularly 'politically sensitive' and vulnerable to policy shifts that may be driven by single high profile incidents.

These features of US policing – its size, diversity, political sensitivity and the frequent movement of senior staff combine to make the integration of research and practice the more difficult.

The situation is different in the United Kingdom, which includes England, Scotland and Wales. Although the criminal justice system in Scotland has some significant differences from the rest of the UK, the policing style is not that different. In this report I will refer to the UK as a whole but acknowledge that there are some differences in Scotland, which are not reflected in the rest of the UK.

There are 43 police agencies in England and Wales, a further eight in Scotland, each with a senior officer (normally called a chief constable) who is appointed on a fixed term contract, perhaps up to ten years. He or she (there are now three female chief constables) is appointed through a fairly complex set of negotiations between the Home Secretary, a central government politician, and the locally based police authority, whose members are themselves appointed through a fairly cumbersome procedure, but whose task it is to ensure that policing in each force area is delivered in an efficient, effective and locally accountable manner. They do not have any authority over the way in which police operations are conducted but are otherwise an influential force in their area. Once appointed, chief constables are not easily removed until they come to the end of their contract. The UK police are therefore, although not totally free, somewhat less susceptible to local political influence than their US colleagues.

Although the Home Secretary does not have any direct control of the way in which police operations are conducted, he does have considerable influence over 'policing', through his oversight of the appointment of chief constables, his significant control of police budgets, and his statutory responsibility for the efficient delivery of policing. In exerting this influence there are a number of 'levers', which it is open to the Home Secretary to pull, and which are simply not available in the United States. In this sense there is, in the UK, far more central control.

The features of UK policing are, therefore, significantly different from those in the USA. There are far fewer agencies and they are each led by a senior management team that appears to be less susceptible to movement of staff than in the USA, although it does occur. They are less overtly political in the way in which they carry out their tasks; indeed the system is designed so that they should be. Finally, because of the potential for central influence over the way in which local policing is delivered, it is easier to see how research can play a part.

4. Defining research

'Research' means different things to different people. Journalists do research, so, for example, do historians, detectives, physicists, and astronomers. In this report I am restricting the definition to social scientists who collect and analyze data, interpret it and report on it. They may use quantitative or qualitative approaches, or both, and they may approach their work from a clear and articulated theoretical position or not. The extent to which they do this will become one of the issues relevant to whether their work is used or judged useful.

There are a number of groups, which fall within this definition of 'research', which I think need special mention and between which I would draw some significant distinctions. These are set out in Table 1 with some of the characteristics, strengths and weaknesses. The list in Table 1 is by no means exhaustive; it does not include survey methodologists, for example, nor statisticians, and it does not distinguish between the different disciplines from which the researcher group may derive – sociology, psychology, geography, and so on. These distinctions are extremely important in influencing the way in which research and evaluation exercises are approached and executed. And they will become more important as we move from evaluating existing initiatives to helping in the design and implementation of new projects.

13

Table 1: Classifying 'Research' *why this?*

	Characteristics	Strengths	Weaknesses
Auditor	Concerned with regulation and compliance. Audit institutions often possess statutory 'clout'. Becoming increasingly involved in performance monitoring.	Authoritative, quantitative, frequently attract media attention.	Lack 'depth', superficial. Concentrate on cost rather than effect.
Consultant	Commercially driven and often operating from large multinational companies.	Able to draw on a large pool of staff and can therefore respond quickly. Used to taking a customer brief.	No depth. Reports tend toward a management style of presentation – bullet points. Can be very expensive. Inflexible.
Not for profit consultancy	Small, entrepreneurial tendency.	Less expensive than a commercial consultancy company. Similar advantages although smaller scale generates other pressures.	Commercial concern to maintain company viability. May over commit to too many projects and deliver late.
Evaluator	Specialist in evaluation. May be associated with a University, consultancy company or 'not for profit' organisation.	Strong on experimental design. Used to listening to customer and interpreting needs.	May lack specialist knowledge of the area subject to evaluation.
Researcher	University based or affiliated. May be from a variety of different social science disciplines.	Strong on experimental design and theory. Able to contribute to the development of initiatives.	Subject to competing demands of university teaching program. Concerned with tenure and next grant application.

Auditors may well use many of the standard social science research techniques but they have typically, and traditionally, concentrated upon the financial side with a strong emphasis on regulation and compliance – a financial audit. They are frequently linked to a statutory requirement for review, which gives them considerable 'clout', and which others generally do not enjoy (Pollitt and Summa, 1997). They are less concerned with the detailed design of a project, tending to use a formulaic methodology, which is repeated consistently from one assignment to the next. They make use of existing data sets wherever possible and could, in the UK at least, be criticized for their lack of attention to the validity of the police data that they use (Diez, 1995).

This traditional role has developed over recent years with a move toward performance audit. According to Pollitt and Summa "auditors are now operating far beyond their

traditional sphere of competence as accountants. They have become interpreters of the expectations, objectives, good practice, and general principles of a range of other groups – policy makers, administrators, managers, and professionals." (Pollitt and Summa, 1997: 102-103). The criteria against which these judgments are made are far more debatable than the financially based criteria of the traditional audit, but the authority, and therefore the

> "[State audit institutions] possess statutory clout of a kind that evaluation organizations rarely if ever achieve."
>
> Pollitt and Summa (1997)

potential effect, remains. Auditors are, *de facto*, moving into research/evaluation areas, and while retaining their authority they are not necessarily developing the expertise to maintain the objectivity from which that original authority derived.

Consultants, including those in the 'not for profit' sector, which is much more developed in the USA than the UK, operate in a not dissimilar fashion to auditors. Those in the six or so major multi-national consultancy companies derive much of their authority from their marketing position, which they zealously guard. The reputation of these major consultancy companies often depends on the work of their audit departments, with the evaluation or public sector work being seen as relatively small beer. In the 'not for profit' companies, there is less emphasis on audit, and more interest in wider public sector work including the evaluation of various government funded programs. But there is both a formulaic approach to assignments and a lack of theory, or any appreciation of existing literature, in the way in which most companies approach their work. They can produce professional reports to tight timeframes, but there is frequently a lack of depth, and they have a tendency to reproduce what interviewees may have reported, without the insights or analysis that might be required.

The discipline of 'evaluation' seems to have developed in the United States during the 1960s and 70s. It has had a rather slower start in the UK. What, according to Weiss (1998:15), distinguishes evaluation from other kinds of research (by which she means 'basic' research) is not the method or subject matter, but the intent – the purpose for which it is done. Evaluation is intended for *use;* the questions it addresses derive from the concerns of the policy or program communities; it has a judgmental component –

15

comparing what is with what should be; it takes place in an 'action' setting; it can generate friction between the evaluator and the practitioner, and the results of an evaluation are not necessarily published. Weiss notes that 'evaluators are often so pressed for time, or so intent on moving on to the next evaluation contract, that they submit the required report to the agency and go on to the next study.' (Page 16)

'Researcher', which is the word used throughout the remainder of this report as referring to the individual who stands in contrast to the policy-adviser and practitioner, is in many respects a hybrid of some of the groups described so far. Researchers use scientific methods and the usual toolbox of techniques and information gathering strategies. Importantly, they also concern themselves with the testing of articulated hypotheses, which are rooted in some view of the genesis of human behavior. They would, therefore, be concerned either to work with a practitioner to develop such hypotheses, or encourage the practitioners to articulate the existing hypotheses before embarking on an evaluation exercise.

Many researchers, under this definition, are associated with a university department. They have an interest, therefore, in publication of the results of their endeavors in academic journals, and through other outlets, as a means of furthering their own careers and (in the UK) of gaining kudos for their university through the research assessment exercise (which scores departments according to the number of publications in refereed journals etc). There is, therefore, a *prior* assumption in favor of publishing the results of any evaluation or policy development exercise with which they may be associated. This, if you like, is part of the deal.

The distinction between researchers and evaluators has been made here in deference to the considerable US, and growing UK literature on 'evaluation'. It is, however, an unhelpful distinction in my view. The so-called discipline of evaluation joins together, as confederates, social scientists from a broad range of base disciplines under a shared banner, which unites them only in the sense that they use the same social scientific techniques. In so doing it obliterates some of the much more fundamental differences

between them such as their approach to the development of the mechanisms through which any particular initiative might work. In approaching any given social problem the kinds of hypotheses and the overall approach taken by a sociologist might differ considerably from that of a psychologist, anthropologist or economist. These distinctions are important if the move toward theory driven evaluation (Chen, 1990), or (my preference) scientific realism (Pawson and Tilley, 1997) is to gain ground. Glossing over the academic roots of potential evaluators puts the 'client group', be it practitioner or policy-advisor, in the position of buying a pig in a poke.

5. The audience

This report is addressed to a range of different audiences. Perhaps it is primarily relevant to those who are responsible for funding government social science research; those who in the United States work in the National Institute of Justice; in Britain, the Home Office, or in Australia, the Institute of Criminology in Canberra. There are similar organizations in other countries, and more are developing. These agencies essentially control the budget, either directly, or by virtue of their proximity to the governments that they serve: they are thus in a highly influential position in delivering change and influencing its direction.

There are three other major 'audiences' for the report – the research community in a broad sense, practitioners and policy-advisers. I have already discussed what I mean by researchers; practitioners and policy-advisers are described more fully in Chapter 2.

6. The study approach

The information on which this report is based was gathered from a variety of sources including the research literature, particularly that on evaluation, discussion (rather than formal interviews) with a wide range of professionals, both from the academic and policing communities, and my personal experience as outlined above. It does not aspire

to be a data based or empirical study but more of a 'think piece' on the inter-relationships between three broad communities – researchers, practitioners and policy-advisers.

6. *A map of the report*

Chapter 2 describes in detail the problem addressed in the report, first setting it in a wider context and then explaining the focus on policing. Chapter 2 also sets out some ideas on the way in which the key players see their different worlds. These views, it will be argued, go a long way toward explaining the problem and its intractability. Some of the key issues are identified which are discussed in greater detail in Chapter 3, where some possible solutions are introduced. This constitutes the major thrust of the report and works through the whole research exercise from commissioning to feedback on the results. Chapter 3 includes descriptions of some of the recent attempts to address various elements of the research/policy/practitioner interfaces – current good practice. In Chapter 4 we look at how far off we are from the ideals discussed and, finally, set out a possible agenda for the future direction of this effort.

18

Chapter 2: What's the Problem?

1. The big picture

In health, education, welfare, housing, criminal justice and transport – in fact, in all areas of public policy – questions are being asked about the relationship between evidence, policy and practice. 'Evidence' is presumed to stem from research; a systematic investigation aimed at increasing the sum of knowledge, rather than expert judgment, anecdote or any one of the many other possible evidential sources. Some areas, like health, have well-established research and development arrangements although even there we see a gap between the evidence on medical efficacy and the practice. And in other areas, like food safety and agriculture, there is a clear expectation that 'science' will guide policy. To quote the incoming Labor government in the UK in May 1997, 'what counts is what works'.

A number of reasons have been suggested for this interest in evidence. First is the increasing public skepticism of the 'professionals' – doctors, police officers, teachers, and particularly politicians, whose judgment was traditionally trusted but who are more and more being seen as at best well-meaning amateurs. An increasingly well educated, better informed and questioning public, together with the massively increased availability of data and other information through the Internet and other sources, fuels this skepticism.

In addition is the concern to see taxes well spent. There has been an explosion of auditors and other regulatory bodies over the past 20 years who have begun with a concern that public money was properly spent but moved on to consider the effectiveness of that expenditure. In the UK for example, total real spending on public audit bodies doubled between 1976 and 1995, there was a tripling of expenditure on inspectorates and equivalents over the same period, and a quadrupling of expenditure on ombudsmen (Hood et al, 1998).

In the United States, the Government, Performance and Results Act (1993) is a good example of the way in which demands in this area have developed. The Act was based upon the view of Congress that 'waste and inefficiency in Federal programs undermine(s) the confidence of the American people in the Government...' and that 'congressional policy making, spending decisions and program oversight are seriously handicapped by insufficient attention to program performance and results.'

Skepticism, accountability for the expenditure of public money, and particularly, an interest in outcome oriented program performance combine to raise the profile of research and its products.

> "We don't need a bridge [between research and practice], we need a superhighway"
> Mario Paparozzi, American Society of Criminology, November, 1999

In principle this increase in attention to research is clearly welcome but it has raised expectations of what science, in this case social science, can actually deliver. There is an uncomfortable expectation that research reaches unambiguous conclusions. Even in the hard sciences this is not universally so, but it is even less the case in many social science enterprises where uncertainly remains and judgments have to be made. This uncertainly has been dealt with in the medical arena in a fairly sophisticated way.

Medical research, which addresses the question of what works in treatment, has made extensive use of true experimental designs requiring the random allocation of subjects to treatment and control groups. These randomized controlled trials (RCTs) have been claimed as setting the 'gold standard' for evaluation. RCTs typically form the backbone of attempts to determine the most effective treatment for various illnesses. Such trials have been enormously influential in that context and there are now estimated to be over one million in print (Sackett and Rosenberg, quoted in Sherman, et al, 1997) covering a broad spectrum of medical practice. Even in the medical field, however, the results of these evaluations have not notably influenced medical practice, especially at the 'front line' where primary care is delivered (Sherman, 1998).

20

There are two issues here. The first is what the research actually says might be effective. Addressing the question of efficacy is not as straightforward as it might appear. Several RCTs need to be undertaken for a satisfactory conclusion to be reached in relation to any particular treatment program. (Unless the treatment is so manifestly effective that clear results emerge from a single experiment, which is an unlikely event.) The second issue is getting the message on what works out to medical practitioners in a form that they can use.

In response to both these issues the Cochrane Collaboration was established in 1992 with its first center in Oxford, England, following the influential work of Archie Cochrane (1972). This began as an attempt to summarize the many RTCs and to make them more accessible to the medical practitioner. The fairly standard procedure involves a review of all the relevant published work in any given area followed by the preparation of a summary paper which sets out the criteria used for inclusion of studies, the results and conclusions. The criteria are strictly adhered to and require randomized or quasi-randomized trials and the reporting of clinically relevant data (Cochrane collaboration web site). There are now over 50 collaborative review groups composed of persons from around the world who share an interest in developing and maintaining systematic reviews relevant to a wide range of medical issues. There are also ten methods groups in support with an interest and expertise in the science of systematic review. Increasingly the reviewers include non-randomized trials in their reports, although the degree of judgment that it required in doing so is stressed in the Cochrane collaboration guidance handbook where it is also conceded that a degree of judgment is required even in regard to randomized projects.

Since its inception the Cochrane Collaboration has developed into a sizable industry on a global scale. Their web pages are extensive and the whole program has proved extremely expensive for its funders. It is, nevertheless, a world-class initiative with huge potential for dissemination through the Internet. The RCT is particularly attractive in this scenario because it is easier (although not easy) to specify what constitutes an acceptable

experimental trial. The classification of completed research, and the task of summarizing it, becomes feasible, almost mechanistic.

So we have a problem with uncertainty. The questions, which policy advisers and practitioners address, are not usefully answered in probabilistic terms. And the techniques, which might reduce that uncertainty, as used in the medical profession so effectively, are less appropriate to the majority of questions raised in the wider public policy arena. Some of the reasons for this are discussed in the methodology section of the next chapter. In these circumstances the policy advisers and the practitioners have to make a judgment on what works and they have to do so in the context of all the other constraints upon the decision making process. These additional constraints, or influences, are set out diagrammatically in Figure 1 overleaf.

It is important that researchers understand the contexts within which policy advisers and practitioners work. As Chelimsky (1997) has noted "... our ability to serve policy depends as much on what we understand about how politics works as it does on the quality and appropriateness of our methods." (Page 55) This might be slightly overstating the point, but it is nevertheless a factor in getting research results utilized, and it applies equally to practice. An illustration of the importance of the political context and the way in which factors other than 'science' can affect policy decisions is given in Box 1 overleaf.

Some of the influences on policy and practice have been around for a long time, but others are either new – like the pressure for outcomes, or they are far more influential than in the past – like the media. These pressures operate within a broader shared context. First, both policy advisers and practitioners tend to have short memories. The rate at which senior staff are replaced, combined with the trend over the past 10-15 years toward 'delayering' with the loss of institutional memory, form part of the contextual backdrop.

22

Figure 1: Factors Influencing Policy and Practice

National budget
Precedent
Media (ntl/interntl)
Politics
Pressure groups
Research
Public perception
legislation

Policy
Local budget
Manager/leader
Performance regime
Training
Culture
Research
Legislation
Local media
Local expectations

Policy

Practice

Outcome
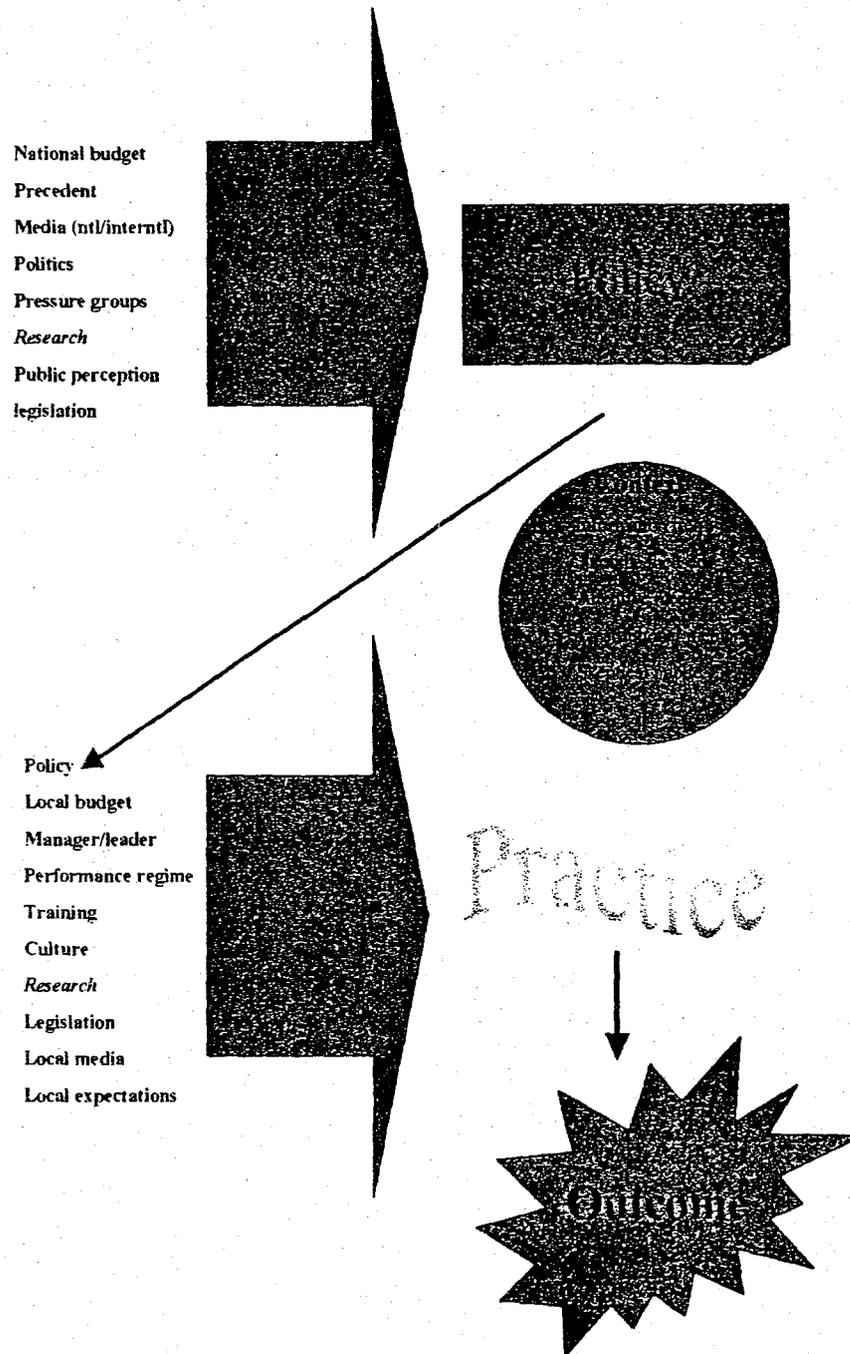
So does the availability of information. There has been a massive growth in the amount of information bombarding all sections of the public sector, including that which has been created as part of the move toward performance management. Further contextual commonalities are the reduced trust of the general population and increasing demand for the justification of public expenditure. These are all factors of which the researcher who wishes to influence policy and practice needs to be aware.

Box 1: Scientific evidence and policy decisions

The furore over British beef, and the ongoing dispute about its safety, provides a good example of the pressures on policy makers. This became an issue of global proportion with bans on the import of beef across Europe and beyond. In the United States, in addition to banning beef, there is a ban on blood donations from anyone who has spent more than six months in the UK between the years 1980 and 1997. The American Red Cross estimates that this policy will cut US blood donations by 2.2% at a time when blood donations are already falling. Cities like New York will be particularly hard hit since a higher proportion of the population there are frequent travelers to Europe. The Canadians followed suit with a similar policy.

The decision illustrates the interplay between adjudged risk and the potential loss of blood supplies. Theoretically, even a brief visit to the UK could lead to infection. The Food and Drug Administration chose the six-month limit as a means of reducing the risk without slashing the US blood supply. Banning donors who had spent less time in Britain would have had a far greater impact. The American Red Cross estimated that almost 23% of recent blood donors had traveled to Britain at least once between 1980 and 1996. The American policy was applied notwithstanding the view of Richard Daly, Chief Medical Officer to the Red Cross, that replacement donors for those refused on the theoretical arguments associate with BSE, are more likely to have an infectious disease.

The existence of these factors is important but so is their differential weight in the decision making process of policy-advisers and practitioners. If 'effectiveness' or 'outcomes' are to be taken seriously then the weighting of 'research' should increase relative to that of the other influences because it is through the process of evaluation and institutional learning that effectiveness can be increased. But the weighting given to research findings, in the judgment decisions of the policy-advisers and practitioners, is not necessarily a passive process, it is open to influence by researchers because, as well as being differently weighted, the factors are interdependent. So, for example, research results might be judiciously pushed toward the media, pressure groups, training programs

or project managers, which would increase their influence in the decision process. Some of these options are discussed more fully in the next chapter.

In the next section we look more closely at these various influences from the perspective of the policy-adviser, practitioner and researcher taking the example of criminal justice policy formation as the focus of attention for the policy-adviser, and policing as an example of a practitioner community.

2. The policy-advisers' tale

As a first step we should be clear what is meant by a policy-adviser. In the United Kingdom we are talking about senior civil servants, in the United States of America, political advisers at Federal, State or even local level whose responsibility it is to advise the governors and legislators on criminal justice matters. They are one step removed from politicians but are nevertheless aware of the political pressures operating at that level and therefore well advised to attend to the expectations and beliefs of the public. In making sensible recommendations about crime control, for example, many will be conscious of the long held public view that the police, acting through the criminal justice system, can and should be expected to reduce crime. And some will share that view. It is an approach, which generally leads to an expansion of the court and prison systems – it fuels cries for more police officers on the streets, and for tougher and longer sentences. It does not point, naturally, to the notion that environmental measures play a major part in generating criminal or disorderly behavior. So any advice reinforcing, for example, the role of opportunity in crime control (Felson and Clarke, 1999), has to overcome these beliefs as a first step.

This message has the added disadvantage that it is not cost neutral to the recipients. The members of the public, and a wide range of agencies in commerce and industry, have to take some responsibility for crime control – they actually have to *do* something rather than sit back and expect the police and others to protect them (Laycock, 1996). This is not an easy message politically. It involves, at least partially, handing crime control back

to the people from the State, with the accompanying responsibility (Christie, 1977). So research recommendations start on the back foot in having to overcome beliefs about the genesis of criminal behavior, which are not necessarily the most helpful in reducing crime.

The world in which the policy-adviser works is also a fast changing and demanding one. This is a common feature of the scene in both the USA and the UK although there are some major differences between these two jurisdictions – in the USA senior policy advisers are often appointed by politicians and may, for example, have an implicit remit to support the politicians' re-election or at least to tow the party line. The advisers are, therefore, not only sensitive to whether the messages from research fit with the public's expectations but they are also expected to address the extent to which research results are compatible with the political complexion of their patrons. In the UK the position is somewhat different. Although there are politically appointed advisers in all government departments, and they have grown in number over recent years, they are still vastly outnumbered by the permanent senior civil service whose advice is *expected* to be politically impartial and factually based. Whether it is, of course, could be debated.

In both the USA and the UK recent administrations have begun to look for 'what works' as a means of improving efficiency and reducing costs. This interest has been something of a two edged sword. On the positive side it has led to substantially increased demand for research, with the associated funding and implication that policy might finally be influenced by it (Home Office, 1999; National Institute of Justice, 1997). More unhelpfully, it has, also exposed the amount of money that has already been spent on research, and the relatively little by way of firm evidence on effectiveness that it has produced (Sherman et al, 1997; Goldblatt and Lewis, 1998).

In both jurisdictions, however, it is probably fair to say that the advisers are looking for 'good news' ie the policies were effective – in the crime prevention world, crime went down, the public was less fearful, everyone was happy. So the first somewhat cynical requirement from research is that it produces what the politicians want to hear. And it

26

needs to do so quickly. There is an impatience about the process of government that leaves research at the starting block. While the average university professor is beginning to contemplate the enormity of the research exercise facing him or her, the politician is expecting an answer, and an unequivocal one at that. Conclusions that 'more research is needed' receive a very bad press.

"Research findings tend to be assessed for their practical and symbolic value, rather than their inherent intellectual worth. Consequently, whether a research study is embraced, ignored, misrepresented, or attacked is likely to be contingent on how various players in the policy process see that research fitting into their own political agendas."
(Brereton, 1996)

Even when a commendably empirical approach to policy development is adopted, with plans for a pilot scheme to be established and evaluated before the process is fully launched, the enthusiasm of the political process to press ahead and announce a new initiative takes precedence over waiting for the results of the pilot scheme (Nutley and Webb, 2000; Walker, 2000).

Despite their wish for good news, and the rush to announce new initiatives, the policy adviser in this new and developing outcome focused world, will increasingly need to have confidence in the conclusions of the research exercise, particularly if large sums of public money are at risk of commitment to new programs. In the 'old days' it mattered less if decisions were made on the basis of political expediency rather than efficacy, but now, with the public, the media and others watching the public sector equivalent of the 'bottom line', confidence in research conclusions matters. And an additional dimension has been added with the interest in reducing the cost of central government activity – not only do policy advisers want to know what works they want to know at what cost. Cost-effectiveness is increasingly being built into the evaluation plans for new studies (Dhiri and Brand, 1999; Colledge et al, 1999).

Next is the delicate matter of money. There is a tendency for policy-advisers in central government to see themselves as the custodians of public money – and he who pays the piper expects to call the tune. In the USA the process is more directly dependent on the views of Congress, where research money is voted for specific purposes, with at present a

relatively small amount for discretionary use. The degree of flexibility within the discretionary allocation varies but the Head of the Federal Government's criminal justice research agency in the United States, is a Presidential appointee, and is expected to take the Government's priorities, and view of the world, into account when setting the agenda.

In the UK the position is again rather different. Here the money voted by Parliament for centrally funded criminological research has generally been for the support of the criminal justice policy process. The policy-advisers were, therefore, seen as the customers for the work (although Ministerial approval was generally expected) and felt that they should have the major say in research expenditure (Cornish and Clarke, 1987). More recently the picture has changed somewhat with the 'legitimate user' being expanded from the policy-advisers to include the practitioners, and with more weight being given to the experience of the government researchers and their views, in setting the research agenda. The relationship is now better characterized as a set of partnerships between policy advisers, practitioners and researchers (Laycock, 2001). Whatever the subtleties of the developing relationship between researcher and policy-advisers, the policy-advisers feel they have a significant role to play on both sides of the Atlantic – and that they should not, therefore, be ignored in determining the research agenda. While this partnership approach dilutes the political influence over the research agenda, in reality it does not remove it. Political ideology at least maintains a 'filtering' role (Doherty, 2000) over what research is done, how it is presented publicly and what influence it may have over future policy direction.

Finally, and in some ways this is hardly worth saying, policy-advisers need to be able to understand the results of a research project. Research papers that are written in obscure technical language, permeable to the chosen few, and covering reams of paper, are less likely to influence policy than a concise, crisp few pages which summarize the important points and spell out the implications for policy. Some researchers are distinctly uncomfortable with this scenario, particularly the idea that they should perhaps go somewhat beyond their data in spelling out the policy implications of what they have done. This latter point is contentious. Whether or not researchers articulate what the *think*

or believe, on the basis of their research experience, rather than what they feel they *know*, on the basis of the conclusions of their work, a policy decision will be probably be made, or some practical steps will be taken. The point at issue is whether or not the researcher contributes to that process by, if necessary, speculating. It may not arise in relation to all studies, or even the majority, but there will be occasions when the researcher knows more than anyone else about the subject of the research, and is therefore in a good position to influence the policy-making or practical process. 'Opinion giving', should not be problematic, if it is made clear at all stages the status of the advice *as* opinion.

The requirements of the researcher from the policy-adviser can therefore be summarized as:

- Good news
- Confidence in the results
- Cost included in evaluations
- A feeling of involvement in the agenda setting process
- Timely production of results
- The identification of 'what works'
- Good communication skills
- A willingness to take risks in making inferences from their data.

The practitioners' tale

The practitioner can be characterized as an individual who is expected to deliver the policies 'on the streets'. We are concentrating on the police who quite literally, come face to face with the beliefs and expectations of the public in a way in which the more remote policy-advisers do not. As in other areas of public service the world is changing rapidly for police officers. It is no longer good enough to point to a list of outputs from the field – more tickets issued, more arrests made, more neighbourhood/block watch groups or partnerships established. The pressure is on for the delivery of outcomes. There are a variety of ways in which this is manifesting itself. In the UK, for example, the Crime and Disorder Act (1998) requires the police and local government to work together and

produce a strategic plan for the reduction of crime and disorder, which increasingly is expected to include hard targets for reduction. Indeed the Government has set a national target of reducing vehicle crime by 30% over five years.

In parallel, still in the UK, a new regime of 'best value' has been introduced under which police forces are expected to examine their policies and procedures on a regular cycle and demonstrate that they are achieving value for money. This does not necessarily mean going for the cheapest option, but going for the most cost effective (Leigh et al, 1999).

> "Society should hold professors and police officers responsible for whether their institutions contribute as effectively as they might to the objectives of education and public safety. Unfortunately, these very institutions often make it difficult for talented, dedicated and far sighted faculty and police officers to be responsible for what they are doing in ways that they and their communities would want them to be."
>
> David Baley (1994)

In the USA one of the more tangible manifestations of this move toward an outcome focus is through the Compstat process first used in the New York City Police Department and now being replicated in a number of police agencies across the country. Compstat is, amongst other things, a management tool forcing the attention of precinct commanders to the crime and disorder statistics in their area. Typically there will be a map of the precinct with the recent crime and disorder incidents displayed. The commander will be expected to give an account of what he or she intends to do about the offending pattern. How will it be reduced? Agencies differ in how they approach the Compstat process (Weisburd, 2000) and there is no clear concensus about what is involves, but on the basis of a recent survey Weisburd and his colleagues estimate that, across the United States, about 33% of police agencies with 100 or more sworn officers are using processes of this kind.

In both the USA and the UK there is also a significant move toward problem oriented policing (Goldstein, 1979, 1990). Here the police are expected to solve problems rather than to continually react to them. Again there is an expectation that what might have been a recurring problem will be addressed in such a way as to reduce the rate at which the incidents occur, or to eliminate them – a focus on outcome rather than output. In the USA this process is generally set within the framework of community policing. In 1997 the

Bureau of Justice Statistics estimated that over 60% of departments serving 100,000 residents or more, had a formal written community-policing plan.

One collective consequence of all these developments is the increasing pressure on practitioners to implement cost effective tactics to control crime. And they should not be expending resources on ineffectual activities. There is, therefore, a need for the practitioners to know what works, where, and at what cost. They are looking to the research community to provide some answers (Sherman et al, 1997; Goldblatt and Lewis, 1998).

In this respect the picture has changed dramatically over the last five or so years. In the past the police researchers and police officers could operate in independent universes for all practical purposes. The researchers could carry out their research *on* the police and report the results in their academic journals. The police could complain about the unfairness of the results and the frequent negative findings – nothing works – and ignore the outcome. This is clearly a caricature of the situation, but it is not that far off. If we calculate the amount of money spent on police related research in both the USA and UK over just the past decade, and then look at how much it has impacted operational policing, then we might wonder where the money went. Although some commentators argue that the influence has been subtle but important, on for example, police responsiveness and accountability (Bayley, 1998), the view taken here is that at best it provides poor value for money – if, that is, changing the way the police do business was the intention of the research, and it often was not.

Like the policy makers, the practitioner could well be involved in the agenda setting process. Although they do not have resources to spend on research to the same significant extent, a great deal of research time can be saved if practitioners are involved in the process of agreeing the research program from the beginning. For example, the policy makers might commission a research project to determine the cause of public disorder in city centers. The researchers might respond by carrying out an extensive study of the data on disorder and discover that it tends to occur on Friday and Saturday nights at the time

31

the public houses, bars and dance halls close. Analyzing data of this kind can be very time consuming and it has to be said that if the police had simply been asked what they *thought* the primary cause of disorder was, that is exactly what they would have said. The conclusion could have been reached in record time by looking at police deployment patterns. Resources would then be freed up to address the more relevant question of what can be done about it.

It is also sensible to involve practitioners in identifying research priorities for other reasons. Because of their proximity to 'the streets', they are often the first to become aware of a rising crime problem, which in the normal course of events, might take months or even years to permeate the policy consciousness. Mobile telephones are a good example. The UK police were painfully aware of the extent to which these devices were involved in crime – being stolen and/or used for drug dealing etc – well before it became a national issue. An ability to pick up these emerging problems early would be a major advantage of involving practitioners in determining the research program. Not only could the emerging problem be quantified and described, but a timely research-based response could nip the problem in the bud.

Many practitioners also appreciate some research support in tackling problems unique to them. They are on the receiving end of policy decisions and are often ill advised on how, for example, the decisions might best be implemented and what others are doing to deliver the latest proposals in a cost effective and timely manner. They need to know not only what works, but where and why. An example of this is neighbourhood or block watch (NW). The policy view, driven by political sensitivity to public interest, was that NW was a good thing and should be expanded. This was in the face of considerable research evidence that it was not noticeably effective in reducing crime (Bennett, 1990; Rosenbaum, 1998). There are, however, a number of reasons for encouraging NW in an area, of which crime reduction is only one. The alternative reasons have resource implications. Where crime is a problem NW might be quite difficult to introduce but *if* successfully implemented could reduce crime (Forrester et al, 1990). Where crime is less of a problem, then NW can maintain or generate police/public relations and can do so with fewer resources than are demanded in high crime areas (Laycock and Tilley, 1993).

NW is an initiative that is highly sensitive to the context within which it is introduced in what it can deliver. Practitioners need to be aware of these subtleties.

What practitioners critically need to understand is *why* an initiative may or may not have worked. Knowing this will help them, and their research colleagues, in deciding where and how to replicate the initiative (Tilley, 1993) and ultimately to 'mainstream' it. At some point the research process needs to lead to the articulation of the 'mechanism' (Pawson and Tilley, 1997) through which the initiative is presumed to be effective. This amounts to saying what the hypothesis is that might lead to a reduction in crime. In the US literature this is called the logic model or program theory (Weiss, 1998).

Like policy-advisers, practitioners also need to be able to understand the results of the research process. Plain English, with the practical implications of projects written clearly. Busy practitioners do not have the time, nor in many cases the interest in immersing themselves in the detail of research, and many will not have the background training to understand complex statistical arguments. A different product is required which meets the needs of the practitioner audience.

Finally, and perhaps a relatively trivial point, practitioners appreciate (even if they do not 'need') feedback, when they have given their time, resources or effort to support a research exercise in their agency. Too often there are complaints that researchers come and go – collecting data, issuing questionnaires or generally making demands, without providing any feedback on the final results of the exercise.

To summarize the requirements of the practitioner, they need:

- To know what works, where and why
- Help in replicating 'what works' – understanding concepts and mechanisms
- Help in generating testable hypotheses
- Timely research
- Involvement in setting the agenda

33

- Reports and recommendations in plain English
- To know of current good practice
- Feedback on the results of research in which they have participated.

None of the requirements outlined above are in any way illegitimate or unfair. They simply reflect the realities of life. If research is to be of use to policy-advisers and practitioners then it needs to be attentive to the needs of these communities.

4. The researchers' tale

If these expectations of research are fair and reasonable why have they not been met as extensively as they might? There are a number of possible reasons, largely stemming from the way in which research is organized and funded both at national level (where decisions about expenditure are made and where the focus is on project management) and within the university or consultancy sector (where the bulk of the research is carried out). There are also reasons related to the way in which researchers in social science have typically been trained, with an emphasis on methodology and relatively little on understanding the complexities of research in the 'real world', which includes a perspective on the pressures under which policy-advisers and practitioner work. And finally, government funded research has been no more clearly outcome focused than the processes and procedures that were the subject of that research. And there was no requirement that it should be.

The majority of research expenditure resulted in published articles, academic books, briefing notes, conference papers – all output measures, not outcomes. Indeed, outcome measures for research are particularly difficult to develop. It is easy enough to count the number of reports circulated, or delegates at a conference, but it is difficult to determine the extent to which research has made a real difference to outcomes in policy or practice, particularly given that research is but one of many possible influences over what is actually done.

34

It may be easier to measure the extent to which research has influenced policy and practice if the research exercise includes a list of action points for the policy-adviser or practitioner. Although this is certainly not without its drawbacks – some recommendations are worthy but hardly earth shattering (better record keeping) and some highly demanding (removal of judicial discretion). Simply summing recommendations is of questionable value. Nor would summing the recommendations enable us to estimate the more subtle influences – more enlightened or critical thinking by the research recipients, a greater use of the ideas and generalizations, a contribution to organizational learning (Argyris and Schon, 1996), changes in the local decision process, later contribution to a meta-analysis, and so on (Weiss, 1998b).

Not many research exercises have been designed explicitly to influence policy or practice. Researchers, like policy-advisers and practitioners, have been working to their own agenda. In their world, they are rewarded for articles published in refereed journals, for the number of citations they amass, and for attracting research grants. They are concerned with tenure policies and, at a professional level, are rightly expected to attend to the quality and professionalism of their work. There are no brownie points for sticking the academic neck out and speculating beyond the collected data set about the implications for either policy or practice of the research exercise on which they may have been engaged. And some might say quite right too. They may be expected to contribute to a policy decision - to give an opinion - but they would not be expected to make a policy decision on the basis of their work. That is for others.

Researchers are also wary of being seen to be 'too close' to either the policy-adviser or the practitioner. On the policy side they may be vulnerable to pressure to come up with a good news story, which supports the party line. On the practitioner side they may be influenced by personalities, be caught up in the detail of the local initiative or, for other reasons, simply find themselves investing personally in the outcome of the study and produce bias in their results as a consequence.

35

Practitioners, in expecting researchers to provide advice not only on what works but where and why, are indirectly influencing the methodological choices open to researchers. RTC's for example, are not an appropriate approach for testing what works for whom since in order to test for interaction effects the sample sizes have to be large, often far too large for all practicable purposes. Furthermore, as researchers are pulled increasingly into the position of helping practitioners to design and develop new tactics to tackle practical problems, they are calling for skills going beyond those of the traditional 'evaluator'. As Weiss (1998b) has noted, there is some pressure on evaluators to take on a broader assignment and become consultants and advisers to programs. She questions whether evaluators have the skills for the expanded role arguing that the evaluators skills are those of the technician in designing an evaluation framework and an ability to listen to what is required and communicate the results. She does not see the generalist evaluator as the best source of new program ideas nor presumably of the hypotheses about human behavior that might generate those ideas. But that is what is required. Neither the practitioners, nor for that matter the policy advisers, need the kind of theoretically sterile evaluation which has too typically been produced. Even the more interactive and responsive evaluation of the kind embraced by Weiss will not fit the bill, although it is certainly a step in the right direction.

All of these issues and concerns affect the way in which researchers carry out their tasks, including their choice of methodology. In the next section we will discuss some of the implications of what might be called a 'new agenda' (Kennedy, 1999) for researchers. This new agenda will need to meet researchers needs, but also those of their 'partners'.

5. A new agenda

I have argued thus far that there is a major new demand on the police to deliver outcomes in terms of reduced crime, and that in order to do so they will need to embrace research and evaluation in a way that they have not so far done – either in the USA or the UK. There are several points to make in relation to this.

First, some police agencies would argue that they already know how to reduce crime and can respond effectively to a crime reduction agenda. How true is this?

The debate on this question has raged on both sides of the Atlantic. In the USA many major cities have experienced a reduction in crime despite the fact that their policing styles are very different. New York, for example, is characterized by its focus on Compstat and an alleged 'zero tolerance' approach to disorder and minor crime. In San Diego, where crime has also fallen, the approach is better seen as 'problem-oriented policing'; in Chicago it is CAPS – 'Chicago's Alternative Policing Strategy'. The extent to which these reductions are attributable to police action or have arisen for other reasons – demography, changing patterns of drug use, or the economy for example, has been hotly debated (Blumstein and Rosenfeld, 1998; Kelling and Bratton, 1998; Bowling, 1999).

In the UK police recorded crime fell for five consecutive 12 month periods (ending September), until September 1999, when there was a 2% rise (Home Office 1/00). The fall had been broadly consistent with what would have been expected on the basis of changes in the economy and demography although there was *some* evidence that it was slightly greater than expected with the extra reduction being attributed to policing activity.

It is not appropriate in this report to go into the details of this debate. It is probably fair to say, however, that the issue has been far from clear-cut, and that a chief officer on either side of the Atlantic would be ill advised to claim full credit for those reductions that we have so far seen. This is particularly true given the possibility (and in the UK the reality) that crime may rise again.

So it seems plausible that police action may have contributed to the observed reductions in crime, but it also seems plausible that police action did not account for the full reduction; that insofar as it contributed, it did so through a variety of mechanisms and

with varying degrees of success, and that nowhere was there an evaluation of police performance which formed part of a deliberate intervention. As Campbell noted:

> "... deliberate, intentional intervention greatly reduces equivocality of causal inference in contrast with efforts to infer causality from passive intervention." (Campbell, 1982: 330)

The majority of claims that police action resulted in reductions of crime or disorder were made on the basis of *post hoc* analyses of observed reductions.

Secondly, some might argue that the police already pay sufficient attention to research. It is certainly contrary to the thrust of the thesis in this report that the police and researchers are working well together, but the veracity of this should perhaps be considered – to what extent do the police and researchers currently work well together? And – a slightly different question – to what extent does present day policing reflect the lessons from research?

In addressing the first question, there is relatively little evidence of a close working relationship between the police and research communities. Some police agencies have 'research' units within them, but they do not do research in the generally accepted sense. Hirst (199?) describes the situation in the UK where there were a plethora of titles for research-type exercises across UK police forces but only one or two forces with staff holding post-graduate research qualifications, and located within the organisation at points where they could realistically influence practice. The police-employed 'researchers' collect information, collate data and support management practices, but in the sense of research as discussed earlier in this report – formulating and testing hypotheses on the basis of sociological or psychological theories, there is very little to show. Nor are there many police agencies with a regular and productive relationship with their local university, although examples do exist. In the UK, the Humberside force established a close working relationship with the University of Hull, and in the US, there has been a long-standing and productive collaboration between the Chicago Police and

Northwestern University, led by Wes Skogan. But these relatively prolonged and productive relationships are rare.

The second question – has research affected policing – is different. This issue was touched upon earlier when discussing the extent to which research is outcome oriented and there it was suggested that the effect of social science research on policing was fairly minimal given the extent of expenditure. There have, however, been some successes. In the United States, the current vogue for community policing can be traced to research-based roots (Rosenbaum, 1998), and the development of problem oriented policing began with the work of Goldstein (1979, 1990) and has been highly influential in policing in most of the English-speaking world. The UK investment in police research, with the establishment of the Home Office Police Research Group (PRG) in 1992, marked a deliberate attempt by central government to influence the way in which police policy and practice was developing across the board. The PRG was commissioned to *increase the influence of research on police policy and practice* and did so by working in close partnership with police agencies in developing the agenda; funding practically oriented research and development and paying increased attention to the ways in which the results might be disseminated. Despite this effort, which met with some success as will be outlined in the next chapter, the *need* for research, by the police agencies themselves, has still not been fully realized.

Furthermore, if the police as an institution, aspire to professional status, rather than being seen as service providers, then they need to concern themselves with the establishment of a body of knowledge from which their professional status might derive. The difference between a professional and a service provider is essentially one of knowledge and a strongly ethical work context. We treat doctors as professionals because they are drawing on a body of well-established knowledge in providing the care we need and we expect them to behave with integrity. Medical ethics is an industry in itself. There are no comparable bodies in policing - no overarching 'ethics police' with the power to strike 'offenders' off. Nor is there a published knowledge base, as the recent review by Sherman et al on what works in crime prevention, for example, has shown. The only way

39

to get a real body of knowledge established is through the process of systematic and prolonged investment in research. This is not to say that academics have not concerned themselves with the concept of policing to date, on the contrary there is a huge international literature on the subject. But there is a very small literature on *what works* in policing and there is even less to say on *where or why*.

There are, therefore, three reasons why the police might be interested in supporting the right kind of research. First, proper, pro-active evaluation of police initiatives against crime would lend credibility to police claims that their action was instrumental in delivering observed reductions. Secondly, where the police are less clear on how to reduce crime, or to deliver any of the other requirements of the police, such as justice and integrity, researchers should be able to work with them to establish research-based programs, which might prove effective in addressing the questions at issue. And finally, the establishment of a body of knowledge on what works in policing is a prerequisite for the firm establishment of a profession.

This new agenda would mark a partnership between police practitioners and academics, which would be driven by hypothesis-based enquiries into what works where and why. There would be a presumption that the work would be published in a form that was readily accessible to practitioners but also readily available to the wider research community where the quality of the work could be fully scrutinized.

In developing this new agenda attention needs to be paid to the whole process of research, from commissioning, through execution to delivery. This process is discussed in the next chapter.

40

Chapter 3: Getting it right

This chapter describes the research process, which applies when a central government organization is charged with procuring a research 'product' from a university, consultancy or not for profit agency. The various stages of production are worked through, from commissioning, carrying out the work, producing the results and disseminating them. There are a number of issues arising at each stage of the process and these are discussed. The chapter ends with a note on 'good practice', which has emerged during the course of this project, mainly from the US or UK but in one case from Australia, and which might reasonably translate to other countries. It also discusses some of the principles on which developing good practice might be based.

1. The commissioning process

a. Involvement of others in topic identification

There are probably four sets of individuals with a direct interest in the way in which the research agenda is set – the policy advisers, practitioners, practicing researchers and government research funders. These groups are variously involved in the agenda setting process in the UK and US.

The British Home Office Research, Development and Statistics Directorate currently manages central government research funds for criminal justice, including policing. They have a remit to address the research needs of policy advisers and involve them in identifying topics for research through an annual process of inviting proposals for new work. As well as seeking the views of policy advisers, the former Home Office Police Research Group involved the police practitioners in the agenda setting process, as represented by a committee of police chiefs (Association of Chief Police Officers (ACPO) Research Committee), or by other, ACPO functional committees (eg Crime Committee, Personnel and Training, Quality of Service), which advise chief officers on elements of the policing policy in the UK. There was no systematic attempt to involve the practicing research community in this process, partly because they were likely to be in competition with each other for any research funds eventually made available.

41

The procedure in the US is very different and more inclusive of practicing researchers. NIJ may, for example, hold a seminar or workshop for a day or more, and invite a selection of academics to discuss emerging trends and possible research topics. Policy advisers from 'main justice' – the administrative arm of the Justice Department would not necessarily be invited but the Head of NIJ, a Presidential appointee, and therefore an influential player in determining the final agenda, would often attend. Practitioners, in the form of a selection of police chiefs, would also be invited although they tend to be selected from a fairly small (given the size of the potential pool of US chiefs) group of individuals known to be generally supportive of research.

b. <u>Projects and programs</u>

A single research project normally covers a fairly tightly defined topic and is carried out by an identifiable research team with an individual professor or 'principal investigator' responsible for its completion. It is normally conducted in isolation from other projects being funded through the same agency. Specific projects of this type may need to be commissioned in response to a political wish for urgent advice on a specific topic – an assessment of the use of police force – for example, which may be in response to a high profile incident or series of incidents.

Some research agencies list all their current projects across a wide field and publish them as a 'program'. There are numerous examples of the Home Office Research and Planning Unit doing this. It provides a publicly available list of all the work being funded by the agency at any given time – a snapshot of how the money is being spent. It is useful for academics in seeing where the current research interests of the agency lies and also offers some potential for public accountability. The National Institute of Justice publishes a similar document on an annual cycle.

In this report 'research programs' are seen as somewhat different and are central to the development of a coherent body of knowledge covering a particular area. A program has a purpose or goal, which may be fairly long term, and it may have intermediate targets or aims. Box 2 gives an example from the Police Research Group. Although it comprises a

list of projects, which may be carried out by a number of disconnected research agencies, they serve a common purpose and are intended to build upon each other and derive synergy thereby.

Box 2: The Police Operations Against Crime Program

A five-year program of research and development aimed at improving police expertise in operations against crime and crime detection

Following consultation with senior police representatives, phase 1 involved three components:

→ **Strategic approaches**: Which policing tactics improve police performance in relation to volume crime?
→ **Resource management**: How can the police service get the most out of the resources it has at its disposal?
→ **Skills development**: What skills are required to improve the effectiveness of police operations against crime and crime detection?

In phase 1 it was decided to concentrate on 'volume crime' – burglary and car crime – on the grounds that if the police could achieve significant reductions in these offences then more time would be freed up to concentrate on other more difficult and more time consuming crimes.

In May 1993 a prospectus was issued to every police force in England and Wales and to over 100 research and consultancy groups. Each of the three components of the program was described in terms of its aims and objectives. Bids for research to address these aims and objectives were invited from academics and consultants, but each bid had to have been discussed and agreed with at least one police agency. Through this process it was also hoped to initiate or improve working relationships between police and local research agencies.

The resulting proposals were considered and agreed by a series of committees comprising a range of practitioner and policy stakeholders and a prioritized list was developed. Twenty projects were funded at a cost of 1.6 million pounds. The results of the research were eventually published in various PRG series.

Box 2 describes the first phase of a 'proactive' program – the aims were set out at the beginning of what was expected to be a protracted and fairly complex process. Although there was no guarantee that the aims could be met, there was some confidence that useful information and an increased knowledge base would result. This 'top down' process sets only the parameters within which academics and others can bid to carry out research. It allows for creativity at the local level in how the requirements might be met. Because it is less prescriptive at the project level, there is a risk that 'gaps' might appear in the final program. These gaps can be filled in a number of ways – with a more specific follow-up prospectus or solicitation outlining the general area that remains deficient; by inviting

bids for a specific project if that is all that is required, or by carrying out the work in-house.

This last option is an important one as it has a useful spin-off in supporting in-house research, which in my view is essential to the maintenance of creative, motivated and well-trained staff within government. It is not the same as asking in-house staff what, if any, research they would like to do, which would be the equivalent of investigator initiated in-house. Rather it is offering the in-house staff the opportunity to contribute to an agreed and important policy research agenda, and providing, through that, training in how to carry out policy relevant work; increasing their expertise as project managers (they will know first hand what some of the problems of research are), and by developing their CV, fitting them better to apply for other posts outside government and thereby creating a healthy turnover of staff.

Commitment to this approach from central management obviously means that they need to ensure there are sufficient staff in-house not only to manage the budgets and external projects, but also to provide sufficient time for the in-house program to be carried out.

An alternative approach to the proactive program – what might be called a 'reactive' program – has been described by Laycock (2001), and involves the development of a series of projects over time, which are dependent upon one another; the first project prompts hypotheses which subsequent projects are intended to address. This leads to the development of a program of work with coherence and internal consistency, but it is reactive in the sense that later projects react to issues raised by their predecessors. A brief description of this program is provided in Box 3 overleaf.

Until the recent injection of significant research funds by the UK government associated with the Crime Reduction Program, the majority of Home Office funded research through the Research and Statistics Directorate comprised ad hoc projects. Policing research was organized differently through the Police Research Group, which had a different remit from the rest of the Home Office research function – to increase the influence of research

Box 3: Hypothesis based research: the repeat victimisation story

A research project, established in 1983, gave researchers the remit to 'find an area with a high burglary rate, make it go down, and tell us how you did it'. An interagency project team was brought together of academics, police, probation staff and others. Their analysis showed that there was a great deal of 'repeat victimisation' on the project estate. If a house had been burgled there was a significantly higher risk of it being burgled again than if it had not been burgled in the first place. This led them to focus on victims as a way of reducing crime. By a variety of means they protected victims and reduced repeat victimisation to zero in seven months. The burglary rate fell on the whole estate by 75% over the following three years.

A series of further research projects were commissioned to test out the repeat victimisation findings in other settings and for crimes other than burglary. These projects confirmed the importance of repeat victimisation in burglary and also showed how it featured in certain types of violent offending and bullying. The conclusions were pulled together in a review report entitled 'Once bitten, twice bitten: repeat victimisation and its implications for crime prevention' (Farrell and Pease, 1993).

The challenge then became to get the findings of this research to impact more generally on crime prevention policy and practice. In rolling out the research findings a great deal of effort was put into engaging practitioners in the repeat victimisation story. There were a series of six 'road shows' on repeat victimisation across the country. This was followed by the establishment of a 'task force': comprising a task force head and a seconded police officer. The head of this task force was a specialist in organisational development and had some marketing expertise.

The task force used a variety of means to reach practitioners:
- A repeat victimisation liaison officer was designated in each police force, whose task it was to ensure that the research was properly disseminated – in effect a local champion.
- A series of liaison officer meetings were arranged to share good practice and iron out any emerging practical difficulties in implementing strategies to tackle repeat victimisation.
- A publication designed specifically for practitioners at middle manager level was produced ('Preventing repeat victimisation: the police officers' guide').
- A computerised database of good practice was established within the PRG for use by UK police forces.

Probably the most significant action in forcing repeat victimisation onto the police agenda, however, was its adoption as one of the Home Secretary's police performance indicators for the prevention of crime. Given the sensitivity of introducing such an indicator, an incremental approach was adopted. This aimed to take the police from the point at which many of them were in 1995 of not being able to measure repeat victimisation, to being able to tackle it and deliver real results:
- 1995-1996 – demonstrate capability of identifying repeat victims
- 1996-1997 – develop a strategy to reduce repeat victimisation for any locally relevant offence.
- 1997-1998 – implement the strategy
- 1998-1999 – set targets for reduction in repeat victimisation.

By 1998 all forces claimed to be able to identify repeat victims to some degree; all but one force was able to identify repeat victims of domestic burglary, and all forces had developed a strategy to tackle such crimes.

The repeat victimisation task force was formally disbanded in 1998, although the staff involved continued the work of ensuring implementation of research results on a broader scale.

(Source: abstracted for Nutley and Davies (2000), from Laycock G. (2001), Hypothesis based research: the repeat victimisation story, *Criminal Policy Journal*, forthcoming.)

on police policy and practice – and it was there that both proactive and reactive programs were developed.

The NIJ sponsors proactive programs and publishes ad hoc project lists some of which are described later in this report as examples of good practice. The main mechanism used by NIJ for generating research proposals is through a process of 'solicitation'. Discretionary research funds are used to support an 'investigator-initiated' program, which consumes the majority of the discretionary funds available. Further, and more substantial funds, are made available through partnerships with other elements of the justice department (the COPS Office and BJA particularly), or through budget which is 'earmarked' by Congress for a specific purpose. A number of substantial research or research-related exercises fall into this category – 'Breaking the Cycle' and the major programs of work associated with the Crime Act are described briefly in Box 4 overleaf.

c. Timing

One of the consistent complaints about government funded social science research is that it is slow to deliver results. This in part arises from the way in which the research is commissioned. In the UK it was normal practice to invite research bids from policy advisors on an annual cycle. Such a cycle might typically start in June or July with an invitation to research 'customers' (ie those in the policy units) to identify their forthcoming research needs over the next 12 months. The 'contractor' (ie the in-house research team), would develop these bids during the autumn with the aim of publishing a finally agreed research 'programme' the following spring, for funding during that financial year. The fact that research was seen to be untimely is hardly surprising under such a system. There was frequently as much as a 12-month delay in starting a project, never mind producing a useful set of conclusions.

This process has its origins in the Rothschild report (1971), which was intended to control the way in which the physical sciences were developed and funded within government. The customer/contractor principle, as it became known, has dominated both physical and social science development of government-funded research for decades. It

was never intended, as Rothschild made clear in a post-script report, to apply in the same
way to the social sciences, but nobody read the post-script.

Policing research was handled differently. The PRG ran a rolling programme of work and
an open door policy on bids for research from policy units; 'call in as a problem arises'
rather than wait, maybe months, for an invitation to bid. This process did not, by any

47

means, account for the available research resources and there was, therefore, plenty left for support of the proactive and reactive strategic programs mentioned above.

In the US the process is different but there is still a built in delay caused by the way in which research funds are generated. In the Spring each year the Federal government researchers are invited to suggest research program topics which can be put to Congress through the Attorney General for funding the following financial year. Even if approved up to the Attorney General, this protracted process means that 'hot topics' identified in Spring 2000, for example, could not be approved by Congress and funded until Fall 2001. The commissioning process could then begin but in practice approval of the Federal budget is often delayed well into the financial year to which it applies. Under such a funding system it is difficult to see how significant research programs can be funded in timely fashion, although it is always possible to squeeze money from existing budgets for urgent, but small scale, ad hoc work.

The criticism of research as late does not stem solely from the funding arrangements. Once let, contracts have to be tightly managed if significant slippage is not to occur. University researchers, for example, have been used to working to their own timeframe and are consequently less 'business-like' in their approach. This is a cultural phenomenon and not easily changed, although some university departments are becoming more entrepreneurial in the way in which they bid for and carry out research there is still a long way to go. It is perhaps not surprising given the other pressures on academics' time (teaching loads in particular are high and rising in many universities), and the fact that their own publication process in academic journals is a protracted one with a wait of maybe two years or more before an article appears in print. Such a process does not encourage a sense of urgency in the completion of work.

An advantage of the major consultancy companies (although less so the not-for-profits) is that they tend to manage their own contracts tightly and are less likely to miss deadlines as a consequence. They are also less interested in the publication of the work – the project ends when the final report is produced for the customer.

In this area the major government research funding agencies have a role to play. They need to reinforce the view that missed deadlines are unacceptable unless there is a particularly good reason. The tendency in NIJ is to extend deadlines fairly freely if there are no cost implications. In the PRG, as a means of encouraging the timely production of well-drafted reports, 20% of the contract price was withheld pending receipt of an acceptable final draft. This was not always popular with the contractors, but it also had the effect of causing them to press the internal contract managers for timely feedback and efficient handling of submitted work. In effect a virtuous circle. Despite this there was still an unacceptable delay in getting final reports out to the field in the form of a publication.

d. Procurement and grant giving

There is a distinction between the procurement of research and the giving of a grant for research purposes. This distinction applies in both the US and UK. Grant giving is far more common in the US. It is an important distinction for a number of reasons and there are benefits and disadvantages to both procedures.

Looking first at procurement (which is a process preferred by Government accountants and procurement units), it is presumed that any procurement contract will include a precise description of what the product of the exercise will be. This requires the commissioning research organization to think through, in precise terms, what they wish to pay for. It also assumes that the contract will be tightly managed and that deadlines for the delivery of final and intermediate products will be preset and met by the contractor. The opportunity for management of such contracts is useful, but the terms of the contract are usually very specific, and there is, therefore, little flexibility if the data are less appropriate than was expected or if the emerging work suggests that changes in approach, timing, output or anything else, would make for a better product.

An advantage of the procurement process is that it is much easier to competitively tender for the work from a range of appropriate agencies. This is obviously more equitable,

although even under this procedure there has to be some selectivity in deciding which agencies are invited to tender in the first place. There is also far greater control over the final product, and with a higher probability of a contract being awarded to a commercial or not for profit organization, where there is relatively less interest in the publication of their final report of subsets of it, there is correspondingly far less likely to be an unexpected publication of either good or bad news.

A procurement process is particularly appropriate when commissioning a social survey, for example, where the sample size, nature of the sample, and the form in which the output is required can all be specified in advance, or when a process or outcome evaluation of an existing initiative is required. A great deal of government social 'research' expenditure falls into one of these categories. Although it would be unfair to compare the process to the normal purchase of chocolate bars or stationery, it is similar in that there is a formal contract between purchaser and provider and the arrangements are more business-like, in that sense.

A grant for research purposes is somewhat different. 'Research' is, or should be, a creative process. It may involve exploring data and testing hypotheses, the outcome of which is not know at the onset of the work. That is *inherent* in the whole notion of research. A serious disadvantage of the procurement process, in my view, is that in being so specific in specifying the schedule of work at the outset, it precludes genuine research. But there are some disadvantages to providing a grant for a research project. Once give, there is relatively little control over how the money is spent – not much chance of close project management therefore, which in some circumstances may be seen as a disadvantage. Money is granted for a specific purpose but there is no formal contract. On the positive side there is greater flexibility to change the potential product as the grant progresses, to modify time schedules and to develop the final product in the light of intermediate results. Recipients of grants can be more flexible. They tend to be based in university departments, and come with the expectation that they will publish the results in academic journals or elsewhere. The option to do this is 'part of the deal'. Grants can also be given on a single tender basis – ie without any competition, which allows for the

protection of the intellectual property rights of the grantee, but they can also be seen as unfair as a consequence, and risk creating a small group of favored, élite grantees with all the potential for abuse that such a system offers.

In practice, in the UK, research grants are being given less frequently and for relatively small amounts. In the US, grants are far more common and may be substantial, but they are given following the issue of a 'solicitation' from the National Institute of Justice, which is a widely distributed invitation to potential grantees to bid for funds in a quasi-competitive situation. 'Quasi' because they are not necessarily bidding to do the same piece of work, but rather to meet the terms of the solicitation as they interpret them. The competitive element arises because there are limited funds available and some comparative judgment has, therefore, to be made.

The UK Home Office no longer invites unsolicited research proposals on any scale, although a new scheme has just begun which is intended to encourage innovation in research and invites proposals for funds from the 'Director's Innovative Research Challenge Fund'. Although bids for funds under this heading are still intended to support the Home Office business plan, they may fall outside any current invitations to tender for research on which the Directorate is currently working. Funds for research which is of special interest to the individual researcher, and which might be called 'pure' research in some contexts, can be requested from the government funded Economic and Social Science Research Council (ESRC). There is no such equivalent in the US; the National Institute of Justice is the sole source of government funds for research of this kind, although there are far more charitable foundations in the US to which funding bids might be made. The entire shrinking discretionary budget of the Office of Research and Evaluation – the social science research arm of NIJ – is devoted to investigator-initiated research.

To summarize these points: a decision needs to be made between whether a contract should follow the procurement process, take the form of an external research grant or be carried out in-house. The decision requires consideration of the extent to which the output

of the project can be clearly stated at the beginning, compared to a more creative research process, where the final product depends upon intermediate outputs or creative input from a research team. The kinds of questions that might be addressed are outlined in Figure 2 overleaf.

Figure 2: Steps in Choosing a Contractor

prescriptive

Yes → Choose any competent agency?

No ↓

Yes → Choose a survey company?

No ↓

Yes → Choose an experienced researcher?

No ↓

Yes → Choose an auditor?

No ↓

Yes → Choose a procurement?

No ↓

Yes → Go for a research grant?

No ↓
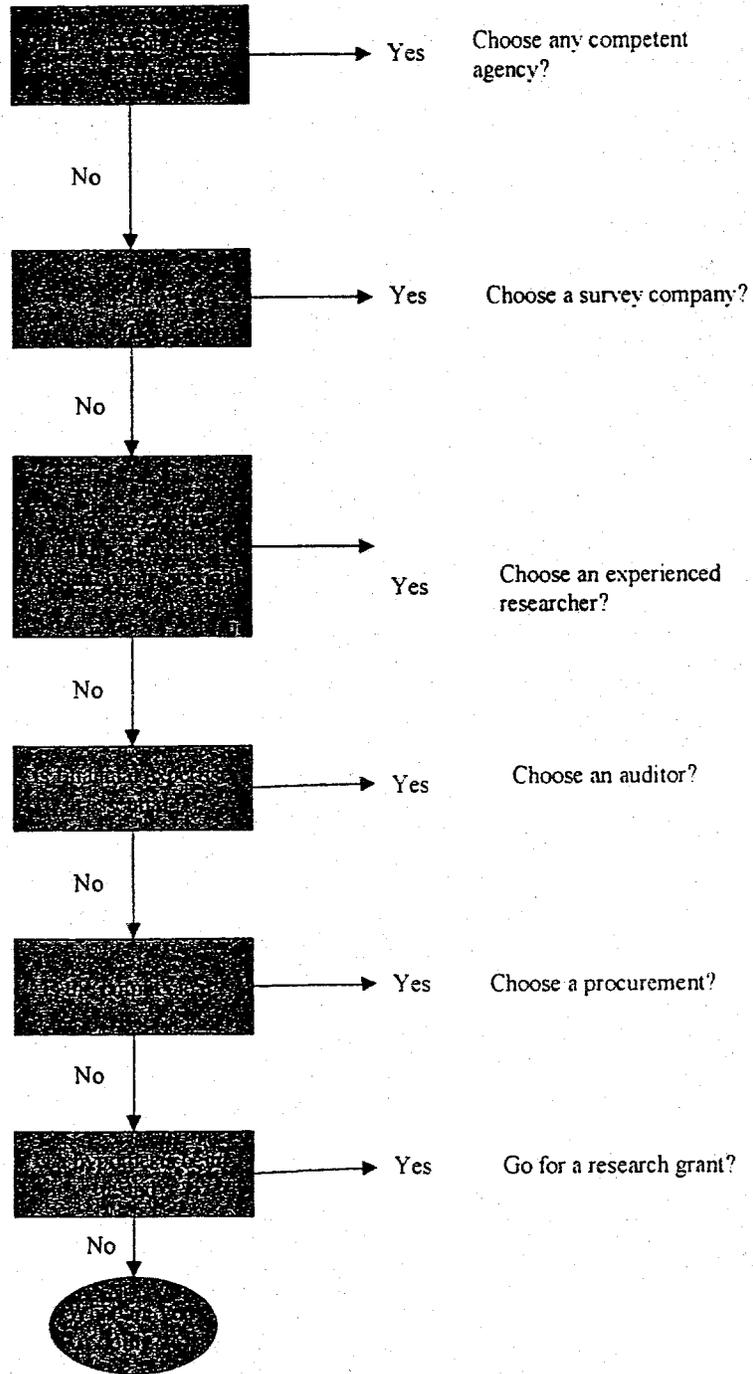
2. Carrying out research

The discussion so far has raised a number of issues with implications for the way in which research is carried out. First, the need to focus on outcomes, and particularly to address the question of what works, where, and why, has implications for research design. Secondly, there are implications for research deriving from the fact that policy makers and practitioners have a variety of more specific 'needs', as discussed in the previous chapter.

They also expect a degree of certainty in the answers they get from research, and they are looking for quick answers. Some of these expectations cannot be met, and it then becomes a matter of managing them and keeping them within reasonable limits. To some extent they arise from what is assumed to be forthcoming from the 'hard' sciences, which are far more advanced than are the 'softer' social sciences in their theoretical foundations and in the certainly with which they can express their conclusions, at least for all practical purposes. They have simply been in the game for much longer. Getting this point across to the recipients of social science research is part of managing their expectations.

In this section we first discuss what's wrong with RCTs and quasi-experimental designs ('experimental methods' for short), which have been strongly supported (Farrington, 1983; Sherman et al, 1997) as the most appropriate approach to research on criminal justice matters (although Farrington concedes that they are most appropriate when applied to experiments involving the random allocation of individuals). We then consider alternative approaches, which are scientifically acceptable but also meet the changing needs of the recipients of research. All the approaches discussed are, or are said to be, capable of assessing an initiative, project or program against an outcome measure such as reducing crime. They do not comprise an exhaustive list of research methodologies! In particular there is no consideration given to survey methodologies, descriptive statistics, or routine monitoring exercises. This is because they are generally not concerned with the identification of why an initiative might be presumed to work, or what the necessary conditions are for this to happen, but with the provision of fairly straightforward

information, which may itself be necessary for a thorough evaluation but is not sufficient. Issues concerned with integrity and ethics are also covered in this section.

a. What's wrong with 'experimental methods'?

Experimental methods have an impressive pedigree. The advancement of knowledge in the field of medicine, for example, owes much to designs of this type, where the question is the extent to which particular medical treatments may or may not 'work'. As such they are obvious candidates for transfer into other areas of social policy where similar 'what works' questions are being asked. The roots of this approach are from the hard sciences – medical researchers are trained as chemists, physicists, pharmacologists and so on. But there are some major differences between the way in which medical issues can be addressed and other social policy areas. In medicine, the treatments themselves have, traditionally, been fairly readily defined in most cases – dosages can be prescribed for instance; and there is a relatively clear outcome – the patients' conditions improved or they did not (defined as mortality, morbidity, quality of life and/or patient satisfaction). (This picture is now becoming more complex, even in the medical field.) In addition, the unit of analysis is typically an individual and it is feasible to allocate individuals to treatment, control, or placebo groups on a random basis. The sample sizes can be manipulated to reflect the expected effect, with large samples being used, for example, where treatment effects are expected to be marginal. It is also feasible to carry out more than one such randomized controlled trial (RCT) in any given research field. Furthermore there is a reasonable degree of certainty that sufficient numbers of patients will follow the treatment regime to make the trials viable.

The medical experience is often seen as the standard to which other social policy research exercises might aspire. It is, however, no accident that RCTs were most noticeably used in the medical area where random allocation of relatively large samples of individuals is feasible and implementation failure is unlikely to significantly threaten the experiment. Although it is always possible that patients may forget or fail to take their treatment for other reasons, the reasonable assumption is that the treatment was, in fact, administered to the majority in the group. In an RCT it is also feasible to randomly assign as many

individuals as statistically necessary to the treatment and control groups. And it is sensible to test the hypothesis that treatment X, a particular dosage of medication, will be to some degree effective for all patients. While there is assumed to be individual variation in the appropriateness of the treatment, the assumption is that it will be more effective than not on aggregate. In other words there will be a relatively small number of patients for whom it may be marginally less effective but they will be sufficiently small in number not to disproportionately affect the outcome when compared with the untreated control group. Where the delivery of the treatment is heavily dependent upon the skills of an individual, in surgery, or psychotherapy, for example, the experimental procedure can be more complex, and larger samples may be required.

It is probably also fair to say that in many if not most of the early clinical trials, researchers were comfortable with the notion that the mechanism through which the treatment was expected to exert its effect was unknown, and it was not necessary that it should be known in order for the experiment to be carried out 'successfully'. So, for example, the efficacy of penicillin could be tested using randomized controlled trials without any real understanding of *how* it was delivering its effect. As knowledge develops in relation to such mechanisms it becomes possible to test efficacy in a more sensitive manner, hypothesizing, for example, that the effects may be greater for some sub-sets of the population than for others, and since samples can be relatively large it is also possible to test for interaction effects in the course of a standard experimental trial.

Perhaps the area of public policy most closely aligned to health is education, where many of the issues relate to the school performance of children. Here classes of children can be randomly allocated to various treatment or control options but the treatment effect generally depends on the skills of an individual teacher. Testing for the effect of class size on attainment, for example, thus becomes complicated by the

> "The traditional assumption is that, if the treatment is properly planned, then a coherent and proper implementation will follow. This assumption may have to change."
>
> Huey-Tsyh Chen (1990:33)

teachers' abilities, and indeed the broader school environment within which the class is located. While these effects can often be dealt with through various statistical manipulations of data, they are more conceptually complex as a result.

Whilst RCTs might, to varying degrees, be an appropriate research model in the medical arena and in some other areas of public policy, there are some major differences in the crime prevention/policing fields, which raise serious doubts about their usefulness there. First, the unit of analysis is often not an individual; it may be a store, community, housing area or parking lot. Random allocation of a sufficiently large sample of stores, communities etc to a treatment and control group is, generally, quite impractical.

Secondly, RTCs are 'black box' experiments where the mechanism through which the effect is taking place is not necessarily known, so even if we were able to randomly allocate a large sample of communities to treatment and control groups, such allocation would not necessarily provide any information on *why* the initiative may have been successful. The outcome of the experiment is, therefore, along the lines that treatment X may have worked, but we do not know how or why. A conclusion of this kind is not what is required at this stage in the development of knowledge about policing and crime prevention.

Thirdly RTC assume that there is relatively little risk of implementation failure, or if it does happen it is in sufficiently small a proportion of the experimental population as to be irrelevant. In policing and crime prevention at the community level, implementation failure is a significant possibility and major efforts have to be made to ensure against it (Laycock and Tilley, 1995).

> "Experience has led to significant methodological shifts to accommodate the needs of social work practice"
>
> "Attempts to fit practice into a particular methodology may distort practice."
>
> Mansoor Kazi (1997:419, 421)

Finally, the experimental treatment has to be maintained throughout the experiment, as originally intended; learning from experience, or making ongoing improvements or adjustments is not permitted. While these assumptions and constraints are plausible in the medical field, with

sufficiently large samples, they are not tenable in communities where there is far less control. Crime moves, evolves and develops with the opportunity structures offered by the immediate situation within which the potential offender finds him or herself and which are beyond experimental control. For all these reasons, then, true experimental designs of the kind ideally deployed in the medical field are generally neither practical nor desirable in the real world of policing. There are far too many variables to even pretend that they can be controlled.

Quasi-experimental designs, where there is no attempt at random allocation, are a rather different matter, although even here the reality necessarily falls somewhat short of the ideal. These designs require a control area, and possibly a displacement area, with which to compare crime rates in the experimental area, in which the innovation under investigation is taking place. The control area is meant to be comparable to the experimental area in social composition etc but is expected to remain free of any innovation or other attempts to affect the dependent variable (the crime rate). The displacement area, usually adjacent to the experimental zone, is meant to test for the possibility that instead of reducing crime, the experimental effect was merely to replace it (Barr and Pease, 1990). Crime figures over a reasonably long period are collected before and after the experimental intervention, in all three areas.

There are some conceptual difficulties with such an approach. First, it is assumed that nothing other than the experimental treatment will affect the dependent variable in the experimental area, and nothing at all will happen in the control or displacement areas that might affect the dependent variable, or if anything does affect the dependent variable it does so in all three areas (experimental, control and displacement) in equal measure. This is an almost impossible condition to guarantee in most real world experiments in communities where there are all sorts of changes and dynamics operating.

It also assumes that the crime rates in all three areas – experimental, control and displacement, are reasonably stable over time. Too much random fluctuation, or 'noise' would make any empirical investigation questionable. In reality, the relatively small

areas, which are usually subject to experimental investigation, are susceptible to random fluctuation and are very difficult to control in the sense of ensuring that no other relevant activity is spontaneously occurring at the same time as the experimental investigation. Thus it is not that unusual to find that crime has reduced more in the so-called 'control' area, for reasons that may well be beyond the experimenter to explain. And, as with true experimental designs, the experimental treatment has to be maintained in its original form throughout the exercise and not modified in the light of experience.

To make matters worse, the notion of 'diffusion of benefits' has been developed in recent years (Clarke and Wiesburd, 1994). Instead of crime moving from the experimental area to the adjacent 'displacement' area, the benefits of the experimental intervention may spread into the displacement area and crime reduces there too. It is possible to offer some plausible speculations as to why this may happen, but the fact that it is a possibility produces problems for quasi-experimental designs.

Finally, the kinds of interventions which practitioners are now claiming might be effective in reducing crime or in improving policing, are often multi-facetted. They suggest that a package of interventions is required in order to turn around disadvantaged communities. Approaches of this kind are particularly difficult to evaluate no matter what the approach, since it is not possible to determine which, if any, of the experimental interventions produced the observed changes, but they are certainly not susceptible to evaluation using quasi-experimental designs.

The criticisms of experimental and quasi-experimental designs have not, so far, been directly related to the extent to which they can provide the kind of information that is required by policy advisers and practitioners. They fail on a number of criteria in this area also. These designs are expensive and time consuming but they are also unrealistic in expecting the control and displacement areas to remain intervention free, and in expecting practitioners to ignore the incoming feedback as an intervention begins to take effect, which may suggest the need for changes to the project. The last thing an experimenter wants is a change in the nature of the intervention mid-evaluation. Finally,

by their very nature they are not able to explain why an initiative might be working o what the context sensitive dimensions of it might be. All these issues go some way toward explaining the relative rarity of such experimental projects in the criminal justice field, and their inappropriateness.

This is a fairly depressing picture for those aspiring to evidence based policy and practice. The picture is not, however, as gloomy as it may appear. Over the past few years a number of alternative approaches have been developed in other disciplines, which are transferable to policing, and new approaches have also been tried within the criminal justice field itself as discussed below.

b. Alternative approaches

There is a huge range of alternatives to experimental methods. It is beyond the scope of this report to go into the detail of how each of these might be used to develop and evaluate programs and projects; whole books have been written on each approach. There are, however, a few recently developed techniques, which seem particularly well suited to the kinds of evaluations now required. They have certain features in common and it is these commonalities which are discussed in this section and which are important in offering alternatives to experimental and quasi-experimental design, which also better meet the needs of the policy and practitioner groups.

Before moving to a discussion of these approaches and their common features, a number of other approaches are noted, which are not seen as alternatives to experimental or quasi-experimental designs, but are either complementary or orthogonal to them. For example, for many years an unhelpful debate raged (the right word) between advocates of qualitative and quantitative evaluation. Qualitative approaches were indeed seen as an alternative to quantitative designs and many of the discussions and debates generated more emotion than information about the relative merits of each approach. The early work of Yin (1984), for example, was an attempt to rehabilitate qualitative research, which suffered in the 1970s and 80s for being seen as subjective and lacking scientific rigor. But as Cook more recently says:

"The case for qualitative methods does not depend on attacking the foundations of quantitative methods; it rests on their utility for answering important evaluation questions either when used alone or when used together with quantitative methods." (1997:33)

The debate over qualitative/quantitative methods has moved on from 'either/or' to 'perhaps both', although qualitative approaches remain vulnerable to the label 'subjective' because of the judgment that is necessarily required of the evaluator in deciding whom to interview or to quote, and how to present the evidence. As such they are particularly susceptible to criticism, and/or to being ignored, by policy advisers who may feel that the results obtained, or the message derived from a qualitative analysis does not accord with their (politically sensitive) view of the world. Ethnographic studies are particularly vulnerable to such criticism because the researchers necessarily operate very closely with the communities or individuals that they are describing. Questionnaire surveys fare rather better since they use standard techniques for gathering and analyzing data, which facilitates replication. All these approaches call for judgments on the part of the investigator, however, which makes it easy for an unhappy policy adviser or practitioner, who does not like the research conclusions, to criticize them. The fact that similar judgments are made in the quantitative field is often overlooked – their numeric content can give them a spurious authority.

When qualitative and quantitative methods are combined they can be particularly powerful in making the results 'real'. Endless tables of data, while no doubt informative and accurate, do not get across what can be the painful reality of a criminal experience for the victim, or the social vulnerability of the offender. Lloyd et al (1994) made good use of qualitative data in presenting a study of domestic violence in Liverpool, in the UK. The report provided support for the use of rapid response alarms, amongst other things, but used a series of descriptions of the victims' experience as a means of illustrating the offences to which they had been subjected and the fear, for themselves and their children, which accompanied the experience. Journalists use the 'human angle' provided by these descriptions in reporting upon research results, in preference to tables of data or statistical

analyses. As such they can be helpful in 'marketing' research and ensuring that it is reported. Qualitative approaches are particularly useful in helping to determine whether initiatives were implemented as planned and in providing additional information on how people feel about a project, which may be important in deciding on its public acceptability or the extent to which it may have contributed to a reduction in the fear of crime. The emphasis on outcome effects does suggest, however, that some quantitative measures need to be taken.

An example of an approach, which could be seen as orthogonal to experimental and quasi-experimental design, is meta-analysis – the systematic analysis of the results of a body of evaluations of similar programs to produce an estimate of overall program or treatment effect (Weiss, 1998a). The Cochrane Collaboration makes extensive use of such analyses in determining the effect of various medical treatments. Although the Cochrane Collaboration uses RCTs as the input to the meta-analyses, it is not necessarily the case that RCTs or quasi-experiments should be used. A meta-analyst would be well advised to adopt a clear criterion or set of criteria in deciding upon which evaluations to include an which to exclude from any given meta-analysis, but insofar as any reasonable criteria could be adopted, meta-analyses are orthogonal to experimental designs they are not an alternative to them.

A more qualitatively oriented alternative to meta-analysis is cluster evaluation, involving the evaluation of a program with projects in multiple sites aimed at bringing about a common general change. Cluster evaluation addresses the following questions:

- ❏ Overall have changes occurred in the desired direction? What is the nature of these changes?
- ❏ In what types of settings have what types of changes occurred, and why?
- ❏ Are there insights to be drawn from the program failures and successes that should inform future initiatives?
- ❏ What is needed to sustain desired changes? (Sanders, 1997:397)

Cluster evaluation differs markedly from meta-analysis in the degree of involvement of the evaluator with the program. In meta-analysis the evaluator may review the published literature for the material of his or her analysis, whereas in cluster evaluation the evaluator is part of the program team, which includes the program directors and project staff. Although each project comprising the program may operate differently in a manner appropriate to the context within which it is located, the overall goals of the program are agreed across sites and the cluster evaluation covers all the sites.

There are a number of other approaches, which might be seen as viable alternatives to experimental methods, but which do not quite fit the bill in that they are more 'mechanistic' – carrying out evaluations in a formulaic manner. The work of an audit department might fall into this category, an external consultancy company called in to assess performance or an inspectorate (in the UK there is a police inspectorate at national level which is required to inspect police forces on a regular basis and ensure that they are giving good value for money and employing good practices). Some of the cost-benefit or cost-effectiveness evaluations also tend to be formulaic in approach (see, for example, Dhiri and Brand, 1999 or Colledge et al, 1999).

There is also a set of evaluation styles, which involve the program participants much more directly, indeed in some cases, centrally. These are called 'self-evaluation', 'empowerment evaluation', 'collaborative evaluation', 'stakeholder evaluation', 'project monitoring', and so on. They vary in the extent to which professional evaluators are involved and in the discretion allowed to the program participants. I prefer to see them as various examples of good management practice, either at the instigation of the project staff themselves or as a requirement of external funders. They vary in their cost to the program staff, with the least expensive, and arguably most appropriate, at the project-monitoring end. There have been a number of manuals and guidebooks produced to assist local project managers to assess their own efforts, often linked to federal or central government funding. So, for example, the Bureau of Justice Assistance (BJA) at the US Department of Justice has produced guidance to its grantees in assessing their efforts.

which is available on the BJA website. In the UK, Crime Concern, which is a charity providing guidance to communities in the development and implementation of crime reduction programs, provides similar guidance and training to project staff concerned to monitor the effects of their initiatives.

These and other approaches, while usefully complementing quantitative work, are not sufficient to assess the effect of a program against an outcome target such as crime reduction, either based on police crime recording systems or on crime surveys. Some of these assessment methods are set out in Table 2 below, where they are divided into those that interact closely with the project team and those that tend to be more remote. Most of these approaches are described in greater detail by Weiss (1998a), where they are critically reviewed.

From the range of possible alternatives to experimental method three are considered in detail below. These are the 'theory of change approach', which derives from the US evaluation literature, 'scientific realism' from the UK, and 'rival explanations', which it

Table 2: Types of assessment

	Interactive with project	Remote from project
Quantitative evaluation	Theory driven evaluation Theory of change approach Realistic evaluation Rival explanations	Randomized controlled trials Quasi-experimental designs
Good management practice	Self evaluation Empowerment evaluation Stakeholder evaluation Project monitoring	Audit Cost benefit/cost effectiveness analyses Consultancy exercises Inspectorate functions
Qualitative evaluation	Ethnographic studies	Questionnaire surveys
Cumulative evaluations	Cluster evaluation	Meta-analyses Cost-benefit/cost effectiveness evaluations

will be agued is a useful adjunct to both. These are quantitative alternatives to experimental methods but are more interactive with the project staff. There is first a brief description of each approach given below, followed by a discussion of their common features and the extent to which they are an improvement, in being more sympathetic to

64

the constraints of the 'real world' and also delivering what is required by the policy advisers and practitioners, over both the traditional approaches and the other possibilities noted in Table 2.

Theory of Change Approach (Weiss, 1998; Fulbright-Anderson et al, 1998)

This approach has a fairly long history; indeed it could be described as experimental psychology, which I was taught as a matter of routine in the 1960s. In recent years it has taken on a new label and has been adopted beyond the psychological specialism as an approach to evaluation more widely. A complex discussion of 'theory-driven evaluations' is provided by Chen (1990), which draws on earlier work by Chen and Rossi (1987) and which has much in common with the theory of change approach, which is more simply described by Weiss (1972, 1998a) and later by Fulbright-Anderson et al, (1998).

The impetus for the present incarnation of this approach derives from the need to evaluate comprehensive community initiatives (CCIs), which the traditional experimental approaches are ill designed to do (Kubish et al, 1995). CCIs are difficult to evaluate because they are complex, they are trying to solve a number of possibly related problems at once; they are flexible, and evolve as the problems themselves change, and there are no control groups with which they can be sensibly compared.

In order for these complex interventions to be evaluated, the program manager or sponsor has to be specific about how and why the program, in its various aspects, might work (Weiss, 1995). The evaluator, with help from the program staff, has to be able to articulate the theories, assumptions and sub-assumptions that might be built into the program. It is these that form the hypothesis or set of hypotheses, which the evaluation will go on to test. Some of these theories can be complicated and inter-related, making firm conclusions and generalization difficult. In some cases the practitioner may not be clear about how or why the program is expected to work, and they may need help in working through the possible alternatives. They may just *know* it is the right approach and resent the analytic approach taken by evaluators in pressing them to work through the underlying theories. There may also be some disagreement between practitioners or

between them and the program sponsors. Finally there is a possibility that being explicit about just how something might have an effect may be politically problematic. For instance, the mechanism may be interpreted as divisive of the community, or may appear to favor one sub-group over another. Whilst such possibilities can be glossed over when the program assumptions are not

> "What if interventions that change only one thing at a time fail ... *because* they change only one thing at a time?"
>
> Lisbeth Schorr (1997)

brought out, it is much more difficult to do this when they are on the table for all, including the media, to see.

If, on the basis of a number of specific theories, a program appears to have worked, how far can we generalize to other communities in what might be quite different areas? The fact that the theory has been set out is helpful in answering this question, and the extent to which it is capable of generalization depends upon the theory itself. If, for example, it relates to some very specific attribute of the particular community, or to the skills of inter-personal relationships of central players there, then the scope for generalization may be less. If generalizations are going to be made, there will almost certainly be a need to replicate the results in other areas. In order to do this, the theory has to be clearly stated. Replications are notoriously susceptible to replicating the wrong thing (Tilley, 1993).

Let us take a specific example from the US literature, to which reference has already been made. Weiss (1998a) describes the program theory behind the proposal that higher teacher pay may increase student achievement. There are a number of possibilities why this may work as set out in Figure 3 overleaf. To quote Weiss directly

"program theory, ...refers to the *mechanisms* that mediate between the delivery (and receipt) of the program and the emergence of the outcomes of interest. The operative mechanism of change isn't the program activities per se but the response that the activities generate." (Weiss, 1998a: 57, original emphasis)

Her decision to emphasize the word 'mechanisms' is significant, as we move on to look at scientific realism.

Figure 3: Mechanisms linking teacher pay to student achievement

Teacher salaries increase

Teacher morale improves ⟶ Classroom social climate becomes more pleasant

Teachers give up their second jobs and put full energy into teaching

Abler people are attracted into teaching

Teachers work harder at teaching and put more effort into preparation and pedagogy

Teachers develop more congenial relationships with students

Teachers more thoroughly prepare their lessons

School districts hire more able teachers

Students' understanding of material improves

Students seek to maintain good relations with their teachers

Teachers employ a greater variety of pedagogical strategies

Abler teachers Teach more effectively

Students work harder

Teachers teach more effectively

Increased student achievement

Scientific realism (Pawson and Tilley, 1997a)

Pawson and Tilley have fully described scientific realism in their book 'Realistic Evaluation' (1997a). It makes use of four concepts – embeddedness, mechanisms, contexts and outcomes. 'Embeddedness' refers to the wider social system within which all human action takes place. Much of it is taken for granted – it is implicit. As an example they give "the act of signing a check is accepted routinely as payment, only because we take for granted its place within the social organization known as the banking system. If we think of this in causal language, it means that causal powers do not reside in particular objects (checks) or individuals (cashiers), but in the social relations and organizational structures they form." (Pawson and Tilley, 1997b:406). This notion is important to the understanding of precisely what a 'program' comprises. It is not, simply, a target hardening initiative or a block watch program, "a program *is* its personnel, its place, its past and its prospects".

When turning to *how* a program might exert its effect, Pawson and Tilley use the notion of 'mechanisms'. It is directly analogous to the notion of program theory as described by Weiss and others. Describing the mechanism means going inside the black box. As Pawson and Tilley say, we can never understand how a clock works by looking at its hands and face, rather we need to go inside the works – and understand the mechanism. It is through the process of understanding, or hypothesizing about mechanisms, that we move from evaluating whether a program works or not, to understanding what it is about a program that makes it work.

But whether it 'works' or not will also depend on the 'context' within which it is introduced. Solutions to crime problems are often context dependent. What might work in one place could be disastrous or prohibitively expensive in another.

As an example, let us look more closely at block watch. If block watch were to reduce crime how would it do so? What is the 'mechanism'? And how is it related to the context within which it may be introduced? The bottom line seems to be that residents agree to 'watch out' in their neighbourhood, and call the police if they see anything unusual,

particularly an offence in progress. Another way in which it may work would be, through the window decals, to alert would be offenders to the fact that they are entering an areas where residents' care about their neighborhood (whether they do or not) – so offenders better look out! There may, of course, be other reasons for block watch with other mechanisms coming into play – perhaps it is going to improve police/public relations for example, but that would be a different outcome measure. Let us stick with crime reduction for now.

So we have two hypothesized 'mechanisms' and both may be correct and contribute to any observed effect. We also have to ask ourselves to what extent the mechanisms behind block watch may be operating anyway, whether or not we introduce the scheme in a particular area. For example, in low crime middle class areas if residents see a crime in progress they already call the police. So what would we be testing? - perhaps the marginal effect of the window decals and street signs. In this case we are saying that the mechanism is context dependent.

> "Some developing countries are discovering that after all their efforts to privatize public utilities, or state-owned enterprises, performance has not improved..... Their expectations had been based on a simple *non-sequitur*. What brings about efficiency is not public or private ownership but competition .... The private sector is not in the business of increasing competition"
>
> Eduardo Wiesner (1997:191)

So the evaluation of block watch has now become quite complex. We may need a high crime area, where residents would not normally call the police, and we may need to do some very specific things to make sure they feel comfortable with doing that – providing telephones for example, and some degree of protection against bullying or threats. We also need to take note of the fact that in areas of this kind it is probably the neighbors that are burgling each other, so looking out for strangers may not be so important as providing a socially acceptable and safe way for residents to call the police. We may also need a low crime area where we are fairly confident that residents do call the police, but we are interested in the effect on burglary of street signs and window stickers – ie publicizing the fact to the would-be offender that the neighborhood is cohesive and working against crime. And we may be interested in the effect of publicity alone in high crime areas regardless of whether the community members sign up to block watch. By setting up a set

of interrelated projects, each with a specified mechanism in different contexts, we can begin to see how a body of knowledge can be built up. The picture is set out in Table 3.

Table 3: mechanisms operating in block watch crime reduction program

Context	Mechanism	
	Call police when crime in progress or when seen a suspicious stranger	Window/door decals Street signs
Low crime area	Already operating	New measure introduced
High crime area 1	Special measures taken by program staff to support this	No decals or street signs
High crime area 2	No measures taken to support this	Decals and street signs introduced in the area
High crime area 3	Special measures taken by program staff to support this	Decals and street signs introduced in the area

The mechanism, then, is hypothesized to have a particular effect, in a given context, which will lead to an observed 'outcome', in our current discussion, a reduction in crime. This can be summarized as Outcome = mechanism + context, or in prose:

"The basic task of social enquiry is to explain interesting, puzzling, socially significant outcomes (O). Explanation takes the form of positing some underlying mechanism (M) that generates the regularity and thus consists of propositions about how the interplay between structure and agency has constituted the outcome. Within realistic investigation there is also investigation of how the workings of such mechanisms are contingent and conditional, and thus only fired in particular local, historical, or institutional contexts." (Pawson and Tilley, 1997,b:412)

By operating in this way, and being prepared to replicate our results in other areas, a body of knowledge can be built up which enables the better understanding of the various methods of crime control.

Rival Explanations (Yin, 1999)

Yin presents his ideas for 'Rival Explanations' as a methodology for evaluating complex community-based programs (Yin, 1999, 2000), although they are equally appropriate for testing more modest community or policing initiatives. Yin describes the features of CCIs as typically involving:

70

- ☐ Systemic change (eg in the norms, infrastructure or service delivery of an agency or set of agencies), not just change in individual behavior

- ☐ Multi-faceted intervention, not just a single variable

- ☐ Unlikely to be standard across sites

- ☐ Multiple, not single outcomes of interest

- ☐ Idiosyncrasies of communities reduce validity of defining 'comparison' sites

In a nutshell, Yin's argument is that a complex set of activities are introduced into a community; there is a desire to attribute subsequent change to those activities (ie to say that they caused them); the threat to this attribution is that some other event or set of events may have been the cause; therefore, let us test plausible rival hypotheses or explanations for the observed changes and see if we can eliminate them. He likens this approach to that or the journalist, detective, forensic scientist or astronomer.

Yin has some persuasive points in support of his approach over experimental methods. He notes that traditional experimental designs try to rule out rival explanations through randomization or the use of control groups, without naming them. Rival explanations methods rule out only those rivals that are named and tested for, but the process is, therefore, more

> "More and more I have come to the conclusion that the core of the scientific method is not experimentation per se, but rather the strategy connoted by the phrase, 'plausible rival hypothesis'."
>
> Donald Campbell in foreword to Yin, 1984

transparent. His main argument, however, is that this approach has greater applicability in that it can address more complicated social change, including CCIs.

The steps are fairly straightforward and are set out in Figure 4 (overleaf).

Yin lists a number of possible rival explanations, which may apply in a fairly wide range of community-based evaluations. These are:

- ☐ Direct rivals in practice or policy – another intervention which was introduced at the same time *caused* the change

- ☐ Commingled rival – another intervention *contributed* to the change
- ☐ Implementation rival – the initiative was *not implemented* properly and could not have caused the change
- ☐ Rival theory – the change was caused by the intervention but for *other theoretical reasons* than those hypothesized
- ☐ Super rival – some new innovation, but external to the project and applying more generally caused the change
- ☐ Societal rival – things are changing anyway

All these various options may need to be considered in any given evaluation and the program evaluator can never, in theory, rule out every alternative potential cause of change. But by making the alternatives explicit, and collecting data relevant to testing them, the process becomes transparent, challengeable and valid as a result.

Figure 4: Steps in Using Rival Hypotheses
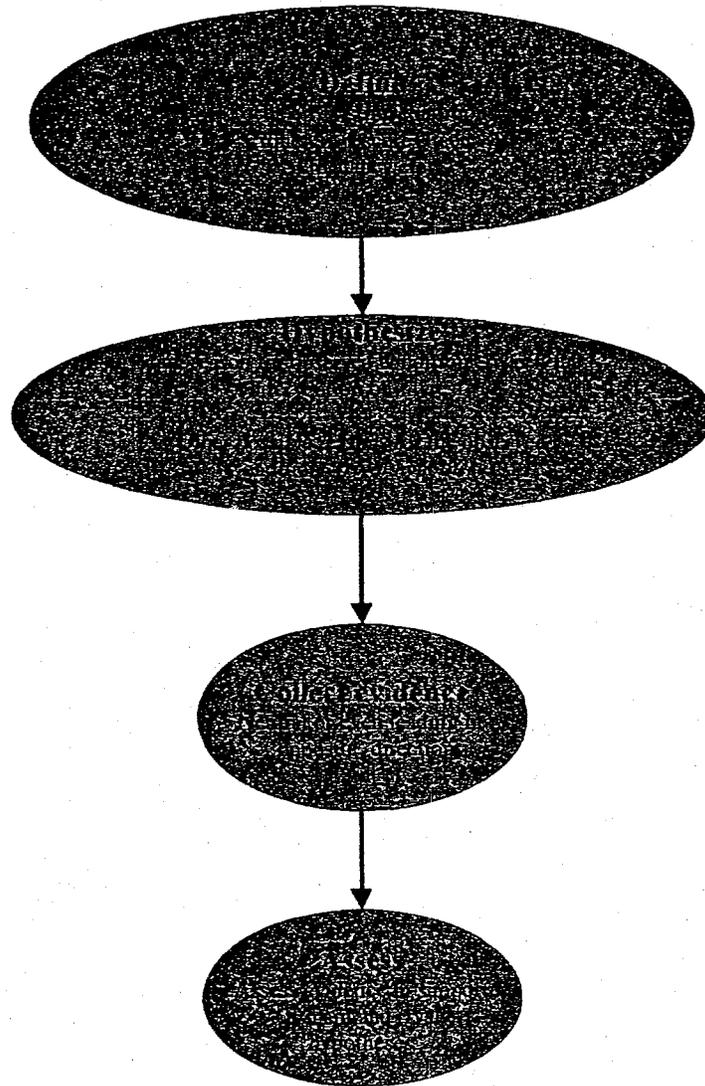
(taken from Yin, 2000).

Commonalities and differences

There are some common features to these approaches, which make them attractive. First they unpack the 'black box'. They require a statement of the causal mechanism between what is being done and what is expected to change. Weiss and her colleagues call this 'program theory'; Pawson and Tilley call it 'mechanisms', and Yin calls it an 'explanation'. It is one of the most important improvements on the traditional experimental methods approach. And it is essential if we are ever to build up a body of knowledge about what works, where and why. It also has other advantages. If, for example, the causal chain is long, or tenuous, then this might cast doubt on the likelihood of the intervention generating a measurable effect, or may lead policy advisers to questions its cost-effectiveness.

'Program theory' and 'rival explanations' approaches have relatively little to say directly about 'context' which is given considerable weight in the discussion of 'scientific realism'. This concept is important in surfacing the relationship between *what* is being done in a program, which may cause the hypothesized effect, and the *circumstances* or social context within which it is being done. These relationships take on an increased significance when the question of generalizability is raised. It then become important to know the extent to which the particular characteristics of the project site were essential to the salience of the mechanism. And this issue becomes event more relevant with the increasing interest of policy advisers in cost-benefit or cost-effectiveness analysis.

To take the Perry Preschool Program (Barnett, 1996) as a concrete example, this involved exposing a group of youngsters at high risk of school failure, involvement in crime and delinquency and so on, in a high quality intensive preschool support program, which included their parents and teachers. The project was highly successful in achieving a number of aims including less involvement in delinquency, better school performance and reduced teenage pregnancy. The cost-benefit ratio is quoted as 1:7 – for every dollar spent 7 were saved. So does that mean that the Perry Preschool Program should be implemented nation-wide? Would the same benefits accrue? The answer is no. The Perry preschool Program was introduced in a high crime disadvantaged area of Chicago. The

context was right for such a program and the potential payoff was high. Introducing a similar scheme in Bethesda, or Hampstead, both upper middle class areas where the involvement of young people in crime is already low, and teenage pregnancies are not a major social concern, would not be cost effective. There would be too many false positives in the population – children who were not going to get into trouble whether they were in the (expensive) program or not. Looking at the mechanism through which the Perry Preschool Program was supposed to be effective, and realizing that in many other contexts those mechanisms, or better, are already in place, leads to the realization that mechanisms and context are inter-related.

A statement about the mechanism or program theory, combined with a discussion of the relevant context, should make it much easier to replicate effective programs. Replication is important in trying to tease out which mechanism, or mechanisms may have delivered any observed change, but it is also important, once a causal chain is established, to ensure that practitioners appreciate the links. Otherwise there is a risk that will think they are replicating a successful initiative when in fact they are implementing some aspect of the initiative, which is irrelevant to the success of the original scheme. The issue of effective replication is discussed more fully by Tilley (1993).

The rival explanations approach stands out from program theory and scientific realism in offering an alternative to the 'control group' used in traditional experimental paradigms. The program theorists note the fact that it is almost impossible to find a control community with which to compare the CCIs, but they do not say how they cope with the criticism that any observed changes in their program area could have been caused by something that was operating outside the program, perhaps to the whole city or state within which the community was located. Yin's approach to this is elegant in its simplicity. To take a straightforward example from policing, if we introduce a crime reduction project in a public housing complex, and we find a reduction in crime, then we may choose to look at crime rates in the whole police precinct, or the neighboring areas, not because they are seen as 'control' areas with the implication that they are somehow comparable, but because we are testing the rival hypothesis that crime was reducing

generally and it is important to discount that possibility if we are to be accepted as correct in attributing the reductions on the project estate to our initiative.

In various combinations these three approaches constitute a viable alternative to the experimental methods traditionally adopted by social science researchers. The extent to which they, and some of the other approaches discussed, meet some of the criteria relevant to delivering what is required of social science research methods is summarized in table 4 below.

Table 4: Research methods and delivery of results

Approach	Theory based	Timely results	Real world sensitive	User friendly	Internal validity	External validity
Audit/inspection	X	✓✓	X	✓	X	X
Cluster evaluation	?	✓	✓	✓	?	?
Empowerment evaluation	✓	✓	✓	✓✓	✓	X
Meta-evaluation	X	X	✓	✓	✓	✓
Monitoring/assessment	X	✓✓	✓	✓✓	X	X
Program theory	✓	✓	✓	✓	✓	?
Quasi-experimental designs	X	✓	X	X	✓✓	X
Randomized trials	X	X	X	X	✓✓	X
Rival hypotheses	✓	✓	✓	✓	✓	?
Scientific realism	✓	✓	✓	✓	✓	?

Table 4 requires some explanation. Working across the categories at the top of the table – *'Theory based'* means that the evaluative approach requires the description of a theory or mechanism through which change will be achieved. Audits and inspections are not generally explicit on this point, and cluster evaluations may or may not be – hence the question mark. Those carrying out empowerment evaluations are expected to specify the mechanism but meta-evaluations are rather more in the black box mode, where they might typically conclude a review of published work with the conclusion that x appears to work but be unable to say why, or to tease out where it might be most effective, depending upon the types of evaluations that have been included in the meta-evaluation. Monitoring and assessment exercises would not, typically, involve a statement of the mechanism because they tend to collect information on process rather than outcome, and practitioners generally have a poor record of articulating the theory behind their

initiatives (Read and Tilley, 2000). Program theory and scientific realist approaches clearly do state the mechanism but experimental designs do not.

The production of *'timely results'* clearly depends upon the scale of the evaluation. Data can only be collected as fast as it becomes available. If a program requires the determination of reconviction rates following release from prison, for example, then there may be a built in delay of some years. Audits and inspections, partly because they tend to take a formulaic approach and to make use of existing data, can be done particularly quickly, and monitoring and assessment exercises tend to be within the control of the practitioner, and to make use of real time data, which means that they, also, are less vulnerable to criticism on the grounds of late delivery. Meta-evaluations, on the other hand can be very long in their delivery because they have to wait until a whole series of other evaluations have been done before they can summarize them. I have also classified RCTs as lacking timeliness because the emphasis on the sampling frame, and getting that right, can lead to long delays in establishing the project.

The extent to which project evaluations are *'real world sensitive'* is an assessment of the extent to which they may make unreasonable or simply undeliverable demands on the practitioners or policy advisers with an interest. For example, the requirements for random allocation can sometimes be quite unacceptable politically, and the need for a program to be 'set in concrete' and not to change, develop or evolve in the interests of the evaluators, is in many cases impractical and unrealistic. Experimental methods come out badly here but so too do audits and inspections, which take a more remote, clinical approach and may not always pay sufficient attention to local realities.

The extent to which research methodologies are *'user friendly'* does, to some extent, depend on how the results are written up and presented to policy advisers and practitioners. It seems reasonable to assume that empowerment evaluation and monitoring or assessment, should score highly on this since the practitioners themselves carry them out. Audits and inspections are written with the media and practitioners in mind, and therefore tend to be in plain English. The experimental methods are at greatest

risk of being less well understood by non-specialists since they can often involve fairly sophisticated sampling strategies or statistical analyses.

The *'internal validity'* relates to the extent to which any changes in the dependent variable ie the thing the program is intended to change (in our case the rate of crime or disorder), can be unambiguously attributed to the program or project activity. It is here that the experimental designs come to the fore. They are specifically intended to cope with threats to internal validity. Audit and inspections do not do well here, nor does monitoring and assessment. The extent to which cluster analysis fits the bill depends upon the specific characteristics of each evaluation at project level. Most other approaches are acceptable, particularly if combined with the rival hypotheses approach.

The *'external validity'* is an assessment of the extent to which the evaluation findings can be generalized to other places, times or populations. Most approaches do not do too well on this criterion, which serves to illustrate the importance of replication as a means of establishing principles of effectiveness, and their relationship with context or population.

Table 4 covers a selection of the kinds of criteria that practitioners or policy advisers might be interested in when considering the various types of evaluation but it does not cover all of their interests nor does it work systematically through their 'needs' more generally, as were identified in Chapter 2. This topic is returned to, therefore, in the final section of this chapter where we look at the whole research process, including methodological issues, and the extent to which they are delivering what is required.

In this section experimental methods have fared rather badly in comparison with other approaches for dealing with 'real world' problems. But one class of criticism of alternatives to experimental methods is that their results are more vulnerable to manipulation or misinterpretation, for a variety of reasons, than are those of the experimental methods school. It is to these issues of integrity and ethics that we now turn.

d. Integrity and ethics

Rather like evaluation methodology, there is a long and philosophically complex literature on integrity and ethics in research. It is beyond the scope of this report to go into the detail of this literature, but it is sensible to address one aspect of the debate, and that is the extent to which the kinds of methods discussed above as appropriate to the evaluation of projects and programs in the real world are disproportionately vulnerable to unethical practices. One of the major criticisms (generally from the advocates of experimental methods) of evaluation approaches bringing researchers and practitioners closer, is that the researcher is vulnerable to 'pressures' as a result, and these pressures threaten the integrity of the work. Its objectivity comes into question. There are a number of points to be made about this.

First, it is presumptuous of experimental methodologist to assume that he or she holds the ethical high ground. There are numerous cases where results have been manipulated and scientific processes ignored during randomized controlled trials – held out as the gold standard in experimental design. Scientists, including 'pure' scientists, are as vulnerable to pressures as the next person, what differs is the source of the pressure. In academia, for example, there is pressure to publish research in journals and this leads to a reinforcement of critical, generally negative findings, and a fiercely competitive context within which research is being carried out; the prize goes to the first in print. If we think of ethics and integrity as one dimension, and consider its relationship with the extent to which researchers and practitioners are working together as a second dimension, while these two dimensions may not be orthogonal, neither should we assume a high positive correlation. This is the assumption of the experimental methodologists when they criticize those who work closely with practitioners and policy advisers.

So the ethical vulnerability of research methods is not unique to any particular approach. But let us set that aside for the moment and look at some of the specific criticisms of made against researchers working more closely with policy advisers and practitioners, and then look at ways in which these criticisms can be met, without prejudice to the need for closer working partnerships between the three groups.

The contemporary discussion of these issues in the evaluation research literature tends to be phrased in terms of 'advocacy' versus 'truth' (Shadish, 1997). If, in the final research report, the authors go beyond the established facts to the point of commending a program, which they had been evaluating (because the people involved are 'nice', and doing their best in what are almost certainly going to be difficult circumstances, and so on), then they have become advocates rather than objective evaluators. This advocacy role is endorsed by Stake (1997) as a necessary evil. He feels that it is not only legitimate but also inevitable and desirable that evaluators should take on such a role. He cites as an example a study in which he and his colleagues were evaluating the Chicago Teachers Academy. In the final report "we have emphasized Academy strengths .... Weaknesses we have not as persistently stated. Of misdirections, we have spoken more privately. Good moves we have repeated to more distant listeners... We have tried not to author a final report that, if distributed, might add to the Academy's insecurity." (1997:473) Surely this is taking 'sympathy' with the goals of the organization too far? Such a report is neither balanced nor 'true' and importantly, would not help in contributing to a body of knowledge about what works in teacher training or how to improve the service delivery in this respect. As Stake himself says – "Many will find this ideology paternalistic, protectionist." – just so.

That is not to say that there is no place for advocacy in the research world. This is what was suggested in the previous chapter as a means of increasing the influence of research. But it was not the results of a particular evaluation that were to be commended, but rather the process of research itself. If an experimental intervention is judged to have failed, in whole or in part, then that is the truth of the matter. There is little justification in obfuscating the result because the project workers tried hard and the cause was 'good' (in the view of the evaluator), or, in the policy evaluation context, because the program has already been counted a success and full-scale implementation is underway. Rather, advocate for further experimentation to develop a program that *does* work.

Scriven (1994, 1997) and others, take the polar position to that of Stake, arguing that truth and objectivity are the *sine qua non* of evaluation, and to take on an advocacy role is

to do a disservice to the client, the sponsor, and the wider research community. Many of the examples given by Scriven are frankly unethical, for example, evaluators who make recommendations for further training and then offer to provide it. But Scriven arguably goes too far in the opposite direction taking the view that the closer the relationship between researcher and practitioner the more *inevitably* contaminated the results will be, and in seeing the only alternative to close working with practitioners (and the inevitable contamination of results) as a totally distanced approach with all its disadvantages as outlined in the previous section.

While a close working relationship between researcher and practitioner does risk casting a rosier glow on outcomes than might otherwise have been the case, there are ways to guard against this without throwing the proximity baby out with the bathwater. Some options have been suggested by Weisburd (1994) whose starting assumption is that the researchers will not pull any punches in their evaluations and that the issue is how best to manage the role tensions that inevitably exist between the practitioners and researchers. Weisburd recommends carefully mapping out the research design from the outset, with the practitioners fully involved, and clear statements made (and agreed) about how emerging information will be handled as the evaluation progresses, a clear definition on what constitutes 'success' and how that will be measured, how the results will be disseminated and how (if at all) program reports might differ from research reports.

If, however, we start with the proposition that the researchers are vulnerable to being 'soft' on the practitioners, as Skriven tends to assume, then a rather different set of options emerges. First, the outcome measure needs to be valid, reliable and independent of the evaluators. If the outcome measure is crime related, for example, then short of actually fiddling the figures, there should be less of a problem than if the study is reporting on a process, or using some other outcome measure, which is dependent upon interpretation. So that is a first step in the corrective direction – choose outcome measures that are valid, reliable and open to independent scrutiny.

81

Next, the process of working effectively with practitioners (and policy advisers) is as much about managing their expectations as anything else. If they believe that the researchers are there simply to rubber stamp the project efficacy then they are hopefully operating under an illusion and it is important to make the terms of engagement clear. An evaluation is an attempt at an objective assessment of a program or project. Unfortunately this is not always as clear as it might be. As Sherman (2000) remarked at an NIJ seminar "if you put forward a hypothesis you are presumed to support it." And therein lies the problem.

> "If you put forward a hypothesis you are presumed to support it."
>
> Sherman: NIJ Seminar April 2000

The same potential difficulties arise for researchers working close to policy advisers, where there are similar pressures to 'prove' that the latest policy initiative works well. Campbell in his seminal paper 'Reforms as Experiments' (1969), discusses the relationship between the researcher and political expectations of the research:

> ".. *specific reforms are advocated as though they were certain to be successful.* For this reason knowing outcomes has immediate political implications. ....If the political and administrative system has committed itself to the correctness and efficacy of its reforms, it cannot tolerate learning of failure." (1969:409, 410, original emphasis)

In this new, outcome focused world the policy advisers are in a major bind. On the one hand they still want the 'good news' story – the confirmation that the already announced new initiative works – but they also want a valid assessment of the value of the program: *Did* it 'work'? Campbell offers a way out of this bind in his 1969 paper, which is as relevant today as it was then, but only marginally more likely to happen. He suggests that we need a shift in political posture away from the advocacy of a specific reform to the advocacy of the seriousness of the problem:

"The political stance would become: This is a serious problem. We proposed to initiate Policy A on an experimental basis. If after five years there has been no significant improvement, we will shift to Policy B." (1969: 410)

Unfortunately Campbell seems to ignore the probability that politicians do not normally work to five-year time frames. If Policy A is really going to take five years to test, then by the time the results arrive there could well be a quite different administration in place, or at least different individuals, with their own prejudices and concerns. The answer, as with the practitioner, is to be clear at the outset that the results will be what they are, and they may or may not confirm the program as a complete or partial success. If the political position is that regardless of the evaluation results the program will be expanded and declared a winner, then the best advice to the politician is to save the money and not bother with the evaluation in the first place.

Perhaps the most direct and important means of guarding against threats to objectivity is through transparency – the results should be published unless pre-publication peer review suggests that there are methodological flaws, which cannot be corrected. The availability of a report on a study should be a requirement regardless of the methodology used, but is less likely to happen if the study is carried out by a consultancy company, not for profit organization or professional evaluator with no personal or professional investment in publication. Academics, on the other hand, have a vested interest in ensuring publication. Sufficient detail to enable replication is also important. Publication and scrutiny of results by the academic community and by the critical and independent media should help to offset any tendencies toward partiality, exaggeration or significant departures from the 'truth'.

3. Writing it down, and other things

If a judgment has been made to spend public money on a research project or program, which may cost hundreds of thousands of dollars, then there seems little point in baulking at spending a reasonable amount in telling people what the results were. This section discusses the ways in which that might be done, and touches on some of the techniques

for disseminating the product. The word 'publication' is used here in the sense of 'made public'; it does not solely mean in the form of a published report, although this is still the most usual form of 'publication'.

a. Publication policy

It is enormously helpful to research contractors and to project mangers, if funding organizations have a clear publication policy. Knowing what is to be done with a final report, or other research product, should affect the way in which it is written and the way in which the contract is managed. This can save a great deal of time and effort – it is not 'good practice' to have the same piece of work written up more than once.

Of course it is in the nature of research that sometimes the results are not as useful as had been hoped, perhaps some mistakes were made in the methodology, or the conclusions of the study were felt to be trivial in some sense, and a decision was taken not to make the results available at public expense (that does not preclude the researchers publishing elsewhere). This should be a fairly rare event, however. In the UK there is a presumption that the results of research funded by the Home Office will be published. The policy in the Police Research Group was to publish all reports unless they were critically reviewed by the external reviewer (a rare event) or were felt to be of operational sensitivity by the police (even more rare). This policy was written down and formally agreed within the Home Office and with the police service. This formal agreement was helpful when, from time-to-time, it was suggested that it might be better not to publish a particular research report.

The formal procedure was that the internal project manager provided immediate feedback on a first draft produced by the researcher(s). Thereafter a revised draft would be sent to relevant policy units and practitioners for comment on the factual accuracy, and to an external reviewer for critical appraisal. A final draft would then be produced and agreed for publication by the series editor. It would be formally approved by the Head of the Police Policy Directorate and then be sent to the printers. In the UK, Ministers are invited to consider the appropriate date for publication of the study, which is the point at which

political considerations might come into play, but they are not asked to approve publication of what is expected to be politically neutral research. A similar process applied to the publication of research by in-house staff: The researcher's supervisor would consider the first draft and thereafter the process was similar to an external project in being externally reviewed and commented upon for factual accuracy. A short executive summary and points for police action would also be produced for dissemination at the same time as the full report.

The publication of research funded by NIJ has a different decision tree. Here, the researcher produces a final report, which is externally reviewed. The comments are passed back to the researcher who may take in any changes felt to be appropriate. The project manager makes recommendations on the dissemination of the work, but it is not necessarily the case that the full report would be published by NIJ. A number of products might be recommended, including the production of a 'Research in Brief' (RIB) publication all of which then have to be separately written. The proposals then go through the supervisor to the publications committee (which meets about every two weeks) and which then makes a recommendation to the Director of NIJ who takes a final decision on whether or not the research will be published by NIJ and in what form. This can be a protracted process.

b. Style of writing

Neither practitioners nor policy advisers find the academic style of writing an easy read. Indeed some academics find themselves struggling. One reason for this is that most disciplines develop a particular language – a shorthand – which helps them to discuss complex issues efficiently. They begin by defining their terms and take it from there. Sometimes these definitions can be highly specific, and part of their purpose is to facilitate the conveyance of nuances of meaning, which might otherwise be lost. One of the difficulties with 'criminology', however, is that it encompasses a number of 'base disciplines' – sociology, psychology, anthropology, geography, and so on, each with their own 'language'. Furthermore, the people for whom criminology might be relevant in the practical sense, the policy advisors and practitioners, may well have been trained as

managers, police officers, lawyers, or in the case of the British civil service, almost anything. They may or may not be familiar with concepts such as 'modernism', 'post-modernism', globalization, 'discriminative stimulus', 'contingencies of reinforcement' or whatever the buzz-words of the time might be. The picture is further complicated by the inclusion of complex statistical analyses, the interpretation of which is debatable even for those in the know.

> "There is a view both in Britain and the US that if you write for, or are accepted by, a large audience, you're writing for the masses rather than for your scholarly peers and that you compromise your scientific values. That danger does exist, but I have always thought it was possible to write intelligently for the public"
>
> J K Galbraith: quoted in 'The Guardian' August 5 1999

So it can be challenging for criminologists to understand each other, and when non-criminologists join the group the challenges can multiply. It is important, therefore, that research reports are written in plain English with the reader in mind.

In order to ensure that the reports, which are planned for circulation to non-specialists, are intelligible to them, a number of approaches can be tried. First guidance can be produced for academics and consultants on house style; feedback can be provided to authors through an editorial process, either from the project manager or from a specialist editor within the central government funding agency, and authors can be encouraged to write articles for submission to specialist journals as a separate exercise from that of communicating to non-specialists. Staff in the former Home Office Police Research Group all attended a plain English course, which is run by a UK organization devoted to the use of plain English generally, but particularly in government publications.

c. Making research results available

The traditional manner of publication for researchers is in academic journals or books with an academic orientation. Consultancy companies or smaller, not for profit organizations, would not necessarily expect to publish their work but would present it to the funding agency as a written report, perhaps accompanied by a disc version these days. Neither the academic output, nor the formal report to the funders, is likely to be read by practitioners, and policy advisers would also be unlikely to read a long consultancy

report. The question arises, therefore, as to how best to transmit the knowledge from research and consultancy exercises to those with a practical or policy interest.

First in line would be paper-based products in the form of books, leaflets, briefing notes, monographs and so on. Clearly any individual research report can be published as an ad hoc paper and circulated to those with an interest. But it is relatively rare that a report is published just at the moment when it is required by a policy adviser or practitioner. And since ad hoc publications are particularly prone to being forgotten, publication in a numbered research series, which can be listed with the list being regularly re-circulated, as the next report in the series emerges, increases the probability that research reports will be remembered and used when needed. An additional consequence of this approach is that reports need to be published in separate series reflecting different customer interests – policing, corrections, probation and so on. To do otherwise results in the provision of reports to those with marginal interest or gaps in the series where, for example, police agencies do not receive the corrections research and vice versa.

But as we have already noted both policy advisers and practitioners are busy people with often neither the time nor the inclination to read long reports. Short summaries, along the lines of the NIJ 'Research in Brief' Series are a good idea, especially if they spell out the policy and practical implications of the study. Asking the researcher to provide these recommendations is a good discipline in focusing the minds of the research team on the practical or policy value of the study.

These days the Internet is an obvious alternative or additional format for the production of full reports or shorted summaries. Central government agencies globally, are making increasing use of the web as a means of disseminating research. There was, for example, a 72% increase in the number of hits on the Police Research Group website from the first quarter of 1998 to the first quarter of 1999. In NIJ hits on the front page have increased from about 500 per week in 1997 to about 4,000 in 1999. The drawback of the use of the Internet in this way is that it is all too easy to do and risks the placement of poorly drafted or overly long reports being placed there. Information overload is a possibility these days

and in the absence of well developed filtering systems the use of the Internet may well start to decline as a source of information for busy professionals if this issue is not addressed.

If research results are intended for groups of professional practitioners, then making it easy for them to summarize the findings in their own in-house journal can be a useful option. In the UK, for example, all police forces have their own force newspaper or other means of dissemination. The easier it is for editors at local level to take results straight into their own publication the more likely it will be to happen. NIJ now publishes 'Research Review', which is a compilation of selected short summaries of recently published, NIJ-funded research. The review is now being made available through the Internet, which will be a major advance in getting it out quickly and cheaply to a wide audience. It has the added advantage that the reports, and the means of accessing them, can be patched directly into police in-house publications and those of other agencies quickly and with minimal effort.

Talks at conferences and seminars offer another obvious means of dissemination. NIJ has developed its 'professional conference series', which is a particularly good example of its kind. It involves funding a commercial organization, which organizes a variety of events on NIJ's behalf, including an annual Research and Evaluation Conference, which brings together a wide range of policy makers, practitioners and researchers from across the criminal justice system; the Perspectives Series, which is targeted at politicians and political advisers on Capitol Hill, and other conferences and seminars, which are arranged on an ad hoc basis.

Another method of dissemination, particularly in the policing field, is through the provision of research based training material, or technical assistance. The Home Office has direct responsibility for police crime prevention training, and much of the material produced in the form of research reports is targeted at police training agencies. In the US, NIJ has established a Crime Mapping Research Center, which provides technical advice and support to police and other agencies with an interest in crime mapping and the use of

88

geographical information systems generally. Initiatives such as these offer effective means of introducing research based ideas into the daily routine of policing, but they can be very time consuming of research expertise if the researchers themselves become involved in training directly. Training the trainers is perhaps a more efficient way to proceed, although this tends to result in a loss of spontaneity in presentation of the information to the trainees.

If all or any of these approaches are to be encouraged as a means of improving the integration of research into practice then there are implications for the funding of researchers. Most contracts fund the research and expect the production of a written final report of some description. This is a fairly universal expectation. If progress is to be made in this area then perhaps we need more flexibility about what constitutes the 'product' of the research exercise. For example, a training program, video, computer-based interactive CD, or a variety of other 'products' could be required as an alternative to or as well as the final report. Such an approach requires the commissioning organization to give more thought to the intended use of the research at the time of commissioning it rather than at the final stage – after the delivery of what may be a long and unmarketable written text.

d. The media

If research is to have an increasing influence in the formation of policy and the development of practice, then greater attention has to be paid by researchers to the media. Most government agencies have their own press office and there are often strict rules about the extent to which individual public servants can interact with journalists and other members of the media. There are good reasons for this. Members of the press can be highly selective in what they focus on, and how they present material, and naïve researchers can find themselves in trouble fairly quickly!

There is, nevertheless, a role for research managers in advising contractors about access to the media and the form it might take, and a role for in-house researchers who may have carried out research themselves, which is of interest to the media, in assisting with drafting press notices. Courses on how to make presentations at press conferences could

89

benefit senior staff, as could advice on TV appearances. All staff might find advice on drafting press notices helpful.

4. Feedback

This section discusses feedback at several levels and between different groups. At the highest level there is feedback provided by the recipients of research (both policy advisers and practitioners) to the research community on the 'quality' of their product. In the course of this project, and throughout a fair part of my earlier career, I have heard a disproportionate amount of grumbling from policy advisers and practitioners about research and its products. But very little seems to have changed as a consequence, and one reasons for this is that systemically, researchers are not 'rewarded' for the production of 'useful' research. In some contexts, and for some research, this is not a problem. Pure mathematics, for example, or astronomy, are not judged on their practical usefulness, and there is a proper expectation that in other fields, including criminology and policing, 'pure' research as it tends to be called, should be supported by central governments on the basis that it is a 'good' thing, and may, in any case, contribute to the greater good in the longer term. But in fields of social policy, where there is a pressing need for evidence based decisions, there needs to be a proper balance between 'pure' and 'applied'. Arguably that balance has not been right for some time. Furthermore, practical or policy oriented researchers are seen as the poor cousins of the 'real' (pure) researcher in the academic community. A certain scientific snobbishness is evident!

If significant improvements are to be made in the value placed on applied research in the academic arena, then some leverage will have to be used. The most obvious is money. In both the US and UK university departments need research funds and more explicit signals from central government on their expectations of publicly funded research are now being sent. In the UK, the ESRC (responsible for funding social science research in Universities) now requires those seeking research funds to provide a statement of the likely effect of the research and a statement of support from practitioners. In the US, NIJ has sponsored a number of events, which offer the opportunity to debate these issues directly (Travis, 1998; NIJ, 1999).

90

Universities could also be ranked on the funds they attract for applied research exercises. In the UK there is a research assessment exercise, which looks at the extent to which university lecturers manage to get articles published in refereed journals. This is the way in which the research is judged. In similar vein, academics count the number of citations they each give each other, and this too is used as a measure of worth. Although it is more difficult to measure, little attention is paid to the extent to which academics contribute to policy or practice. There are some unusual difficulties in doing this. For example, some of the more widely used activities in policing stem from research although this is not common knowledge to the police officers involved. The SARA (Scanning, Analysis, response, Assessment) model, for example, was developed from a problem solving exercise in Newport News, Virginia (Eck and Spelman, 1987) and is now quite common in police agencies in English speaking communities. But there are not many police officers able to quote the source.

At the specific project level, rather than at the level of research in general, different feedback issues arise. I have already argued for more, and more critical communication between researchers and practitioners in the course of research evaluation exercises. There is little to be gained from researchers maintaining their distance while projects drift off the rails heading for implementation failure, when a judicious input from the research team could keep the initiative on track, or lead to an improvement in program delivery.

At the end of an evaluation the lessons learned by the research team need to be conveyed to the policy advisers or program funders following the terms of the research contract. But they also need to be conveyed to the practitioners in a sensitive manner. Feedback, or communication generally from researchers to practitioners during the research exercise can help in the transmission of what may be difficult or critical final messages. Objective evaluation can sometimes come with a heavy price tag to the hardworking practitioner. Bad news is easier to handle if it is not completely unexpected.

Feedback from researchers to practitioners is not that unusual but feedback from practitioners to researchers about the *experience* of being the subject of research is less common. It may be a useful discipline if research and evaluation teams adopted the practice of actively seeking feedback from the subjects of their research on a more systematic basis. Was the experience useful, enlightening, or constructive? Did they learn anything they did not already know? How could it have been done better? The answers to some of these questions may emerge during the course of a project if there is a move toward partnership between researchers, policy advisers and practitioners with a shared agenda to learn more about the process of outcome delivery.

5. Research/policy/practice models

There have been a number of attempts to develop the relationship between researchers and their partners through specific funding arrangements and/or by ensuring that researchers work in the same parts of the organization as their policy colleagues. Some of these are described below.

a. Locally Initiated Research Partnerships

One of the more ambitious and recent attempts to bring the police and the research community closer was through the NIJ/COPS funded Locally Initiated Research Partnerships program (LIRPS).

It marked a significant shift from the normal funding arrangements under which the researcher would typically identify a problem and perhaps a cooperative police agency in which to locate the work, and apply for central funding. The researchers would, under this arrangement, retain control of the process. The LIRPS program was different in this respect. Instead of the researchers being in control of the process it was a partnership between the police and the research team, with agreement on the problem definition and research process. This was a 'bottom up' program in which the researchers and the police worked together to identify a problem or project on which they were keen to work together rather than one which may have been set centrally or one on which the researchers have a specific interest.

A total of 41 projects, costing approximately $6 million, were supported from 1995 with funding from the Crime Act (1994), which *inter alia* was intended to support community policing and problem solving. The LIRPS program complements the basic premise of community policing: working as partners achieves more than working alone (McEwan, 1999). The Institute for Law and Justice is evaluating the whole program centrally but their final report is not yet available. Interim results reported to the NIJ in June 2000 (McEwan and Spence), suggest that the program has been successful on a number of fronts, although there are also lessons learned for similar exercises in the future, which might improve the final product. Some of the results are summarized in Table 5 under the headings 'good' or 'bad' news.

Table 5: Selected results from the evaluation of LIRPS

Good news	Bad news
Most relationships have evolved into genuine, equal partnerships, but in different ways	Turnover has been a significant problem in several partnerships and is probably the greatest threat to their long-term stability
Police commanders appreciate the objectivity that 'outsiders' bring	The term 'research' is very broadly used to cover almost anything the department considers important
Researchers have been welcomed (because most departments are understaffed for planning and research)	Descriptive analysis is the most common analytic approach, only a few partnerships are doing experimental designs
Partnership products come in a variety of shapes and sizes – slide shows, training packs, tables, etc	Many projects are behind schedule
Successful partnerships have several spin-offs such as requests for further work	Availability and quality of data is a problem
Results have been used in unanticipated ways	Some partnerships have experienced difficulties and delays in getting contracts between the city and research organisation
Useful work experience for interns in police departments, opportunity for research dissertations and Master's Theses	Researchers have not well understood the political environment of police departments, especially in large cities
Wide variety of problems addressed	Initial lack of trust

The decision under which heading to put some of the items is not necessarily obvious. For example, the fact that research is very broadly defined and is used to cover almost anything a department thinks would be useful, is not seen as a positive attribute of the partnerships. There are, as was discussed in Chapter 1, a whole range of possible

approaches to the 'research' exercise, and their appropriateness varies with the problem, as do the skills of the researcher. If the police agencies and researchers are to become effective partners, then not only do the researchers have to learn about the way in which the police operate (a point fully conceded and emphasized by McEwan), but the police agencies need better to understand the options before them from the research field; they need to become clearer about what they need and more demanding in ensuring that they get it (Laycock, 2000).

McEwan and Spence lists some of 'key factors for successful partnerships', which include key personnel remaining in post during the entire period (which may not be achievable); researchers acquiring an understanding of police culture, and police personnel learning the benefits and limitations of research. The criteria adopted for judging the success of the LIRPS program by the central evaluation team include at least one of the following three measures:

- ☐ Whether the department changes as a result of the research
- ☐ Whether information systems have been developed or improved
- ☐ Whether the partnership continues beyond the life of the initial research project. (McEwan, 1999:7)

It is difficult to assess the extent to which the LIRPS program addresses some of the issues raised as important in this report. It is, for example, disappointing that hardly any of the 41 projects have taken a data driven or hypothesis based approach to problem solving. Most of the studies have involved organizational development/implementation (13 projects), surveys or evaluations (15 projects) and compstat or computer mapping (6 projects); four have addressed domestic violence, and a further three projects involve 'multiple activities'. It seems reasonable to assume that this round of LIRPS projects will not advance the 'what works, where or why' agenda, although the continuing involvement of researchers with the police agencies may lead to a move in that direction.

94

b. Strategic Approaches to Community Safety Initiative

Another NIJ initiative, launched in 1998, brings researcher together with other partners, and is known as the Strategic Approaches to Community Safety Initiative (SACSI) program. This was first established in five sites, New Haven, Connecticut; Winston-Salem, North Carolina; Portland Oregon; Memphis, Tennessee, and Indianapolis, Indiana. The US Attorney takes the lead in establishing an inter-agency working group, gathering information about local crime problems, designing a strategy and interventions to tackle the crime problem, implementing the interventions and evaluating and modifying the strategy as necessary. The local researcher acts as a full partner on the group and assists in the analysis of data and development of the initiative.

The sites are tackling a range of different problems although there is an emphasis on gun violence, particularly following the publication of an influential report on tackling youth gun violence in Boston (Kennedy, et al, 1996). The Department of Justice is supporting the initiative by providing grants to local research partners, cluster meetings, technical assistance and the development and installation of the Community Safety Information System, which provides a crime mapping facility.

The SACSI program is being evaluated at national level by an independent team of researchers who have yet to report their findings. The program has, nevertheless, attracted a great deal of interest from US Attorneys around the country, who are keen to be involved in their own areas and there are plans for its expansion.

c. Researchers working within a policy unit

In 1983 the Home Office established a new policy unit with specific responsibility for crime prevention (the CPU). When the Unit was established a small team of researchers were seconded to it from the main Home Office Research and Planning Unit. Their management was split between the Head of the CPU, on day-to-day matters, and the Head of the Research and Planning Unit, for professional purposes. This was initially a three-year experiment; the more normal arrangement was for all social science research staff to be centrally located and to carry out research on a customer/contractor basis.

Laycock (2001) has described some of the consequences of this arrangement. It had a number of advantages in helping the researchers better to understand the policy process, and in ensuring that the research agenda was more closely aligned to the developing agenda, ultimately driving it. During the course of the experiment a new series of research papers was established (Crime Prevention Unit Papers), which were specifically designed to influence police policy and were widely distributed to the police practitioners in the UK and abroad.

The experiment was reviewed in 1986 and considered a success. The researchers remained in the CPU until 1992, when they were transferred to the newly created Police Research Group (PRG), which had a wider remit to carry out research on policing issues beyond, but including crime prevention.

The establishment of the PRG meant that the researchers no longer had the close relationship with the policy unit and this resulted in a lost training opportunity for the researchers, particularly the young recruits who were new to central government. As a way of filling this gap, a new system was introduced. Junior staff were seconded to work in a policy unit for about 10% of their time over a three-month period during the first few years of their careers. They were there to learn about the work of a policy unit, not to carry out research. This experience enabled them better to understand the world in which policy makers operate.

In these posts they wrote or contributed to speeches, provided briefing for parliamentary debates, answered parliamentary questions or Ministers' cases and generally absorbed the work of a typical Home Office policy unit. There were also a number of attachments to Ministerial private offices, including that of the Home Secretary. These high level attachments were particularly useful in the context of the British Civil Service in exposing young research staff to the pressures of policy making at the highest levels of Government. They also had the incidental advantage of demonstrating the high intellectual quality of many of the specialist research staff and offered the opportunity for

the researchers to explain how existing research might be relevant to ongoing policy issues.

6. Good practice   *Prescriptive*

In this section we look at some existing good practice and some principles, which might guide further developments in the process of commissioning, executing and disseminating research. Some of the examples already described above, such as the LIRPS program, or SACSI, constitute good practice, although both would benefit from some 'fine tuning', as is intended should be provided through the national evaluations of both programs, which are still underway. But they are operating at a fairly high level; the remainder of this section considers good practice in a more detailed sense, working again through the process from commissioning to dissemination.

Good practice in the **commissioning process** depends upon the commissioners' aims. If the intention is to provide funds for 'pure' research, then it may be less appropriate for the commissioning authority to involve practitioners or policy advisers in the process of developing the agenda, although if only for pragmatic reasons of requiring their co-operation, it may be wise to keep them informed. Even with a heavily research driven agenda, there is always a co-operative chief officer who is prepared to 'host' a project, or to make data available.

But the argument in much of this report is that this way of operating is no longer as appropriate as it once was. There is a trend at national level, which is being reflected at State and local level, for an increasing interest in evidence relevant to policy and practice, and this requires a different approach to commissioning research. In this scenario, the following rule apply:

> A conscious decision needs to be made between whether a project should follow the procurement process, take the form of a research grant, or be carried out by in-house staff.

- Practitioners and policy advisers need to be involved in setting the research agenda. The question is how best to do this, not whether to do it.
- There should be less emphasis on the evaluation of existing projects and programs and more on the development of new, research based ones.
- There should be fewer resources devoted to process evaluations, valuable though these may sometimes be, and more on outcome focused evaluations.
- There should be fewer ad hoc projects, or lists of projects masquerading as programs; more strategically driven programs are need.
- There should be capacity for in-house research, which contributes to the social science research agenda. In this way young research staff can be trained not only to carry out research, but to do so in a policy relevant manner.
- Research results need to be delivered in a timely manner: this means that funding agreement for research should be efficient; projects should be tightly managed; dissemination should be rapid; actionable interim reports should be encouraged.

Moving on to the **execution of research**, and bearing in mind the need to provide information on what works, where and why, there are a whole plethora of possible approaches that might be used, some of which were discussed above. The conclusion there was that the scientific realist or theory of change approaches, combined with the ideas of Yin on testing plausible alternative hypotheses, offer the most appropriate approaches to the development of the kind of knowledge that is required. There are some caveats to these conclusions, however:

- The results need to be published to ensure the integrity of the research itself and to facilitate much needed replication.
- Outcome measures need to be valid, reliable and transparent, and as little subject to interpretation as possible.
- The expectations of practitioners and policy advisers need to be managed.

That is not to say that randomized controlled trials are no longer of use. They remain the best way to ensure the internal validity of experimental interventions, and where individual treatment methods are being tested, their status as setting the 'gold standard' remains unchallenged.

Good practice in the **dissemination of research** results starts with a clear statement of the publication policy of the organization, with guidelines available for authors on what is expected of them for the different outlets that may be available. Beyond that there needs to be:

- ☐ A policy of using plain English
- ☐ Writing style appropriate to the target audience
- ☐ An efficient publication process designed to reduce the delay in disseminating results
- ☐ A 'write it once' policy ie agreement on the publication outlet, or other means of dissemination before the researchers begin to draft their reports
- ☐ Sufficient funds in the grant or contract to fund the dissemination of the results, including attendance at any appropriate conferences
- ☐ *Judicious* use of the Internet as a means of dissemination (ie no unedited or poorly drafted material simply being 'made available')

Not only do the research results have to be made available at the project or program level, but information on 'what works' might usefully be pulled together periodically to provide information to practitioners and policy advisers but also as a means of checking that the research funding agencies are getting the balance right between process and outcome evaluations. One way of doing this is through a Cochrane-type exercise and this has now been proposed. The 'Campbell Collaboration', which compares with the Cochrane Collaboration described earlier, is now being developed as a means of assessing evidence on what works in education, crime and other areas of social policy. A specific proposal for a Criminal Justice Group has been made. It will be important to ensure that defensible criteria are developed for the inclusion of studies in meta-analyses. As argued above restriction to RCTs would, on the basis of the arguments set out in this report, lead to the

exclusion of a considerable amount of useful work, and the inclusion of a number of unhelpful 'nothing-works' or frankly misleading evaluations.

It is also worth developing a media handling strategy associated not only with individual publications but also with a strategic program. Part of the purpose of research is to raise the level of public debate on crime and criminal justice issues generally, and learning to communicate well with the media is part of that process.

Ensuring that research results are brought to the attention of others groups or individuals of influence is also helpful. In Australia, for example, Commissioner Tony Fitzgerald carried out a review of corruption and misconduct in the Queensland Police Service and the public sector, which was extremely critical of the way in which criminal justice policy was formulated in Queensland. His solution, which was accepted in full, was to create an independent agency – the Criminal Justice Commission – with an integral Research Division. Professor Ross Homel, of Griffith University, Brisbane, served as a part-time Commissioner from 1994-1999. In this capacity, and working with the staff of the Research Division headed by David Brereton, it was possible to raise the profile of research carried out. Despite these favorable structural arrangements there was still felt to be insufficient integration of research and policy – it fell a long way short of the ideal envisaged by Fitzgerald (Brereton, 1996).

A similar process, in this case of linking research to a high profile inspection process, has been used in the UK more recently. Researchers have worked closely with the Police Inspectorate in gathering information during the course of inspections and have published a research report alongside the inspection report (Hough and Tilley, 1998 and Her Majesty's Inspectorate of Constabulary report, *Beating Crime*, and Read and Tilley, 2000 and Her Majesty's Inspectorate of Constabulary report, *Calling Time on Crime*). This has resulted in more publicity for the research and a wider distribution of reports and recommendations arising from it with promised follow-up from the Secretary of State.

There have been a lot of 'recommendations' made in the latter part of this report and it is perhaps appropriate to return to the 'needs' of the practitioners and policy advisers, as outlined in Chapter 2, and ask to what extent these needs might be met if the recommendations were followed. Table 6 below summarizes the needs that were identified (without distinguishing between practitioners and policy advisers) and comments on how they might be met.

Table 6: Meeting the needs of practitioners and policy advisers

Needs	How will the needs be met?
To know what works, where and why	More investment in experiments with an outcome focus and more use of research designs which are explicit on mechanisms and contexts.
Help in replicating 'what works' – understanding contexts and mechanisms	Greater preparedness on the part of funding agencies to support replication of apparently effective projects in different contexts. Inclusion of these concepts in training programs for practitioners and researchers.
Cost included in evaluations	Encouragement by central funding agencies to include at least basic cost measures in new projects.
Help in generating testable hypotheses	Closer working relationships between evaluators and practitioners before evaluations are commenced to ensure that the hypotheses being tested are clear and agreed.
Timely research	Funding arrangements, which do not lead to excessive delays in commissioning 'hot topics'. Tight management of research contracts to minimize delay in completing work and a clear publication or dissemination strategy aimed at getting results out quickly.
Involvement in setting the agenda	Mechanisms in place to ensure that policy advisers and practitioners are involved in the research agenda setting process.
Reports and recommendations in plain English	Advice to report writers on 'house style' and training in report writing where appropriate.
To know of current good practice	Regular reviews of 'what works, where and why'. Support for the Campbell Collaboration and similar activities. Attention to the means through which good practice is developed and disseminated.
Feedback on the results of research in which they have participated	Routine mechanism established to ensure that research results are reported back to those who have contributed to their development.
Good news	Sensitive drafting, and early warning of potential bad news. Essentially managing the expectations of the recipients of research results.
Confidence in the results	External scrutiny of research reports; publication of results; transparency; high standards in training and support for researchers.

	practitioners operate, which in turn may require specific training opportunities for researchers. A clear statement should also be required of the extent to which any advice is speculative rather than deriving directly from the research data.

Some of the mechanisms, which are suggested in the table are, of course, already operating or are being developed. Research directors may, however, find it useful to work through a checklist along the lines of that set out in Box 5 below.

Box 5: Checklist for research directors

1. Is at least some part of your budget set aside for strategic research programs?
2. What proportion of your budget addresses 'what works, where and why' questions?
3. Do you help practitioners to develop testable, theory-based hypotheses?
4. Are some measures of cost included in evaluations?
5. Do your funding policies distinguish between grants, procurements and in-house research? Is it clear why you choose one rather than another?
6. Could you make the whole research process more efficient, and thus deliver results faster?
7. Who is involved in setting the research agenda?
8. Does your agency actively encourage 'action research'? If so, how?
9. Do you have a systematic way of identifying good practice?
10. Do you train your staff to write in plain English?
11. Do you provide advice on report writing to your contractors?
12. Do your contractors routinely provide feedback to practitioners with whom they have worked?
13. Do your staff members have formal training on the way in which policy is made?
14. Do your staff members have secondments to work with practitioners?
15. Do you prepare partners for good/bad news?
16. Do you have training programs for working with the media?
17. Do you have a written publication policy?
18. Do you have a marketing strategy for each project?
19. Do you know the key people who should receive your research summaries?
20. How do you ensue the quality of the work you commission?
21. Does your agency have its own 'experts'? How do you manage them?
22. Do you formally encourage feedback from policy advisers and practitioners with whom you have worked?

Chapter 4: How far off are we?

To summarize the picture presented so far in this report (and as was said earlier to slightly caricature reality), practitioners and researchers might well have been operating in different universes. The police have had research done *on* them by researchers who criticized what they found (because that is what researchers are trained to do). The researchers worked to their own agenda. They published their work in journals, worried about their tenure, their next grant, the number of citations they had amassed and the purity of their methodology. Contributing to the development of policing was not their job. The police complained and ignored the research. Basically they didn't need it. They didn't need the hassle it caused and they could carry out the tasks that were required of them without any help from researchers –thank you very much.

> "Contemporary criminologists although far more numerous, put most of their energy into educating specialist criminology students and writing for other specialists."
>
> Wiles, (1999)

The basis of the thesis in this report is that some fundamental changes are now underway which will change this scenario; it will change what I am calling the 'deep structure' of both the research community and that of the police. Changes, which should mean that research, or the techniques associated with it, will become far more central to policing than ever before. The police will need good research. And they will, therefore, have to become more aggressive in making sure they get it. The report points to three reasons why the police may now need good research in a way that they have not to date -

☐ The outcome focus and the need to demonstrate that police action can affect crime rates.

☐ The professionalization of policing - the need for a body of knowledge on what works.

☐ The shift to data driven problem solving.

These three reasons mean that the relationship between the police practitioner and the researcher is under pressure to change. Similar pressures apply at the policy level particularly driven by the 'what works' agenda and the control exercised by policy advisers over the budgets.

In this Chapter we look briefly at the philosophical and structural implications of getting social science research 'right' for policy advisers and practitioners. The chapter concludes with a brief consideration of the probability of being able to achieve the kinds of changes argued as necessary under the heading – what next?

Philosophy and structure

a. <u>In policing</u>

For many if not most of the police officers in the United States, and their colleagues in the United Kingdom, their understanding of their role still centers on the arrest and prosecution of offenders. This is despite the massive push from the Federal government through the 1994 Crime Act (in the UK the Crime and Disorder Act, 1998) to change the way in which policing is delivered. Although the majority of US police agencies now claim to be community oriented[1], and many are trying to 'solve problems', there is still, nevertheless, very little hard evidence that what they are doing under these headings is having any real effect, or that the claims to community policing are any more than rhetorical. The resistance of the police on the front line to problem solving is considerable, as UK research has shown (Leigh et al, 1996; Leigh et al, 1998).

Leigh and his colleagues argue that in order for problem solving to develop police agencies need a geo-coded data base; creative thinking; an appropriate police management structure; devolved decision making to a low organizational level; sustained senior management support, and a comprehensive training program (on which more below). Although some exceptional agencies meet these criteria, most do not. Geo-coded

---

[1] 60% of the police departments serving 100,000 or more residents had a formal written community policing plan by 1997 and, overall, about 90% of local police officers worked for a department with some kind of community policing plan (Community Policing Consortium website).

data, to take an example, is still at the development stage in many areas although it is being heavily supported through the NIJ's Crime Mapping Research Center, which has overseen a series of increasingly well-attended annual conferences, supported training, sponsored research and developed and disseminated a range of publications.

There is support at the Federal level for these developments, and some State and local agencies are encouraging moves toward outcome focused problem solving, but there are relatively few levers available to Federal Government within the United States with which to press this agenda. The situation in the UK is somewhat different where there is a greater degree of central control facilitated by the relatively small number of police agencies, at 43. Although the UK police at local level retain operational independence from central and local government (which means that day-to-day operational decisions are essentially free of political interference), they are accountable for the delivery of local policing through the locally based police authority, and centrally, through the Home Secretary (who is the Secretary of State appointed to the Cabinet with responsibility to ensure the effective and efficient delivery of police services throughout England and Wales). He himself is responsible to Parliament. The effect of this system is that a range of 'levers' is available to central government with which to exert influence over the police service as set out in Table 7.

Table 7: Centrally available 'levers' over UK policing

Category	Lever	Commentary
Financial	General police funding	The Home Office provides 51% of police funding.
	Project money	The crime reduction program has provided a major injection of funds in the crime reduction field, which are being used to influence the delivery of policing.
Legal	Police Inspectorate	There is a high profile central Police Inspectorate in the UK, which regularly inspects forces and reports publicly on performance. They have completed two recent inspections on crime reduction.
	Legislation	In addition to general legislation covering policing, the Crime and Disorder Act requires the police to work with local government and develop a strategy to tackle local crime and disorder based on a crime audit and consultation with the community.

Managerial	Training	There is considerable central influence over police training in the UK. Crime prevention training through a specially designated college is under direct Home Office control. This has the advantage that is has been easier to involve other agencies in the training programs, running joint training for example, but it has had the disadvantage of keeping crime prevention away from the 'mainstream' police training which is done elsewhere.
	Audit Commission	The Audit Commission has a statutory responsibility to report on the extent to which the police are carrying out their responsibilities in an effective and efficient manner. They carry out regular inspections of forces on various themes and publish their reports.
	Performance regime	Police performance against a range of criteria is assessed on an annual cycle. The Home Secretary sets some of the targets for crime reduction.
Intellectual	Research results	The free availability and wide dissemination of Home Office research papers has had increasing influence over UK policing since the early 1990s.
	Persuasion	Advice and guidance is provided to police forces on a regular basis from central government. Although much of this has no statutory force, it is often seen to be setting out good practice in the area, which it addresses.

The levers available to the Federal government in the US are relatively few and, because of the fragmentary nature of US policing such levers as there are, are fairly blunt. They are set out in Table 8. The major lever, which the Federal government does have, is money. The COPS office has overseen expenditure of $8.8 billion from 1994-2000 and the President has requested $1.3billion for the COPS program in FY2001 of which $135million is specifically for 'crime prevention'.

Table 8: Centrally available 'levers' over US policing

Category	Lever	Commentary
Financial	General police funding	No central funds are routinely provided to support local policing.
	Project money	The COPS program has provided substantial financial support to local agencies for a range of community policing initiatives including, but not exclusively, problem solving.
Legal	Police Inspectorate	There is no Federal oversight of the delivery of US policing at local level.
	Legislation	There is no Federal legal requirement for the local agencies to attend to crime prevention.
Managerial	Training	Police crime prevention training is provided nationally at the crime prevention college in Louisville, Kentucky and in other locations but there is no central control over the agenda. It remains marginal to 'mainstream' policing.
	Audit Commission	There is no national auditing of the delivery of policing at local level.
	Performance regime	NIJ has invested in a number of research projects under the general heading 'Measuring What Matters', which is intended to assist agencies in developing their own performance regimes from the top of the agency down to the performance of individual officers. This work is ongoing.

	Research results	The COPS Office, in partnership with NIJ, has funded $44million worth of research under the Crime Act. Much of this has been process, rather than outcome focused. The results have yet to be made widely available.
Intellectual	Persuasion	Extensive advice and guidance is provided through the COPS Office web site.

A range of Home Office reports have demonstrated potentially effective tactics for crime prevention, and the policing structure in the UK has meant that they have been fairly easy to disseminate. There has been no equivalent research in the US, and the fragmentary nature of US policing makes dissemination more difficult. There are two major advantages to the US system however. First it is more locally democratic with communities feeling they have, and in practice having, more direct control over the police. It is also easier to test a much wider variety of hypotheses about policing and its delivery because of the large number of agencies with a whole spectrum of inclinations and approaches.

One of the levers available to those wishing to influence police behavior is through training. Most reports call for more training and this one is no exception. There are a number of areas to which this applies. The police need to be much more aware of the range of problem solutions beyond mere enforcement. Most police training centers on the legal process – the powers of the officer on patrol – what he or she can and cannot legally do. Clearly there are good reasons for this and suggesting more for already overloaded training schedules is not going to be welcome news. But if the police are to raise their game strategically then this may be part of the price.

As was noted earlier they also need to be more assertive in ensuring that they get what they need from the research community. Although it is not a major exercise, and many police chiefs have research qualifications themselves (so it may be even less of an issue than is supposed), there is still a need to ensure that those making the decisions about police agencies working in partnership with researchers appreciate the sometimes subtle differences between what one researcher might produce rather than another. At its most simple, the police need to seek out 'applied researchers'. And somewhere in their training programs, at some level, they need to be told this.

Although attention to these issues through the training programs would be a step in the right direction, it would not be sufficient unless those officers returning from their training programs, with new ideas on problem solving and a more sophisticated view of what research may and may not produce, were to return to the same organization that they had left, with the same expectations of them and their research partners. The level of organizational change required is considerable and although a lot has been and is continuing to be done to change the nature of policing, the changes aspired to are deep in the structure and culture of the organizations. It will take a great deal of time and consistent effort to achieve the scale of change desired. It is, in the USA, a fairly high-risk strategy that appears to have been adopted. It is assumed that the COPS money will dry up soon and the question then will be whether a significant paradigm shift will have been achieved. If not, the community policing, and problem solving of which it is a part, will almost certainly atrophy. The energy required to maintain them, given the lack of structural and cultural support will be too great. This is less likely in the UK because the requirements that the police work in partnership to reduce crime is now embedded in legislation.

b. In research

There is also a need for a change in the behavior and to some extent the value system of, in this context, the criminological research community. As Joan Petersilia said in her Presidential Address to the American Society of Criminology (ASC), which focused on the relevance of criminology to policy:

> "The potential for policy 'irrelevance' is inherent in the scientific advancement of which we are all justly proud. It is also inherent in how we are trained, how we do our research, how we communicate our results and how we are rewarded."

> (Petersilia, 1991:8)

This theme was also selected by Zahn in his 1998 Presidential Address to the ASC, who in an otherwise fairly congratulatory speech, called for stronger research ties with practitioners, victims and offenders. He also quoted John Timoney, now Chief of Police

108

in Philadelphia, as saying, when asked what could be done about crime, – "criminologists could stop being part of the problem".

Also at the 1998 ASC meeting, Jeremy Travis, then Director of the National Institute of Justice, gave a plenary presentation in which he called for a new relationship between researcher and practitioner – "a relationship of constructive engagement, of partnership in the development of useful knowledge, of symbiosis in the testing of ideas." (Travis, 2000:85)

Similar concerns have been expressed in the UK and Australia. In a recent speech to a criminological audience, Paul Wiles, Director of Research, Development and Statistics in the Home Office, said:

> "Criminologists have ceased to play a significant part in the public debates about crime and crime policy, and the consequence has been that these debates have become less sophisticated and more simplistic." (Wiles, 1999)

And in Australia, Brereton (1996) while noting that there has been a marked increase in the quality and quantity of criminological research available to the governments in Australia, argues that key areas of crime and justice policy continue to be developed with little apparent regard to that research. More recently Braithwaite (2000) goes further in suggesting, "Criminology… is destined for decline." Although conceding that for the moment it is booming, "not fed by intellectual accomplishments of the field, but by the continuous growth in public sector employment in the criminal justice system combined with new expectations that police should be university graduates, and by even stronger growth in private policing." (2000:223) All of which resonates with Wiles comments that most criminologists spend their time training their students and writing for other specialists (Wiles, 2000).

This apparent pessimism about the future of criminology is clearly at odds with what Garland and Sparks (2000) describe as "the rude health of contemporary criminology"

(2000:190). They discuss this paradox at some length conceding that criminology has made a minimal contribution to the development of criminal justice policy, certainly far less that the founding fathers of the discipline in the UK might have hoped (Radzinowicz, (1999).

So on the one hand we have a booming discipline fuelled by a sudden influx of money through the Crime Act in the US and the Crime Reduction program in the UK, on the other we see little evidence from past performance that the bulk of practicing criminologists in universities have been able to influence policy or practice to date, or have the inclination or indeed the expertise to do so in the future: a grim picture.

There are, however, pockets of research activity, which suggest a more productive set of relationships might be developing. There is, for example, more 'applied' research being done. The work completed under the general heading of situational crime prevention (Clarke, 1983) is a good example, and Clarke and others have now edited a substantial collection of research papers in the 'Crime Prevention Studies' series of books, of which there are now 10 volumes in print. In Australia too the work of Homel and his colleagues has an applied orientation and has proved effective in influencing policy and practice (see for example the Surfers Paradise Safety Project, Hauritz, et al, 1998).

There is also a core of individuals, occupying key positions in terms of research funding, who are influencing the development of this agenda to a significant extent. These are the Heads of research in the National Institute of Justice in the US, the Home Office in the UK, the Institute of Criminology in Canberra and others. They now meet increasingly often to discuss emerging issues and share ideas. The American Society of Criminology meeting in 1999 was preceded by a meeting of research heads from a whole host of countries where in future just these issues might be developed. Although early days in the development of these relationships it was agreed that a further meeting would take place in 2000.

If attention needs to be paid to the training of police officers, similar arguments apply to the training of researchers. They not only need specific research-based training in the techniques of the applied researcher, they also need better to understand the ways in which policy advisers and practitioners work. This information either needs to be built into the training programs at post-graduate level, for those wishing to enter the applied research field, or, better yet, they need to be given the opportunity to work alongside policy advisers and practitioners as part of the ongoing professional experience.

The way in which they write their research reports is also an area in which training can help. As noted earlier the use of simple English needs to be encouraged, but this should be set in the context of greater understanding of the audience at which the report is directed.

The possibility that both policy advisers and practitioners are likely to become better 'customers' should help to create a virtuous circle. Both policy advisers and practitioners should become clearer in specifying what they want from research and more assertive in making sure they get it (Laycock, 2000). There is too much money on the table, and too much pressure on the delivery of results, for them to do otherwise.

What Next?

This report has described an evolving scene, one in which huge changes are underway in the relationship between government and the people governed. Aspirations have been expressed for more rational policy making and an evidence-base to practice. It is possible that these aspirations will prove to be transitory, as new governments are elected and a new agenda develops. But at the moment there is support at the very highest level of government, both in the US and UK, for the kind of rational decision making which would raise the profile of research. The question is, how long will it last? And will the research community be able to respond to these challenges in timely fashion?

We will have to wait and see. But if research is to take on a more responsible and responsive role in relation to the development of social policy and practice, there has

never been a better time to do so. My personal hope is that we can all rise to the

challenge.

REFERENCES

Argyris, C. and Schon, D. A. (1996) *"Organizational Learning II: theory, method and practice"* Reading, MA: Addison-Wesley.

Barnett, W. S. (1996) *"Lives in the Balance: Age-27 Benefit-Cost Analysis of the High Scope Perry Pre-School Program"* Monographs of the High/Scope Educational research Foundation, 11, Ypsilanti, MI: High/Scope Press

Barr, R. and Pease, K. (1990) *"Crime Placement, Displacement, and Deflection"* In Michael Tonry and Norval Morris (eds) Crime and Justice: A Review of Research, vol. 12, Chicago and London: The University of Chicago Press.

Bayley, D. H. (1994) *"Police for the Future"* New York, Oxford University Press.

Bayley, D. H. (1998) *"Policing in America: Assessment and Prospects"* Ideas in American Policing, Police Foundation, 1201 Connecticut Ave, NW, Washington DC 20036

Bennett, T. (1990) *"Evaluating Neighbourhood Watch"* Aldershot, Hants: Gower

Blumstein, A. and Rosenfeld, R. (1998) *"Explaining Recent Trends in Homicide Rates"* The Journal of Criminal Law and Criminology, Vol 88, No. 4

Bowling, B. (1999) *"The Rise and Fall of New York Murder: Zero Tolerance or Crack's Decline?"* British Journal of Criminology, 39 (4)

Braithwaite, J. (2000) *"The New Regulatory State and the Transformation of Criminology"* British Journal of Criminology, 40 (2) 222-238

Brereton, D. (1996) *"Does Criminology Matter? Politics and the Policy Process"* Current Issues in Criminal Justice, 8, (1)

Bureau of Justice Statistics (1999) *"Local Police Departments, 1997"* US Department of Justice, Office of Justice Programs

Campbell, D. T. (1969) *"Reforms as Experiments"* American Psychologist, 24, 409-429

Campbell, D. T. (1982) *"Experiments as Arguments"* Knowledge: Creation, Diffusion, Utilization, vol. 3, March 1982, 327-337.

Chen, H-T. (1990) *"Theory Driven Evaluations"* Sage Publications

Chen, H-T. and Rossi, P. H. (1987) *"The Theory Driven Approach to Validity"* evaluation and Program Planning, 10, 95-103.

Christie, N. (1997) *"Four Blocks Against Insight: notes on the oversocialization of criminologists"* Theoretical Criminology, Vol 1 (1) 13-23

Clarke, R.V. (1983). *"Situational Crime Prevention: Its Theoretical Basis and practical Scope"* In: Crime and Justice: A Review of Research, vol. 4, edited by Michael Tonry and Norval Morris. Chicago and London: The University of Chicago Press.

Clarke, R.V. and Weisburd, D. (1994). *"Diffusion of Crime Control Benefits: Observations on the Reverse of Displacement"* In R.V. Clarke (ed) Crime Prevention Studies, vol 2, Monsey, New York: Criminal Justice Press.

Cochrane, A. L. (1972) *"Effectiveness and Efficiency: Random Reflections on Health Services"* London: Nuffield Provincial Hospitals Trust, 1972, reprinted in 1989 in association with the British Medical Journal.

Colledge, M., Collier, P. and Brand, S. (1999) *"Programmes for Offenders: guidance for evaluators"* Crime Reduction Programme Guidance Note 2. London: Home Office.

Connell, J. P., Kubish, A. C., Schorr, L. B. and Weiss, C. H. (1995) (eds) *"New Approaches to Evaluating Community Initiatives: Concepts, Methods and Contexts"* The Aspen Institute.

Cook, T. D. (1997) *"Lessons Learned in Evaluation Over the Past 25 Years"* In Eleanor Chelimsky and William R. Shadish (eds) 'Evaluation for the 21st Century' Sage Publications

Cook and Campbell (1979) *"Quasi-experimentation: design and analysis issues for field settings"* Chicago: Rand McNally.

Cornish, D. and Clarke, R.V. (1987) Social Science in Government: the case of the Home Office Research and Planning Unit *In M. Bulmer (ed) Social Science Research and Government, Cambridge University Press.*

Chelimsky, E. (1997) *"The Political Environment of Evaluation and What It Means for the Development of the Field"* In Eleanor Chelimsky and William R. Shadish (eds) 'Evaluation for the 21st Century' Sage Publications.

Crime and Disorder Act (1998) (c.37) ISBN 0 10 543798 0 London: The Stationery Office.

Crime Prevention Studies Series, Edited by R. V. Clarke, xxxxMonsey, NY: Criminal Justice Press

Desautels, L. D. (1997) *"Evaluation as an essential Component of 'Value-for-Money'"* In Eleanor Chelimsky and William R. Shadish (eds) 'Evaluation for the 21st Century' Sage Publications.

Dhiri, S. and Brand, S. (1999) *"Analysis of Costs and Benefits: guidance for evaluators"* Crime Reduction Programme Guidance Note 1. London: Home Office.

Diez, L. (1995) *"The Use of Call Grading: how calls to the police are graded and resourced"* Police research Series, Paper 13, London: Home Office

Doherty, J. (2000) *"Housing: linking theory and practice"* In Davies, H. T. O., Nutley, S. M. and Smith, P. C. (eds) "What Works? Evidence based policy and practice in public service" Bristol: The Policy Press.

Dowds, L. and Mayhew, P. (1994) *"Participation in Neighbourhood Watch: Findings from the 1992 British Crime Survey"* Home Office Research and Statistical Department, Research Findings No. 12, London: Home Office.

Eck, J. (1997) *Preventing Crime at Places* in Sherman, L., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P. and Bushway, S. (1997) *Preventing Crime: What Works, What Doesn't and What's Promising* A report to the United States Congress. US Department of Justice.

Eck, J. E. and Spelman, W. (1987) *"Problem Solving: Problem-Oriented Policing in Newport News"* Police Executive Research Forum, Washington DC.

Farrington, D. (1983) *"Randomized experiments on Crime and Justice"* In Michael Tonry and Norval Morris (eds) Crime and Justice: An Annual review of Research, vol. 4 The University of Chicago Press, Chicago and London.

Felson, M. and Clarke, R. V. (1999) *"Opportunity makes the Thief: practical theory for crime prevention"* Police research Series, paper no. 98, London: Home Office.

Field, S. (1990) *"Trends in Crime and their Interpretation: A Study of Recorded Crime in Post War England and Wales"* Home Office Research Study 119. London: HMSO.

Forrester, D., Chatterton, M., and Pease, K. with the assistance of Robin Brown (1988) *"The Kirkholt Burglary Prevention Project, Rochdale"* Crime Prevention Unit Paper 13. London: Home Office.

Forrester, D., Frenz, S., O'Connell, M., and Pease, K. (1990) *"The Kirkholt Burglary Prevention Project: Phase II"* Crime Prevention Unit Paper 23. London: Home Office.

Fulbright-Anderson, K., Kubish, A. C. and Connell, J. P. (1998) (eds) *"New Approaches to Evaluating Community Initiatives: Volume 2 – Theory, Measurement and Analysis"* The Aspen Institute.

Garland D. and Sparks, R. (2000) *"Criminology, Social Theory and the Challenges of Our Times"* British Journal of Criminology, 40 (2) 189-204

Goldblatt, P. and Lewis, C. (1998) (eds) *"Reducing Offending: an assessment of research evidence on ways of dealing with offending behaviour"* Home Office research Study, 187, London: Home Office.

Goldstein, H. (1979) *"Improving Policing: A Problem Oriented Approach"* Crime and Delinquency (April) 236-258.

Goldstein, H. (1990) *"Problem-Oriented Policing"* New York: McGraw Hill.

Government Performance and Results Act (1993) Public Law 103-62, United States of America

Hauritz, M., Homel, R.,McIlwain, G.,Burrows, T. and Townsley, M. (1998) *"Reducing Alcohol related Harm Through Local Government Safety Action Projects: The Queensland Experience"* Contemporary Drug Problems, 25, 511-551

Her Majesty's Inspectorate of Constabulary (1998) *"Beating Crime"* London, Home Office

Her Majesty's Inspectorate of Constabulary, (2000) *"Calling Time on Crime"* London: Home Office

Hirst, J. (1994) *"Report on the National Police Research and Development Conference"* Ad Hoc Police Research Group report, London: Home Office

Hood, C., Scott, C., James, O., Travers, T and Jones, G. (1998) *"Bureaucratic Gamekeeping"* www.ncl.ac.uk/politics/whitehall/briefing8

Home Office (1993) *"A Practical Guide to Crime Prevention for Local Partnerships"* prepared for the Home Office by Crime Concern. London: Home Office.

Home Office (1999) *Reducing Crime and Tackling its Causes – a briefing note on the Crime Reduction Programme.* London: Home Office

Home Office (2000) *"Recorded Crime Statistics England and Wales, October 1998 to September 1999"*Home Office Statistical Bulletin, 1/00, London: Home Office.

Hough, M. and Tilley, N. (1998) *"Getting the Grease to the Squeak: Research Lessons for Crime Prevention"* Crime Prevention and Detection Series Paper 85, London: Home Office

Husain, S. (1988) *"Neighbourhood Watch in England and Wales: A Locational Analysis"* Crime Prevention Unit Paper 12. London: Home Office.

Kazi, M. A. F. (1997) *"Single-Case Evaluation in British Social Services"* In Eleanor Chelimsky and William R. Shadish (eds) 'Evaluation for the 21st Century' Sage Publications

Kelling, G. L. and Bratton, W. J. (1998) *Declining Crime Rates: Insiders' Views of the New York City Story* The Journal of Criminal Law and Criminology, Vol. 88, No. 4

Kennedy, D., Piehl, A., M. and Braga, A. A. (1996) *"Youth Violence in Boston: Gun Markets, Serious Youth Offenders, and a Use-Reduction Strategy"* Law and Contemporary Problems, 59, (1), 147-196

Kennedy, D. (1999) *Research for Problem Solving and the New Collaborations* in Viewing Crime and Justice from a Collaborative Perspective: Plenary Papers of the 1998 Conference on Criminal Justice Research and Evaluation. US Department of Justice.

Kinsinger, P. (1999) *"The Roles of Practitioner-based and Scholarly Professional Associations in Enhancing Public Safety Through Principled Approaches to Corrections"* Paper presented at the American Society of Criminology meeting, Toronto, 1999.

Laycock, G. K. (1977) *"Absconding from Borstals"* Home Office Research Study No. 41 London: HMSO

Laycock, G. (1985) *"Property Marking: A Deterrent to Domestic Burglary?"* Crime Prevention Unit Paper 3. London: Home Office

Laycock, G.K. (1992) Operation Identification or the Power of Publicity?' Clarke, R. (Ed.) Crime Prevention: Successful Case Studies New York: Harrow and Heston.

Laycock, G. K. (1996) 'Rights, Roles and responsibilities in the Prevention of Crime' In 'Preventing Crime and Disorder – targeting strategies and responsibilities' Trevor Bennett (Ed). University of Cambridge Press.

Laycock, G. (2000) *"Becoming More Assertive About Good Research: what police practitioners should ask of the academic community"* Subject to Debate: Police Executive Research Forum, 14 (7)

Laycock, G. (2001) *"Hypothesis Based Research: the repeat victimisation story"* Criminal Justice, Volume 1 forthcoming

Laycock, G., and Tilley, N. (1993) *"Policing and Neighbourhood Watch: strategic issues"* Crime Prevention and Detection Paper 60, London: Home Office.

Laycock, G. and Tilley, N. (1995) *Implementing Crime Prevention Programmes*, in M. Tonry and D. Farrington (Eds) *Building a Safer Society*, Crime and Justice: A Review of Research. Chicago: University of Chicago Press.

Leigh, A., Mundy, G., and Tuffin, R. (1999) *"Best Value Policing: making preparations"* Police Research Series, Paper 116, London: Home Office.

117

Lloyd, S., Farrell, G. and Pease, K. (1994) *"Preventing Repeated Domestic Violence: a demonstration project on Merseyside"* Crime Prevention Unit Paper 49. London: Home Office.

National Institute of Justice, (1997) *Criminal Justice Research Under the Crime Act – 1995 to 1996* US Department of Justice.

National Institute of Justice (1999) *"Viewing Crime and Justice from a Collaborative Perspective: Plenary Papers of the 1998 Conference on Criminal Justice Research and Evaluation"* National Institute of Justice, 1999

Nutley, S. and Davies, H. (2000) *"Making a Reality of Evidence-Based Practice"* In Davies, H. T. O., Nutley, S. M. and Smith, P. C. (eds) "What Works? Evidence based policy and practice in public service" Bristol: The Policy Press.

Nutley, S. and Webb, J. (2000) *"Evidence and the Policy Process"* In Davies, H.T.O., Nutley, S. M. and Smith, P. C. (eds) "What Works? Evidence based policy and practice in public service" Bristol: The Policy Press.

Paparozzi, M. A. (1999) *"The Roles of Scholars, Practitioners, Politicians and the General Public in Establishing Correctional Goals and Influencing Practice"* Paper presented at the American Society of Criminology Meeting, Toronto, 1999.

Pawson, R. and Tilley, N. (1997a) *Realistic Evaluation.* London: Sage.

Pawson, R. and Tilley, N. (1997b) *An Introduction to Scientific Realist Evaluation"* In Eleanor Chelimsky and William R. Shadish (eds) 'Evaluation for the 21st Century' Sage Publications

Petersilia, J. (1991) *"Policy Relevance and the Future of Criminology – ASC Presidential Address, 1990"* Criminology, 29 (1)

Pollitt, C. and Summa, H. (1997) *"Performance Auditing: Travellers' Tales"* In Eleanor Chelimsky and William R. Shadish (eds) 'Evaluation for the 21st Century' Sage Publications

Read, T. and Tilley, N. (2000) *"Not Rocket Science? Problem-solving and crime reduction"* Crime Reduction Series Paper 6, London: Home Office.

Rosenbaum, D. (1998a) *"A Critical Eye on Neighbourhood Watch: does it reduce crime and fear?"* in Hope, T. and Shaw, M. (eds) *Communities and Crime Reduction* London: HMSO

Rosenbaum, D. (1998b) *"The Changing Role of the Police: Assessing the Current Transition to Community Policing"* Jean-Paul Brodeur (ed) How to Recognise Good Policing: Problems and Issues. Police Executive Research Agency, Sage Publications.

Rothschild (1971) *Framework for Government Research and Development* Cmnd. 4814 Civil Service Department, London, HMSO.

Sanders, J. R. (1997) *"Cluster Evaluation"* In Eleanor Chelimsky and William R. Shadish (eds) 'Evaluation for the 21st Century' Sage Publications

Schorr, L. (1997) *"Common Purpose"* New York: Doubleday

Schorr, L. (1999) *Replicating Complex Community Partnerships* in Viewing Crime and Justice from a Collaborative Perspective: Plenary Papers of the 1998 Conference on Criminal Justice Research and Evaluation. US Department of Justice.
Shadish, (1997)

Sherman, L., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P. and Bushway, S. (1997) *Preventing Crime: What Works, What Doesn't and What's Promising* A report to the United States Congress. US Department of Justice.

Scriven, M. (1994) *"The Final Synthesis"* Evaluation Practice, 15, 367-382.

Scriven, M. (1997) *"Truth and Objectivity in Evaluation"* In Eleanor Chelimsky and William R. Shadish (eds) 'Evaluation for the 21st Century' Sage Publications

Stake, R. E. (1997) *"Advocacy in Evaluation: A Necessary Evil?"* In Eleanor Chelimsky and William R. Shadish (eds) 'Evaluation for the 21st Century' Sage Publications

Tilley, N. (1993) *"After Kirkholt: Theory, Methods and Results of Replication Evaluations"* Crime Prevention Unit Paper 47. London: Home Office.

Travis, J. (1998) *"Research and Practice: Can We Talk?"* Plenary presentation to the American Society of Criminology, November 1998, Washington DC, in *"Selected Speeches of Jeremy Travis"* National Institute of Justice, 2000

Walker, R. (2000) *"Welfare Policy: tendering for evidence"* In Davies, H. T. O., Nutley, S. M. and Smith, P. C. (eds) "What Works? Evidence based policy and practice in public service" Bristol: The Policy Press.

Weisburd, D. (1994) *"Evaluating Community Policing: Role Tensions Between Practitioners and Evluators"* In Rosenbam, D. P. (ed) *The Challenge of Community Policing: Testing the Promises* Thousand Oaks,CA: Sage

Weisburd, D. (2000) *"Compstat and Organizational Change: a national assessment"* Paper presented at the National Institute of Justice Annual Conference on Criminal Justice Research and Evaluation, Washington DC, 2000

Weiss, C. H. (1972) *"Evaluation Research: methods for assessing program effectiveness"* Englewood Cliffs, NJ: Prentice Hall.

Weiss, C. H. (1995) *"Nothing So Practical as a Good Theory: exploring theory-based evaluation for comprehensive community initiatives"* In Connell, J. P., Kubish, A. C., Schorr, L. B. and Weiss, C. H. (eds) *"New Approaches to Evaluating Community Initiatives: Concepts, Methods and Contexts"* The Aspen Institute.

Weiss, C. H. (1998a) *"Evaluation"* second edition, Prentice Hall, New Jersey

Weiss, C. H. (1998b) *"Have We Learned Anything New About the Use of Evaluation?"* American Journal of Evaluation, 19 (1) 21-33.

Weisner, E. (1997) *"Evaluation, Markets and Institutions in the Reform Agenda of Developing Countries"* In Eleanor Chelimsky and William R. Shadish (eds) 'Evaluation for the 21st Century' Sage Publications

Welsh B. C. and Farrington, D. P. (1999) *"Value for Money? A Review of the Costs and Benefits of Situational Crime Prevention"* 39, 345-368

Wiles, P. (1999) Speech given at Centre for Crime and Justice, King's College, London, November 1999. Text available on Home Office website www.homeoffice.gov.uk

Yin, R. K. (1984) *"Case Study Research: design and methods"* Applied Social Research Methods Series, vol. 5. Sage Publications.

Yin, R. K. (1999) *"Rival Explanations as an Alternative to Reforms as Experiments"* Cosmos Corporation.

Yin, R. K. (2000) *"Using the Rival Explanations Method for Evaluating Complex, Community-based Programs"* Paper presented at the Annual Conference on Criminal Justice Research and Evaluation, Washington DC, July 2000.

Zahn, M. A. (1998) "Thoughts on the Future of Criminology – Presidential address to the American Society of Criminology, 1998" Criminology, 37 (1), 1-16.

SOCIAL RESEARCH - GETTING IT RIGHT FOR PRACTITIONERS AND POLICY MAKERS

A proposal to the National Institute of Justice, Washington DC

This proposal is intended to lead to the development of a framework within which three key communities can work together more effectively - researchers, practitioners and policy makers. The work will be focused in the criminal justice field, centring on policing.

The main question to be addressed is whether and how policing research can be made more relevant to both police practitioners and major policy makers at local, State and Federal levels.

Although the project will be developed in the policing context the lessons learned should be relevant to other parts of the criminal justice system and to the development of policy and practice in other areas of major governmental concern such as health and education.

Background

There is a developing interest in Western nations generally, but in the USA and UK particularly, in evidence-based policy. The move away from dogmatic policy assertions to a more rational information led system is making fresh demands on the way in which business is conducted across the board. There is an increase in the demand for information on crime at national and local level and for research that can inform and develop both policy and practice. The public is becoming bored with argumentative politicians who do not deliver on their promises of less crime and safer communities. In the UK crime control field there is a real pressure to find our what works, at what cost. The crime figures are predicted to rise; dealing positively with this expectation is becoming an imperative.

It is also fairly clear that despite what has been a significant investment of public money in a huge range of crime prevention programs over a 20 year period, we are no closer to being able to tell policy makers or practitioners what works, why it works, or where it works on anything other than a project basis. We know, for example, that we can reduce gun crime in Boston, or burglary in Huddersfield, but we cannot advise on how to implement project-based good practice on a broad, national or regional basis thus driving the top line crime figures down. These headline figures appear to be driven by demographics and the economy, and remain doggedly outside the control of local and central government activity.

This scenario represents a huge challenge to the research community. If we are to increase the influence of research on policy and practice then we need to do three things better -

we need better to understand the culture of the local police practitioner, and the drivers for crime faced at local level, so as to be able to provide the kind of information needed in an understandable format with the aim of improving the performance of the police and other agencies;

we need better to understand the political context within which policy is formulated if we are to be equally influential in that arena, and

we need, bluntly, to get our act together as researchers and refocus effort on hypothesis based studies which inform policy and practice by providing clear evidence of what works, where, and why.

These points reflect the fundamental lack of congruence between the research, policy and practice agendas. The players are working within different operational contingencies – what is understood and valued within one community is incomprehensible or irrelevant in another. But if government funded research aims to contribute to the public good, then government funded researchers need to begin to operate under a different set of contingencies. Their work needs to be more outcome-focused in delivering the central government agenda, which in the case of policing includes a real contribution to the reduction in crime within an ethically bounded, corruption free structure.

In order to achieve this shift a new set of understandings need to be developed including a greater appreciation of the way in which changes in the behaviour of front line practitioners, reflecting evidence-based good practice, can be encouraged. The development of effective 'levers' including cutting edge communication and training methods; local police performance regimes; budgetary controls and particularly the improved use and analysis of local data can all assist in improving local delivery of services. The bottom line is that the researchers need to go beyond the execution of the funded research programs to consider how the output of their work can affect the governmentally defined outcomes.

This agenda is challenging but it is probably no exaggeration to say that both the US and UK Governments have invested heavily in social, criminal justice research and are continuing to do so. They will rightly expect a handsome payoff from that investment. We will get one shot at demonstrating what we can deliver - and we need to get it right.

How will the work be done?

The project will draw primarily on information in the USA and UK although it will also touch on alternative models in Australia and Northern Europe (probably Sweden and Norway, possibly Germany) depending on the time available.

The process through which police-related research is commissioned; conducted; disseminated and acted upon will be systematically covered in both primary research jurisdictions. The following questions will be addressed –

Commissioning
How is research commissioned?
Are key stakeholders involved? How? At what stage?
Is the work set in a strategic framework? Are we aiming for a long-term build-up of knowledge within a coherent strategy?
Are stakeholders receptive to research? Is it seen as a threat or an opportunity?

Conduct of research
Do we use the most appropriate methodology to achieve the desired outcome?
Is sufficient attention paid to the testing of hypotheses rather than the passive evaluation of existing programs?
When is the randomised control trial appropriate and when not? Can we learn equally useful lessons using other approaches?
Is there a greater role for 'action research' and development projects? If so, how do we maintain the objectivity of the researchers and integrity of the programs?
What can we say about the cost effectiveness and comparative cost of programs? Do social researchers need to work more with economists and other disciplines?

Dissemination of research
How are the results of research best disseminated? And to whom?
Are we missing tricks? Should we be writing for a wide variety of audiences in a range of different formats rather than for eachother?
What is good practice here? Do the demands of the academic community facilitate or hinder the process of putting research results to use?
How do we make use of the lessons from marketing and sales to increase the effectiveness of research findings?

Acting on the research
What can we say about the effect of the research effort in the 'real world'?
How can we ensure better implementation of research results?
What are the key components of a 'good' project?
What 'levers' can be pulled to increase the influence of research?
Most importantly, what can we say about moving from effective project based work to rolling out what works at regional or national level?

These, and other questions arising during the course of the study, will be approached by taking a series of case studies as examples. The 'repeat victimisation story' will be taken as a primary example in the UK. The way in which this work was progressed from a development project to a mainstream police activity will be described. In the USA the development of the COPS program will be taken as an overarching example within which specific studies of crime pattern analysis, problem oriented policing and the Boston gun control project will be considered in detail. The Chicago work will also be studied as an example of a major and highly productive initiative.

This process will involve analysis of solicitation papers and the process and products of resulting research. It will also involve discussions with key academics in the US and UK (eg Mark Moore, David Kennedy, Herman Goldstein, Ron Clarke, Dennis Rosenbaum, Larry Sherman, George Kelling, Wes Skogan, Ken Pease and Nick Tilley), police practitioners in San Diego, New York, Chicago, Boston and Los Angeles and a selection of relevant policy makers ideally at city, State and Federal levels.

The purpose of these contacts and the consideration of information outlined above will be to draw a picture of the contexts within which the current key communities operate. Standard research techniques will be used during this process – structured and semi-structured interviews, questionnaires to police practitioners, and quantitative data where available. The purpose of the exercise is not to develop new research tools but to address some key issues using existing methods. The work will be more of a think piece than an empirically driven exercise although where available hard data will certainly be used. This would include, for example, an analysis of the distribution lists for Home Office research briefing notes and NIJ's Research in Action papers with perhaps a follow-up of a sample of recipients to determine to what use the briefing papers were put.

Proposal output

The proposal should lead to a strategic action plan aimed at increasing the ability of the social research community to deliver relevant, timely, pragmatic research with a significant payoff for policy makers at all levels and for practitioners in the crime control field. It will challenge the current working practices of researchers, practitioners and policy leaders and set out proposals for increasing the mutual understanding of all three communities.

Outcome

The outcome should be an increased understanding of the process through which government funded research can be translated into more effective, outcome oriented policies and practices. This should be cross-cultural and cross-jurisdictional identifying fundamental issues which need to be addressed if government funded social research is to contribute fully to the development of national social agendas.

Gloria K Laycock
Head of Policing and Reducing Crime Unit
London, England
June 1999

4

Acknowledgments

I am grateful to a large number of people who have generously given their time for discussions with me during the course of this fellowship. These include:

Angela Smith	Gil Kerlikowski	Mike Wagers
Bill Sousa	Graham Farrell	Nick Tilley
Brenda Buren	Herman Goldstein	Noel Brennan
Chip Coldren	Jeremy Travis	Phyllis McDonald
Connie Stine	Jim Finkenhauer	Ron Burns
Dan Reynolds	John Buchanan	Ron Clarke
Darrell Stevens	Ken Pease	Sally Hillsman
David Bayley	Kevin Kotsur	Steve Edwards
David Kennedy	Lois Mock	Terry Dunworth
David Weisburd	Marcus Felson	Tom McEwan
Dennis Nowicki	Marlon Defillo	Wes Skogan
Dennis Rosenbaum	Marvene O'Rourke	Winnie Read
Ed Connors	Maureen Casey	
Gil Kerlickowski	Mike Scott	

Since the beginning of the Fellowship on 1 September 1999 I have made the following visits:

Australia: Brisbane; Melbourne; Canberra; Tasmania; Sydney; Perth

Israel: Israeli police in Tel Aviv and Jerusalem

Canada: American society of Criminology meeting and Justice Department Seminar on cost effectiveness in the criminal justice system.

Police Departments in New York, Newark, Montclair, Tempe, Charlotte, Savannah, New Orleans

University of Wisconsin; Rutgers State University, NJ; Boston

Products arising from NIJ Grant No. 1999-IJ-CX-0050

The following papers, book chapters or reports have been, or are in the process of being published with NIJ support through this grant:

Published

Hobbs, L., Farrell, G., Edmunds, A. and Laycock, G. (2000) *"RV Snapshot: UK Policing and Repeat Victimization"* Crime Reduction Research Series Paper 5. London: Home Office

Davies, H., Laycock, G., Nutley, S., Sebba, J. and Sheldon, T. (2000) *"A Strategic Approach to Research and Development"* in H Davies, S Nutley and P C Smith (eds) "What Works? Evidence-Based Policy and Practice in Public Services" London: The Policy Press

Laycock, G. (2000) *"Becoming More Assertive about Good Research: what police practitioners should ask of the research community"* Subject to Debate, Police Executive Research Forum Newsletter, 14, No 7

Accepted and forthcoming

Tilley, N. and Laycock, G. (2000) *"Joining Up Research, Policy and Practice About Crime"* Policy Studies, Volume 21 No.3

Laycock, G. (2000) *"From Central Research to Local Practice: identifying and addressing repeat victimisation"* Management Matters: special edition

Laycock, G. (2001) *"Hypothesis Based Research: the Repeat Victimization Story"* Criminal Justice: The International Journal of Policy and Practice, Volume 1

In preparation

Laycock, G. (2001) *"Methodological Issues in Working with Policy-Advisers and Practitioners"* Chapter for inclusion in Tilley, N. (ed) 'Methodological Issues in Crime Prevention Research and Practice' Crime Prevention Studies Series, Criminal Justice Press.

Laycock, G. and Clarke, R. V., (2001) *"Policing and crime prevention in the UK and US"* for a special issue of the International Journal of Comparative Sociology.

Conference presentations made or invitations accepted

<u>Completed</u>

Griffith University, Brisbane Seminar on UK crime reduction program.

Queensland Police, Brisbane Lecture on "Recent Developments in Crime Prevention" to Queensland Crime Prevention Task Force and Senior Officers Group.

Queensland Police, Brisbane Lecture on "Key issues in Policing"

19th Biennial International Conference on Preventing Crime, Australian Crime Prevention Council

Keynote speaker: "Crime Prevention: Time to Deliver"

Plenary presentation: "Unpacking Crime"

These presentations were followed by a live radio interview.

Australian Institute of Criminology, Canberra Seminar on "Current Research Issues"

Tasmanian Police, Hobart

Lecture on "Current Issues in Policing" to Tasmanian Police

Lecture on "Crime and Crime Prevention" to police, local State politicians and community leaders

These presentations were followed by two TV interviews and one radio interview

Charlotte-Mecklenburg Police Lecture on current issues in crime prevention

PERF, Washington DC, Annual Conference Gary Hayes Memorial lecture

Police Foundation, Washington DC Brown Bag, "Crime Prevention: Time to Deliver"

Tel Aviv, Israel (funded by the Israeli Government) Two lectures to Israeli Police Conference

"Community Policing in the US and UK"
"Crime Prevention Models"

Perth, Australia Environmental Criminology Conference paper on "What's Wrong with Situational Crime Prevention?"

NIJ Research and Evaluation Conference, Washington, DC Moderator, session on "New Approaches to Criminal Justice Research Methodology"

Cost effectiveness in criminal justice seminar organized by the Frazer Institute, Ottawa for the Canadian Government. (Funded by the Canadian Department of Justice)

Accepted but not yet given

American Evaluation Association, Hawaii Paper with Nick Tilley to Criminal Justice Topical Interest Group, arranged by Winnie Reed

American Society of Criminology, San Francisco Discussant on repeat victimization panel arranged by Deborah Lamm-Weisel

Crime Mapping Conference, San Diego Discussant on broken windows panel arranged by Nancy La Vigne