Developing the evidence-base for probation programmes

John Wilkinson
University of Surrey
July 1998

PhD Thesis
Abstract

Developing the evidence-base for probation programmes

This thesis considers how research evidence can support the probation service in its work with offenders, addressing the questions: 'What methods are effective in offender rehabilitation?', 'How can we know these methods are working in practice?' and 'How can research evidence underpin policy and practice in work with offenders?'. The thesis explores whether crime-reduction via work with individual offenders is a feasible aim, ethically and practically, and reviews how this fits with theoretical approaches to understanding crime, and as an aim of the probation service. The characteristics of effective and ineffective methods of intervention as described in research literature are reviewed, revealing a low level of consistency between research-syntheses as to the characteristics of effective interventions. Methods of research synthesis and meta-analysis are critically reviewed. Issues in evaluating offender-interventions are reviewed, focusing on the use of reconviction as an outcome indicator.

The thesis presents research case-studies of five offender-interventions, selected to be representative of probation programmes. The impact of each intervention on reconviction is evaluated using a quasi-experimental methodology. The offending careers of over 700 offenders, in intervention and comparison groups, are analyzed. Three interventions show clear evidence of effectiveness, with a lower proportion of offenders in the intervention groups than in relevant comparison groups being reconvicted: 48 percent, 34 percent (these finding were statistically significant), and 18 percent fewer offenders in the intervention group were reconvicted within one year. In this latter case-study of an evidence-based cognitive-skills training programme, offenders who showed positive change in attitudinal measures were more likely to be reconvicted. The implications of this finding for programme-evaluation methods are discussed. The thesis concludes by considering what these findings convey about effective ways of working with offenders, and how the findings contribute to an evidence-based approach in the probation service's work with offenders.
Contents

Acknowledgements

1. Probation and Rehabilitation
   i. Pressure on the probation service to demonstrate effectiveness
   ii. Probation and rehabilitation of offenders
   iii. The rise of the evidence-based approach in planning and delivering services
   iv. What works: the findings of research syntheses
   v. The current research

2. How can we tell if interventions reduce offending?
   i. Methods of research synthesis
   ii. Problems with interpreting meta-analytic findings
   iii. Programme evaluation methods
   iv. Statistical methods
   v. Reconviction as an indicator of the success of sentencing
   vi. Making sense of reconviction outcomes
   vii. Comparison group issues
   viii. Extending the evidence-base: selecting programmes for evaluation

3. A long-term follow-up of the first Inner London Probation Service Demonstration Unit
   i. The inception of the Demonstration Unit
   ii. Developing methods of supervision: the offence-focused approach
   iii. Summary of Demonstration Unit aims and methods
   iv. Evaluation methods
4. Offence-specific work with offenders who take cars: Ilderton Motor Project

i. Responses to car crime
ii. Motor projects and crime reduction
iii. Car crime and offence specialisation
iv. Making sense of Ilderton's impact on reconviction
v. Reoffending outcomes
vi. Conclusions

5. Probation centres' impact on reconviction: an evaluation of the Sherborne House programme for young offenders

i. Probation centres in probation practice
ii. Probation centre programmes
iii. Sherborne House: background and earlier research
iv. The Sherborne House programme: aims and methods
v. Research design
vi. Results and discussion
vii. Reconviction outcomes
viii. Conclusions

6. Reasoning and rehabilitation: an evidence-based programme for offenders

i. The origin and design of the Reasoning and Rehabilitation
 programme 174
ii. The Reasoning and Rehabilitation programme 177
iii. Evidence for the effectiveness of the Reasoning and Rehabilitation programme 181
iv. The Camberwell Probation Centre programme 183
v. Evaluation method 186
vi. Results and discussion 193
vii. Completing the Camberwell Probation Centre programme 197
viii. Reconviction 204
ix. Intermediate outcome measures and reconviction 213
x. Summary 222
x. Conclusions 225

7. Building the evidence-base: summary and conclusions 233
i. Why the research was carried out: pressure to demonstrate the effectiveness of probation 234
ii. The research: programmes included 235
iii. Evaluation methods 239
iv. Results and discussion 242
v. Reviewing findings in the light of evidence about 'what works' 250
vi. Conclusions 256

References 272
Acknowledgements

This thesis is based on a programme of research I was responsible for carrying out whilst employed as Research Officer for Inner London Probation Service. Many people have supported this work. In particular I must thank my PhD supervisors Nigel Fielding and Alan Clarke, and my collaborative supervisor and manager at the time of this research, Stephen Stanley for their guidance and support. I am also grateful for the support of my managers since leaving the probation service, in encouraging me to write up this thesis, these are Rex Hewitt of Turning Point, and Susan Lonsdale and Jenny Griffin of the Department of Health.

I would like to thank staff of the Inner London Probation Service Research and Intelligence Unit for their help with some of the original research, in particular I must thank Christopher Hignet for his advice on the Camberwell Probation Centre Study, and for work on the reconviction data; David Morgan, for data-input on the Demonstration Unit, Ilderton Motor Project and Sherborne House studies, and preliminary data-cleaning and analysis on these studies, and for carrying out most of the structured interviews in the Camberwell Probation Centre studies; and Jill Whiteman for her help in obtaining literature.

I would like to thank the then managers of each of the programmes included in the research for helping me obtain information, and for their advice on the conduct of the research. These were Peter Harraway, of the Demonstration Unit; Peter West of Ilderton Motor Project; Barry Snelgrove of Sherborne House and Delia Mitchell of Camberwell Probation Centre. I also am grateful to the staff of these programmes for their support, in particular the staff of Camberwell Probation Centre, who were active in collecting data. I must thank the offenders participating in the Camberwell Probation Centre programme, for giving their time. I must finally thank Home Office and Metropolitan police staff for allowing me access to data about offending.
Chapter 1

Probation and rehabilitation

This thesis describes a programme of research, carried out with the aim of adding to knowledge about effective ways of working with offenders. The research takes the form of a set of outcome-evaluations, done primarily to identify how effective a set of probation programmes were in reducing reoffending amongst their target client-groups. These programmes were run by Inner London Probation Service between the early 1980s and early 1990s, a time of significant change for the probation service.

Much remains to be learned about effective ways of working with offenders. The new research reported here seeks to fill some gaps, and to stand as an example of good practice in evaluating projects which seek to change individuals' behaviour. Three themes run through this work:

1. what works? The first theme is concerned with identifying effective ways of working with individual offenders to reduce their offending. The probation service is small compared to other criminal justice agencies, but has a central role in working with offenders. Part of this role is working with offenders to reduce crime.

2. how do we know it works and is working? The second theme is concerned with the origin and quality of the evidence on which assertions about 'what works' are based. It also is concerned with how it is possible to know if programmes are succeeding in reducing crime. This means seeing if effective methods are being used, and monitoring work with offenders to determine its impact on offending.
3. how can 'what works' be implemented? The way research and empirical information influence policy and practice in work with offenders is far from straightforward. This theme involves looking at how the evidence-based approach can be developed in work with offenders.

The evidence-based approach, represented here by the key aims of this thesis: 'what works' and 'how do we know it works, and is working', can be seen as embodying a particular view of the work of the probation service rooted in the language of managerialism, and leading to objectives for probation officers' work not all would accept. Without putting these aims in theoretical and organisational context, there is also a danger the new research could be seen as what Jefferson and Shapland, in their (1994) survey of criminological research, characterise as 'narrowly-focused concrete examinations of particular bits of the criminal justice system.'

The rapid growth of evidence-based approaches whilst presented increasingly as orthodoxy (Sheldon, 1998), can be also criticised because this approach focuses on value-for-money, can be used to undermine professional autonomy, and according to Sacket (1998) can regard as unnecessary that which cannot easily be measured. Measuring the work of the probation service is far from straightforward as Humphrey (1991) has observed. This uneasy relationship between research evidence, and probation policy and practice constitutes a third strand running through the set of evaluative studies described in later chapters, which can be encapsulated as 'how can what works be implemented?' Implementing evidence-based practice can entail changing organisational cultures (Roberts, 1995, Centre for Evidence-Based Social Care, 1998.) The research findings set out in later chapters were disseminated internally in the service in initiatives aimed at raising awareness about 'what works' (see for example Wilkinson, Ed. Stanley 1995.)

This research was initiated at a time in the early 1990s when the probation service was regaining confidence that reducing the offending of individuals supervised by the service was a feasible and legitimate objective. Reducing reoffending had been
reaffirmed as the *raison d'etre* of the service (Inner London Probation Service, 1993.) Research evidence as to which ways of working with offenders were likely to be successful in reducing reoffending was starting to be disseminated within the probation service (see for example Roberts, 1990, McIvor, 1990, Lipsey, 1990.) These and other positive reviews provided impetus for the current research, inspiring the organisation to seek to evaluate impartially mainstream probation programmes in the hope (if not always the expectation), of identifying successful practice in reducing reoffending.

Though the primary aim of this work was to add to knowledge about effective ways of working with offenders the secondary aim, of developing ways of monitoring whether probation programmes achieve desired outcomes, is also significant. Referring to evidence that a given method of intervention should reduce offending is quite different to being able to show that in practice a programme is actually achieving this result. This secondary aim entails developing appropriate outcome measures for probation programmes, generally involving some measure of reconviction, or factor predictive of reconviction.

There exists a significant literature which looks at the effectiveness of interventions in reducing delinquency and offending (see McGuire, 1995, for a review, and Losel, 1995, for a synthesis of structured reviews of this literature.) However relationships are complex between this evidence-base, the probation service's organisational policies, and individual probation officers' work with offenders. Relating research evidence to policy and practice in a professional organisation like the probation service is conceptually difficult. It necessitates analysing theory and research findings in the hybrid fields of criminology and penology; it involves policy analysis and value-based political argument. Furthermore, this discourse takes place in a context of organisational and policy change.
Pressure on the probation service to demonstrate effectiveness

In recent years the probation service, in common with many organisations funded from the public purse, has come under increasing pressure to demonstrate the efficiency and effectiveness of the services it offers. As Thomas (1995) has observed

'a public sector organisation that demonstrates clearly quantified costs but no similar unequivocal benefits must feel insecure.' (Page 25)

Similarly Sutton (1994) comments:

'many observers of the Probation Service are expressing the view that unless the service can quite quickly demonstrate its effectiveness at reducing offending, and that social work based community sentences can add value, then it is unlikely that the service will survive in its current form.' (Page 1.)

This insecurity was exacerbated by legislative changes in the early 1990s. The 1991 Criminal Justice Act introduced significant sentencing reform on a model of 'just deserts', intended to reduce the prison population, and enhance the role of the probation service. However Nellis (1995) describes how within a year the main provisions of the act were reversed, in the face of populist criticism, leading to the introduction of the Criminal Justice and Public Order Bill 1993. Other factors adding to insecurity were the removal of probation training from universities (see for example Downes, 1997), and the possibility in the background of mergers with the prison service (see for example Fulwood et al, 1997, Guardian, 17 July, 1997.)

There has been an increased focus on accountability for the content and outputs of work, conveyed via the HM Probation Inspectorate, and the Home Office, and embodied in prescriptive guidance such as National Standards for Supervision in the Community (Home Office, Department of Health and Welsh Office, 1995)
The need to demonstrate cost effectiveness has increasingly been a dominating issue in probation, leading to pressure to articulate outputs in the form of Key Performance Indicators (Home Office, 1992); to obtain better information about work inputs, outputs and costs via the Financial Management Initiative (see Humphrey, 1991, for an account of this); and to moves to obtain information about the unit cost of work via the Resource Management Information System. There has also been pressure to show the effectiveness of the outcome of work with offenders, in terms of impact on reoffending (see for example Audit Commission, 1989.)

Increasingly the work of publicly funded agencies is monitored in terms of the outcomes achieved for service users (see for example Department of Health, 1995, or see Williams and Webb, 1992, for an influential introduction to the outcome funding approach.) The probation service is itself a funding organisation, spending at least 7 percent of its revenue budget on partnership arrangements to deliver services (ACOP 1997.) Thus probation services have become increasingly accustomed to monitoring outcomes achieved for probation clients, by voluntary sector organisations. This has increased the desire for information about outcomes of probation programmes. There has also been an increased focus on monitoring the quality of services offered to potential users (see for example Alaszewski and Manthorpe, 1993.)

A related issue is the rise of the evidence-based approach in health and social care (see Sacket, 1996, Sheldon, 1998), and latterly in work with offenders (Underdown, 1998.) This is significant and is discussed below. As the Centre for Evidence Based Social Care (1998) observe, evidence-based approaches have a long history in social care. Similarly in criminal justice, research evidence has long been linked to policy-making. Particularly pertinent is the work of Mannheim and Wilkins in the late 1950s, whose use of prediction techniques to make sense of the outcome of borstal training led to the establishment of what is now the Home Office Research and Statistics Directorate. However the evidence-based approach implies a need for more systematic relation of research findings to policy and practice (though this can be problematic, see Tizzard, 1989, Pitts, 1992.)
The mid-'90s were a time of increased confidence on the part of the probation service that it could be effective in reducing offending, (see for example Underdown, 1995). Reducing offending became firmly established as a primary goal of the service (Home Office, 1992), and was increasingly embraced as a way of maintaining the service's long-standing concern with helping offenders achieve rehabilitation (see for example Raynor and Vanstone, 1994.) These pressures for information about the effectiveness of work with offenders, coupled with renewed optimism that reducing offending might be possible, came at a time when the probation service, in common with other organisations in the public sector at this time, was becoming a more formally managed organisation (see Stratham and Whitehead, 1992; though the development of management structures was not entirely uncontested, see for example McWilliams, 1992.)

Increased demands for monitoring and effectiveness information, coupled with a management infra-structure led to the development of a research infra-structure in probation and recruitment of research and information staff, manifested by NPRIE, a body made up of social researchers, library and information workers, and statisticians and information technology workers. Much of the work of these staff was concerned with providing management information to meet Home Office requirements, to meet the needs of the new management structure, and to implement the information systems needed to provide this information (see for example Home Office / Probation Service Information Systems Working Group, 1993.)

A significant amount of work to evaluate the effectiveness of work with offenders was carried out, but very little probation-based research was of high quality or was published (see Ellis and Underdown, 1998.) Whilst Underdown (1998) emphasises the importance of an evidence base for work with offenders to reduce their offending, surveys of evaluative research show very little high quality evaluative research into probation work with offenders has recently been carried out (Ellis and Underdown, 1998, Hedderman and Sugg, 1997.)
Probation and rehabilitation of offenders

A significant assumption underlying probation is the idea it should reduce crime, thereby benefiting society, and benefiting the offender by helping her or him out of a life of offending (see for example a Association of Chief Officers of Probation document with the self-explanatory title 'Probation Works and Everybody Benefits', ACOP, 1994.) Over the past 25 years there has been considerable debate as to how well founded are these assumptions.

The probation service was largely founded on a paradigm of offending being characterised as pathological behaviour, with offenders being offered 'curative social work' (see e.g. Donnan, 1961). Raynor and Vanstone (1994) attribute the adoption of psychotherapeutic approaches to a need by probation officers to acquire a professional knowledge-base, to help justify professional status. Through the 1950s and 1960s social work expanded as a profession, and was adopted as the professional base of the probation service. It then seemed a truism that social work which was well intentioned, carried out by trained practitioners, and which aimed to improve people's lives, must be successful in achieving this. Sheldon, 1994, gives an account of this era as relates to work with offenders. Bottoms and McWilliams (1978), observed:

Most probation officers, and certainly almost all policymakers, have always tacitly assumed that the advice and assistance offered [by probation] did have an effect in steering at least a proportion of their charges away from criminal acts. (page 160)

In the language of targets and objectives of the 1990s, protecting the public from crime, reducing offending and rehabilitating offenders are clearly established as aims for the probation service, although rehabilitation is now less conceived as 'curing' offending, and more seen within the probation service at least, as being restoration to full citizenship (Nellis, 1995.) However the assumption that probation can reduce crime by working with offenders, and the legitimacy of this as a goal for the service, has been challenged on the grounds, that:
- treatment / intervention to reduce offending does not work;
- work with convicted offenders (even if effective) cannot appreciably reduce the amount of crime experienced by the community;
- crime rates are the result of deep social forces, not amenable to change by superficial intervention;
- most crime is an ideological construct, better understood as 'deviancy';
- coercing people to change denies their free will, and is not ethical.

Taken together these criticisms provide powerful arguments against offender treatment, which remain influential. These criticisms led to the probation service's disenchantment with rehabilitation. The next five sections examine and challenge each of these assertions.

1. Treatment or interventions to reduce offending do not work

The 'treatment' model of probation work has been criticised with most effect on the grounds treatment does not work. As a social work-based profession which works with offenders, probation was dealt a double blow in the 1970s. Research into the effectiveness of social work showed that the psychodynamic methods to which the profession was wedded were ineffective (see for example Fischer, 1976, 1973). Likewise research into the effectiveness of work with offenders found nothing seemed to reduce crime (see for example Martinson, 1974, Brody, 1976.) In particular Folkard's (1976) study, which showed that more intensive probation did not lead to less offending, had a major influence on the probation service. The phrase 'nothing works' in respect of the treatment of offenders was coined by Martinson (1974) in a widely cited review of 231 research studies, which posed two questions:

'do all of these studies lead us irrevocably to the conclusion that nothing works, that we haven't the faintest clue about how to rehabilitate offenders and reduce residivism? And if so what shall we do?' (page 48)
Martinson's answer to the first of these questions was, in fact, more equivocal than the statement 'nothing works', which his conclusion is usually represented as having been:

'these data... give us very little reason to hope that we have in fact found a sure way of reducing recidivism through rehabilitation. This is not to say we found no instances of success or partial success; it is only to say these instances have been isolated, producing no clear pattern to indicate the efficacy of any particular method of treatment.' (page 49)

The findings on which Martinson (1974) based his conclusions are hardly more negative than those of more recent research (for example Lloyd et al, 1994), or more recent reviews (for example Losel, 1995.) But at the time these findings shocked the probation service, and had a major influence in moving the service away from the goal of reducing crime.

The most influential rebuttals of the 'nothing works' message are rooted in new research (e.g. Ross et al, 1988), and reinterpretations of existing research evidence (e.g. McIvor, 1990.) These show that a large number of interventions with offenders do reduce offending. No recent large scale reviews have reached negative conclusions. Lab and Whitehead (1988), and Whitehead and Lab (1989), have been portrayed as 'supporting a very firm version of 'nothing works'' (Andrews et al 1990, page 371). However Lab and Whitehead (1990) reject this portrayal, citing studies which show treatment is effective.

Even when the 'nothing works' argument was at the peak of its acceptance, research and literature reviews were being published showing that some interventions with offenders did reduce offending; see for example Palmer (1975, in Palmer, 1980), Jones (1981), Walker (1983) Stanley and Goldberg (1984), Priestley and McGuire (1986), Raynor (1988), Maitland and Keegan, (1989). Bottoms and McWilliams, 1979 offered "straws in the wind" (page 175) of evidence some interventions were associated with reduced offending. Martinson (1979) recanted his earlier more pessimistic views. Indeed Martinson's 1974 review included studies, deemed methodologically sound, which reduced recidivism. For example Massimo, (1963,
reported in Martinson, 1974) found a decline in recidivism rates amongst offenders treated to 'pragmatic psychotherapy' including a focus on vocational problems in a community setting.

From the mid-1980s onwards a stream of positive reviews were published (most notably Andrews, 1990, Lipsey, 1991, Sheldon, 1994.) By the early 1990s commentators referred to the 'end of an era' (of 'nothing works', Pitts, 1992) and Mair (1994) could declare 'nothing works' was in retreat. Bottoms (1995) refers to a 'new orthodoxy', that interventions with offenders could reduce their offending. Advocates for this apparent new consensus amongst researchers advised probation programmes be planned along the lines of 'principles of what works' (McGuire, 1995.) The principles of effective offender-programmes, proposed by various reviewers, are described in a later section.

2. Intervening with individual offenders cannot reduce the amount of crime experienced by the community

There are strong arguments that no interventions with individuals convicted of crime can have a significant effect on crime rates experienced by the community, based on the fact that most crimes do not result in conviction. Barclay (1993) estimates only 3 percent of recorded offences lead to a successful prosecution. Pitts (1993) reports the comment of a senior police officer that to expect the police to prevent crime was like expecting the staff of a hospital emergency ward to stop road accidents. And that:

'It would be remarkable if anything that the police, the courts or social workers and probation officers did with, for or to, those few apprehended offenders they meet had much impact on the phenomenon of crime.' (Pitts 1992 p. 141)

From this it follows that it does not much matter what is done with an offender once apprehended, as this can have only a negligible effect on overall levels of
crime (see for example Bright, 1997.) Similarly the notion prison can reduce crime has been disputed on the grounds incarceration can only temporarily interrupt the offending careers of the minority of offenders who are apprehended and successfully prosecuted (Farrington and Langan, 1992).

However there are two ways in which successful rehabilitation might significantly reduce crime experienced by the community. Firstly a relatively small number of men are responsible for a large proportion of crime. Nearly 60 per cent of all court appearances before the age of 40 were attributed to about 20 per cent of offenders, or 7 per cent of the male population (Home Office, 1997). Graham and Bowling (1995) report that 3 percent of offenders are responsible for about a quarter of all offences. In the Cambridge Study of Delinquent Development 6 percent of the sample of 400 accounted for half of convictions (reported in Tarling, 1993). Of males in the Home Office's 1953 cohort, 7 percent accounted for 65 percent of the convictions amassed by the cohort (reported in Tarling 1993). A third of males in this cohort were convicted at least once, meaning two thirds were not convicted, so the chronic offenders represent 18 percent of offenders. Tarling (1993) concludes:

'18 percent accounting for 65 percent of convictions is still an important disproportionality but with much less impact than 7 percent.' (Page 43).

While it is a moot point whether these offenders are responsible for a similarly large proportion of unrecorded offending, any reduction in the offending of this group, could have a significant impact on crime experienced by the public. Losel, 1995, offers a cost-utility analysis for German social-therapeutic prisons, showing that financial savings result when offenders have fewer prison sentences over the course of their lives, as would happen with treatment effect sizes of the order of a 10 percent reduction in the proportion of offenders reconvicted in a given time.

Secondly, and as importantly, successful rehabilitation might spare society some of the most serious offences. Barclay (1993) reports that the more serious an offence is, the more likely it is to lead to a conviction. Violent and sexual offences are
relatively rare and committed by a relatively small number of 'dangerous' offenders. The probation service supervises many of these people in the community on life-licence and conditional release, and so has a role in protecting the public from serious crime (see for example HMIP 1995.)

These findings show that the criminal justice system, and hence the probation service, can reach a large proportion of the people who commit crime, especially serious crime; and choices exist as to how to work with these offenders, which may have a significant impact on the community's experience of crime.

3. Crime rates are the result of deep social forces which cannot readily be changed

McIvor (1990) cites a reason for the shift away from rehabilitationism as being rising crime rates, which led to public and political concern that a hard line should be taken with offenders. Crime rates have been shown to be linked to unemployment, and to cycles in the economy (Wells, 1995.) Downes (1997) states that social, economic and cultural sources of crime are of much more causal significance than the operation of the criminal justice system. Bright (1997) adds that crime levels relate to the number of young males in the population, and the extent to which people feel they have a stake in society. This leads to the conclusion that whilst treatment may yield gains for individuals, significant inroads into crime as experienced by the community can only come from changes in societal conditions. As Bottoms and McWilliams (1979) observe:

'the treatment model applied is one of individual treatment, while many of the assumed causes of crime are social: treatment-orientated criminology has never learned the lesson of social medicine that better drains may be worth scores of doctors.' (page 161.)

However this is no rationale for not treating disease (or offending.) Andrews (1994, 1995), presents strong evidence that indicators of social deprivation are weaker
predictors of individuals' offending than are psychological factors. Andrews (1995) does not deny that societal and environmental factors are important sources of crime, just that individual internal factors are stronger predictors.

If social deprivation is the main cause of offending, this does not explain why post-war improvements in living standards internationally, rather than leading to a fall in crime were associated with increasing offending. This paradox led to what Young (1994) terms the 'etiological crisis' responded to by policy-makers initially by the provision of more police and more prisons.

4. **Crime is a social or ideological construct, better understood as 'deviancy'.**

In the early 1970s a strand of sociological thought came to question whether crime was a problem as popularly conceived, arguing instead that crime is a construct revealing as much about the state and society, as it does about the people labelled as criminals. Young (1994) summarises these approaches as 'left idealist', including deviancy theory, critical and radical criminology, labelling theory and subcultural theories of crime. In the interest of brevity this grouping of quite distinct approaches to understanding offending may verge on caricature (see Downes and Rock, 1982, for a summary of these theoretical approaches.) But generally these approaches characterise most crime as an ideological construct, which suits the powerful in society; whereby crime is a consequence of poor people's struggle to live, and of rich people's greed (though these latter forms of victimising behaviour are less likely to be treated by the system as criminal, see for example Box, 1983.) According to these approaches crime statistics are a product of police preferences and prisons are full of working class people because their behaviour is defined as criminal. Deviancy theorists thus see crime figures as unscientific, or at best as measures of diverse and complex social interactions. For example Bottomley and Coleman (1981) refer to the amount of discretion and selectivity the police have in deciding whether an event constitutes a crime, and also note that certain classes of
individuals are more likely to be targeted by the police. Taylor, Walton & Young (1975) state

'an examination of those [official] statistics reveals the highly patterned, and indeed the class nature, of society and its law enforcement (and something of the patterned nature of the crimes committed).'
(Page 33).

Given these views of the nature of crime, with many theorists taking the position that to characterise crime as a problem is to misunderstand social phenomena, (see for example Hester and Eglin, 1992), it is not surprising that what Young (1994) terms 'left idealists' have paid little attention to rehabilitation. New deviency theorists (see for example Taylor, Walton and Young, 1973), reacting against correctionalist criminology had no hesitation, initially, in dismissing the notion crime was based on individual pathology and therefore amenable to individualised treatment. Whilst many proponents of 'new deviency' modified their initial views (see for example Young, 1975), probation officers steeped in this type of thinking (as many were as part of their training) are unlikely to see their role as being to challenge individuals' offending.

There are two main criticisms of radical deviency and labelling theories, relevant to practitioners charged with doing something about crime. Firstly these approaches appear to deny the reality of much crime, whereby most victims as well as perpetrators are working class. These approaches play down the harm done to victims of crime, who are often powerless and often women. Secondly these approaches are top-down, invoking an all powerful state and media which are able to create and consign people to sub-cultures, and control how the public thinks of crime. But how people think of crime is unlikely to be entirely governed by the media, and as Schwartz (1997) states:

'we can talk about corporate crime all that we want, but people are just plain more worried about being raped or mugged than being killed by cigarettes, polluted air, or a lack of industrial safety'.

Young (1994) sets out a 'left realist' approach to crime, which accepts the common
view that crime is a problem, seeing it as a process involving victim and perpetrator. According to this view crime is defined in a dynamic social context, whereby the community can recognise 'new' crimes, becoming more intolerant of violence against women, child abuse, racial violence and drink-driving. The causes of crime may include relative deprivation (perhaps interpreted through sub-cultures), but also include individual predisposition. Left realist approaches to crime admit the possibility and legitimacy of rehabilitative work with offenders, and chime with probation officers' experience of the harm done to victims of crime.

Psychological approaches to understanding crime similarly locate crime in an interaction of individual and environment, concentrating more on internal factors (see Farrington, 1992), though some may incorporate sociological theories (e.g. Elliot, 1985.) According to Andrews (1994) crime describes a category of human behaviour that can be defined with a high degree of cross cultural agreement. Criminal behaviour can be explained by general theories of human behaviour, which also suggest ways of responding to crime. These theories and suggested responses are increasingly influential over probation practice (see for example Hedderman and Sugg, 1997), and generate testable predictions relevant to the research reported in later chapters.

5. Rehabilitation is not ethical.

The treatment model of probation, weakened by negative research findings, came to be criticised on ethical grounds. Firstly there is the issue of to whose definition of normality should offenders be rehabilitated. Lewis (1953) offers a vision of offenders made to conform to 'some condition of normality hatched in a Viennese laboratory'. Most influentially, Bottoms and McWilliams (1979) argued correctionalism is unethical as it implies disregard for the basic human rights of self-determination and free choice. In this paradigm correctional approaches are seen as coercing offenders into treatment, denying their ability and right to make free choices. Jones, 1981, notes that rehabilitation implies a disregard for people's
self determination. McWilliams (1992) comments:

"From the very beginning of the probation system the pressures towards coercive reformation have been powerfully manifest...The probation ideal has always stood at the mercy of those determined to see reform as a matter for coercion, and this has been most clearly manifest in the confusion which has surrounded the enforcement of this requirement of orders and licences and the desire that offenders should be reformed.' (page 15)

These ethical criticisms of rehabilitation were also in part rooted in the fact that rehabilitation appeared not to work in practice (see most importantly Bottoms and McWilliams, 1978, though as described above, these authors in fact cited research evidence that seemed to offer support to the idea that helping offenders might reduce their offending.)

However these criticisms of rehabilitative ideals on ethical grounds can be answered (see for example Rotman, 1992). McWilliams and Pease (1990) contend that probation officers have an ethical duty to strive to help offenders become restored to full citizenship (i.e. rehabilitated), because it would be wrong to stand by whilst offenders occasion more harm to themselves or others. Sheldon (1991) argues that probation officers charged with supervising offenders should use methods known to be effective, rather than ineffective, because the first ethical duty of the social worker is to know he or she is using effective methods. If ways of working with offenders are known to exist which make reoffending less likely and do not compromise offenders' human rights, it would be hard to justify not using such methods.

Positive research findings prompted further reassessment of the ethics of coerced rehabilitation. Raynor and Vanstone (1994) develop Bottoms and McWilliams' (1979) non-treatment paradigm to encompass effective intervention, which can empower offenders who want to be helped away from crime, and thereby also help the community. The upshot is that despite dissent by Neary (1992a, 1992b) rehabilitative methods have come to be seen as entirely consistent with probation values (see for example Bhui, 1995.)
This thesis does not revisit the care versus control debate which has for long been a preoccupation of many in the probation service (see for example Fielding, 1984.) The caring professions, as Sheldon (1994) has noted, also have significant controlling powers. Social workers can take away people's children, or have people forcibly taken to mental hospital. Doctors regularly prescribe psychotropic drugs to alter mood and behaviour; psychiatrists can do this without consent; and forensic psychiatrists can decide to divert a person out of the criminal justice system, and hold that person indefinitely in secure psychiatric care (see for example Hinton, 1983 for a discussion of these issues). This context does not diminish the importance of ethical debate about the validity of control in probation practice, although a probation officer seeing someone for a hour a week, to try to help them think about why they keep getting into trouble, does seem qualitatively different to the examples given above. What matters most is whether those probation officers could be making use of techniques and interventions which actually would reduce the extent to which their clients get into trouble with the law.

The impact of anti-rehabilitation arguments on probation

The five reasons why the probation service moved away from rehabilitation are set out separately above, but their impact on the probation service was compounded. Taken together they make a powerful case against offender-treatment, which remains influential. For example Bright (1997) uses the first three of the reasons set out above to 'demolish the myth' that the criminal justice system can have an impact on crime. However the most powerful blows, from the point of view of probation, were delivered by the negative findings of empirical research. This finding, that rehabilitative methods did not appear to rehabilitate, caused a generation of practitioners to turn their backs on research, the impact of this remains in the "folk memory" of probation officers. For example in their paper "A view from the troops", Humphrey and Pease (1991), report that very few probation officers accept that reducing offending is an aim of their work. A typical quotation taken from an interviews with a probation officers is:
"the thing is we can’t change how people offend - the research shows that" (unnamed probation officer quoted in Humphrey and Pease, 1991).

Mair (1991, 1994) likens the phrase “nothing works” to an advertising slogan, its impact arising from its articulation by the media as a simple idea. Martinson’s (1974) review was equivocal in its conclusions, but the more sophisticated messages were lost. For example one intended conclusion was that crime should not be seen as curable disease, because to do so:

'...denies - both the normality of crime in society and the personal normality of a very large proportion of offenders, criminals who are merely responding to the facts and conditions of our society.' (page 49)

On the grounds they could not be treated, Martinson (1974) called for decarceration of low risk offenders, and called for more research to be carried out so:

'we will be in a better position... to judge to what degree the prison has become an anachronism and can be replaced by more effective means of social control' (page 50).

However these conclusions, which were in tune with radical deviancy thinking of the time, were drowned out by a more comprehensible message that 'nothing works'.

The abandonment of rehabilitative aims, which effectively took place in the probation service in the late 1970s suited both conservatives and radicals in the criminal justice system. To some on the left rehabilitation was a bourgeois con-trick. To some on the right the absence of rehabilitative possibilities meant that justice could and should be applied. Andrews et al (1990) refer to a 'marriage of conservative politics and leftist social science' (p 371). Sheldon (1994) describes how the more conservative probation officers were able to pursue their role as officers of the court without needing to worry about trying to rehabilitate offenders, whereas the radicals, taking a 'new deviancy' view of offending were able to
'adopt the offender's view of the justice system as a rough game... at which it is useful to have sympathetic lines persons" (Sheldon, 1994).

Theoretical approaches to crime and deviancy were also important in this move away from rehabilitationism. Many probation officers were trained in 'new deviancy' approaches to criminality. Probation officers who view criminality as being at least in part what Sheldon (1994) terms a 'sort of primitive revolution' are less likely to accept ways of working with offenders (proposed for example by Hedderman and Sugg, 1997) predicated in notions of individual deficits. Furthermore probation officers subscribing to deviancy theory are not likely to accept criteria for the effectiveness of their work which are based on the extent to which an offender classed as 'deviant' continues to commit infractions of the criminal justice system's supposedly arbitrary boundaries.

In the face of anti-rehabilitationist theorising and findings, many in the probation service sought alternative formulations to help conceptualise their work. Most influential was the non-treatment paradigm for probation work offered by Bottoms and McWilliams (1979). This proposed a form of probation practice which acknowledged the parameters imposed by statutory requirements, but within which framework offenders can make genuine choices. They saw the role of probation as being to offer help to offenders, to reach a shared assessment of the offender's position, and together to define a task as the basis for social work action.

If the anti-rehabilitationist position was most strongly rooted in research evidence represented as showing rehabilitation was not possible, the apparent new consensus that intervention can reduce individuals' offending has been the main influence in leading the probation service back towards correctional intervention. However uneasiness remains about treatment methods based on notions of pathological thinking on the part of offenders (see for Neary 1992a 1992b). Raynor and Vanstone, 1994, extend Bottoms and McWilliams' (1979) non-treatment paradigm to encompass the possibility (and ethical imperative) of offering offenders effective help so they can stop committing crime. However these authors continue to express
concern that methods of behaviour modification might be used without offenders' understanding or consent.

The 'what works' approach has rapidly attained the status of orthodoxy for reasons other than the support of research evidence. A virtue of what can be called the 'what works' approach from the point of view of governmental policy-makers, pointed out by Pitts (1992) is that regardless of effects on reoffending, projects conforming to the 'what works' orthodoxy tend to be community-based and cheap compared to custody. Pitts (1992) also notes that this approach is attractive to policy-makers because it locates a sizeable portion of the blame for offending with the offender, rather than on factors such as unemployment or poor housing. However for practitioners the 'what works' approach can be empowering, promising that probation officers can, after all, make a difference to people's lives. From a managerial perspective this approach is attractive as it involves closer control over aspects of practice, in the pursuit of programme integrity (see for example Hollin, 1995.)

The rise of the evidence-based approach in planning and delivering services

The need to base decision-making at all levels on empirical evidence is becoming increasingly well understood in the field of offender-interventions, with calls to 'develop systematically evidence-based practice' (Underdown, 1998). Research-evidence is better integrated into decision-making about policy and patient-treatment in the field of healthcare. Hicks (1997) defines evidence-based health-care as taking place when:

'decisions that affect the care of patients are taken with due weight accorded to all valid, relevant information.'
In Hicks' formulation such decisions may be taken by clinicians, managers or policy makers, with the 'relevant information' including information about the relative effectiveness of various therapies with different patient groups, according to patient preferences, and external factors such as resources. Sacket et al (1996) define evidence-based medicine as being:

the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research. Evidence-based medicine is not restricted to randomised trials and meta-analyses. It involves tracking down the best external evidence with which to answer our clinical questions...

The Centre for Evidence-Based Social Care, (1998) adapt this formula to define evidence-based social care, substituting 'service-users and carers' for 'patients'. However the evidence-based approach is not uncontroversial in medicine, and has been criticised for undermining professional accountability and autonomy. Salvage, 1998, notes that this approach:

'...owes much to the value-for-money culture fostered by Thatcherialism, which at its extreme regarded that which cannot easily be measured as unnecessary or ineffective.' (page 63).

The evidence-based approach in criminal justice and probation

The evidence-based approach has deep roots in criminal justice-related research in Britain. A task of the Home Office Research and Statistics Directorate has been to offer an evidence base for the work of the probation service. This has included advising the probation service on how to work with offenders, for example Probation Circular 77/1995 which required probation services to review their offender-programmes against a list of 'critical success factors'.

The use of research-evidence to inform policy development in a political context is visible in initiatives such as Testing and Treatment Orders for offenders on
probation. This was proposed by the Labour Party, (see Russel, 1996; Labour Party Manifesto, 1997) before becoming policy in the 1998 Crime and Public Order Bill. Research indicated the nature and scale of the problem, showing high levels of drug use amongst people convicted of acquisitive crime. Research synthesis also indicated possible solutions, by showing the likely effectiveness of treatment, including coerced treatment delivered via the criminal justice system (Sowers and Daley, 1993), monitored via drug testing (Hough, 1996.) However translating coerced testing and treatment into probation policy involves many considerations other than evidence of possible effectiveness, including level of public and political support, attitudes of civil servants, and the attitudes of agencies involved in operationalising policy, in this case probation, sentencers and treatment agencies. Their attitudes are mediated by the organisational, political and ethical issues which stem from forcing people into, or rationing, medical treatment on the basis of rule-breaking (Turning Point, 1997.)

How research findings influence policy and practice.

Links between research and social policy are often unclear. Tizzard (1989), expresses this in her pessimistically entitled paper "Research and policy: is there a link?" In this she suggests to influence policy, researchers should identify 'gateways' in the form of policy questions and policy makers amenable to findings presented in appropriate form. Bulmer (1982) also explores these issues.

It is also not enough that research findings are simply known about. Planning developments in service organisation to deliver services takes time. Making these happen takes longer (see Mair 1994). More importantly simply knowing about findings with regard to effective methods is not enough. As Sheldon (1994) points out, using evidence-based methods may require rigorous training on the part of the practitioner, beyond that provided by a brief 'appreciation course' (see for example the Reasoning and Rehabilitation Programme, Ross et al, 1986.) Thus using research evidence may require time and resources.
There is also a problem of time lag between research findings being known amongst researchers, and becoming influential amongst those who deliver services. Probation officers may be about five to ten years behind in their awareness of research findings (Humphrey and Pease, 1991). Similarly in the medical field, where there is little controversy in identifying effective treatments, getting the findings of research acted upon remains problematic. Coiera (1997) observes:

For example, the first trial to show that streptokinase was useful in the treatment of myocardial infarction was published in 1959. Convincing evidence mounted in the early seventies, and the first multi-trial meta-analysis proving its value was published in the early eighties. However, formal advice that streptokinase was useful in the routine treatment of myocardial infarction only appeared in the late eighties... a full thirteen years after a close examination of the published literature would have indicated the treatment's value.

The priority given by government, supported by the medical establishment, to getting research acted upon is illustrated by the National Health Service White Paper (1997), which proposes a National Institute of Clinical Excellence, to draw up guidelines from the latest scientific evidence and ensure they reach all parts of the NHS. However criminal justice-related research lacks theoretical coherence compared to health-care research. Lack of academic consensus within criminology, sometimes even as to the terms of debate (see for example Hestler and Eglin, 1992), dilutes the messages of research.

Lipsey, 1992, observes that the translation of research into policy and practice is plagued with difficulties, but sees this as a technical problem. He goes on to note that in the criminal justice field research is perceived as often irrelevant, always ambiguous and inconsistent. Lipsey suggests this is the fault of research, and that clearer less equivocal findings would be more likely to be taken up. However there are numerous instances, in the criminal justice field, of unequivocal findings not being heeded. Three recent instances are:

- Harsher regimes for young offenders: (see Thornton et al 1984) in the early 80s the “short sharp shock” for young offenders was found to have no practical
value in crime reduction, either in discouraging further offending, or in deterring others from offending. It was not even found to be experienced as punishment by the offenders. This research was glossed over at the time (Jupp, 1989), but these regimes were quietly abandoned. However in the mid-90s the call returned for more “austere” regimes for prisons, and for secure accommodation for young offenders (see for example Nellis, 1995.) Despite further policy-focused research showing the inefficiency of this approach (Audit Commission, 1996) the establishment of STCs followed.

**Electronic tagging:** (Mair and Nee, 1990): this research was a thorough evaluation of a set of pilot schemes. It concluded that regardless of technical problems, tagging was unlikely ever to be cost-effective as a disposal. It also came to the conclusion tagging was most suitable for older offenders who do not go out much in the evenings. Regardless of these findings, tagging remained on the agenda, with further small-scale piloting being carried out (Mair and Mortimer, 1996). Tagging appears to be part of a strategy for reinforcing public and sentencers' confidence in community sentences, embodied in the 1998 Crime and Public Order Bill, by automatising them via the use of 'objective' measures such as tagging and drug-testing, rather than via the professionalism of the service administering the sentence.

**'Prison works':** (reported for example in Nellis, 1995) in 1993 the then Home Secretary reversed policies of decarceration carried out over the previous decade, declaring that 'prison works', in the teeth of a professional consensus that research evidence shows no evidence that incarceration rates protects the community from crime.

**Problems with research which is driven by policy rather than by theory**

Criminology and penology are hybrid fields in which a considerable degree of research is driven by policy considerations, and paid for by policy-makers. The
absence of unifying theory in the social sciences means much policy-focused research appears atheoretical. Yet all actions taken in relation to crime must rest on some ideas about crime, and about why individuals act as they do, whether or not this is explicit. For example measures to prevent crime such as closed circuit video arrangements in car parks, intended as a deterrent to motor vehicle crime (see for example Tilley 1993), are predicated on a notion of rational offenders calculating likelihood of detection. Similarly work by probation officers rests on some sort of theory however tacit, or naive (as such theories may have been in probation casework, see for example Fielding, 1984).

Criminology and penology are diverse fields containing distinct and irreconcilable paradigms. 'What works' in offender-rehabilitation is of central concern to policy-makers, and to probation, but within much academic criminology it is a minor concern. Jefferson and Shapland (1994) review criminological research in the UK without mentioning rehabilitation, or offender-interventions. Young (1994) describes 'recent paradigms in criminology' without mentioning how any deal with changing offenders' behaviour. (For a review of the place of rehabilitation in current theories of crime and deviancy see Palmer, 1992). The theoretical context of criminal justice related research is complex, as Morgan et al (1994) observe:

'Criminological discourse is extremely varied in character, some of it highly abstract and theoretical, some of it narrowly policy-orientated or technical. It ranges from work that is avowedly politically radical, committed to challenging seats of state power, to that which takes the authority of the law as the only necessary basis for evaluating what 'the problem' is and whether further control is necessary. The disciplinary orientations of those working in criminology include sociology, political science, law, psychology, psychiatry, geography, econometrics and systems analysis. Yet whatever their intellectual backgrounds they tend increasingly to draw on a variety of perspectives when forming and testing their hypotheses'. (Page 8.)

Similarly crime is complex. Downes (1997) offers a useful summary:

'Crime is a vast, complex and ill-charted array of activities, clumped together on the sole common denominator that they are infractions of the criminal law. There is no one theory, and no neat solution, that begins to encompass them all, (page 1)
In this context, Pitts (1992) describes the problematic relationship between the output of social science research and criminal justice policy-making. He sees difficulties in this relationship arising from the different way knowledge is constructed in the social sciences and in law. Law, and policy making requires a 'common sense' consensual view of the world, whereas social science undermines common sense constructions of reality. This is especially so in the heterogeneous field of criminology. Jefferson and Shapland (1994) describe governmental attack on social sciences, in the 1980's, resulting in the growth of 'safe' research, narrowly-focused on policy, and a decline in critical research. Jupp (1989) notes that radical or critical approaches are by their nature unlikely to have a direct effect on official policies because:

'Theoretical ideas which challenge, and sometimes encourage the replacement of existing institutional structures are unlikely to find favour with those who control and manage such structures.'

However contrary to this view, critical criminology has had a significant impact on probation practice, via its presence in probation officer training. More usually for social research output to contribute to policy-making it must have credibility with, and be enlisted by, decision makers (see for example Jupp, 1989, McIvor 1991). Even when this happens the likelihood is findings will be overlaid by ideology. What Young (1994) terms 'new administrative criminology' or 'neo-positivism' has become the major paradigm in governmental approaches to crime, in Britain emanating from the Home Office. This approach is concerned less with the causes of crime than with cures. It conceives crime as fairly rational behaviour, pathology as over-played, and is little concerned with wider problems of social order.

The danger is that when research is driven entirely by policy considerations, it becomes evaluative rather than exploratory, and is less likely to add to broader understanding of the world. Glaser (1975) notes that a problem of evaluation research in general is that it tends to answer narrow questions beginning with "what" rather than with "how". Evaluation research tends to be method driven and
theoretically unsophisticated (Pawson and Tilley, 1994). Jupp (1989) provides a critique of programme evaluation on the grounds that research done to an organisation's brief by its own researchers is owned by the organisation, which then is able to ignore its own findings. Pawson and Tilley, 1994, comment:

'in government and, perhaps more surprisingly, some academic circles, there is a working expectation that evaluators are technical labourers unearthing meticulous data to be fed to the expert minds of policy makers for interpretation and pronouncement. These sentiments have for too long acted as a shelter for theoretically bland experimentalism.' page 306).

The risk is that theoretically sophisticated research may fail to influence policy, and that policy focused research may ultimately prove uninformative.

**What works: the findings of research syntheses**

Many reviews of research have now been published which claim to set out which interventions are the most effective with which types of offender. These include qualitative literature reviews, which may apply varying degrees of structure (for example Sheldon, 1994, McIvor, 1991, Roberts, 1990). But the most influential research syntheses have used the quantitative technique of meta-analysis, most widely cited perhaps being Andrews (1990) and Lipsey (1992). There also exist syntheses of, and commentaries on these and other research syntheses (Gerandreau et al 1994, Losel, 1995, McGuire, 1995, Mair, 1994, McIvor, 1997.) These reviews and syntheses assert, with varying degrees of confidence, what seem to be the characteristics of successful programmes. However the empirical evidence on which these assertions are founded remains of questionable quality (see for example Mair, 1994, Antonowicz and Ross 1994, Mair and Copas, 1997). The sources and reliability of this evidence are examined in some detail in the next chapter, with a particular focus on the doubtful quality of much primary research and the way meta-analysts construct their findings. Furthermore the level of consensus as to the claimed characteristics of effective programmes is not as high as often claimed (see
for example Losel, 1995).

This section surveys and brings together conclusions from an extensive literature, drawing on primary research studies, some of which (for example Robinson, 1995, Lloyd et al, 1995) were relatively large scale studies. This literature also includes qualitative secondary reviews (e.g. McIvor, 1990), or quantitative meta-analysis (Lipsey, 1992.) There also exist reviews of reviews (Losel, 1995, McGuire, 1995, Palmer, 1992). Each of these syntheses, or reviews of reviews rests on different subsets of the original research studies. This present review brings together conclusions and messages from all of this material.

Difficulties of drawing prescriptive advice from research syntheses

Whilst all recent syntheses show that intervention with offenders can reduce offending, some reviewers find that present research evidence supports only the most general of conclusions, from which little prescriptive advice can yet be drawn (see for example Losel, 1995). However other reviewers claim that existing research evidence can support detailed prescriptive 'principles of what works' (see e.g. Andrews, 1992, Gendreau, 1996, McGuire 1995, Home Office, 1995b.)

Evaluative research can describe what happened in a given offender-intervention. But most interesting, from the perspective of finding effective interventions, is the extent to which findings can be generalised, and underlying principles discovered which when applied to future interventions, mean they will likewise be effective. Statistically significant findings, by definition, are generalisable to populations (see next chapter.) But there are dangers in generalising from specific programmes, carried out in a particular context (national or organisational), or from primary research which is of uncertain quality, or where little is known about programme context, content or treatment integrity.

To illustrate the risk of over-generalising, the largest reduction in offending
reported by Lipsey (1990) was achieved by employment programmes run in North American juvenile justice systems. On average these programmes achieved significantly improved outcomes, equivalent to a reconviction rate of 32 percent, where 50 percent of an equivalent control group were reconvicted. However it is not clear how this widely reported finding (see for example McGuire, 1994), should be interpreted. The finding is based on just four programmes, and so as Lipsey (1995) notes, could be a 'statistical fluke'. Furthermore these four studies were carried out in North America, with juveniles, and could have been carried out at any time from the 1960s to the mid-1980s. This finding is of doubtful relevance to adult offenders in Britain in the late 1990s, yet based on this finding, Underdown (1998) cites 'employment related' interventions as being amongst the most effective.

There is divergence amongst reviewers as to what characterises effective interventions, however there is a high degree of consensus as to what forms of interventions are ineffective in terms of reducing reoffending. (Though the same warnings with regard to the dangers of over-generalisation should be registered as with positive findings.)

**Ineffective programmes**

The following types of intervention and sentencing have been found to be ineffective:

- judicial sanctions, deterrence, punishment and imposed sentences

Andrews *et al* (1990) report that

'to our knowledge, not a single review of the effects of judicial sanctioning on criminal recidivism has reached positive conclusions except when the extremes of incapacitation are tested, or when additional reference is made to moderators..' (page 373)
Andrews et al (1990) also report on criminal sanctions which seek only to offer an alternative to custody, and which do not attempt to help offenders change, or address their offending. This category of sanctions includes for example sentences such as community service, or probation with an enhanced surveillance or reporting element. Such sentences have at best only a slight and inconsistent impact on recidivism. McGuire (1995) and McIvor (1997) report on outcome studies of such Intensive Probation Supervision studies in the United States. These were characterised by increased reporting on the part of offenders, home confinement, surveillance, curfews, random drug testing, and electronic tagging. As Petersilia and Turner (1993) report, such programmes have no impact on recidivism. In fact some are associated with increased offending, through breaches of supervision conditions, and through the fact that more crimes are detected by probation officers. Gendreau et al (1994) reach similar conclusions.

McIvor (1997) states that interventions have proved ineffective:

'which are imposed upon offenders, rather than negotiated as part of a contractual agreement and implemented with the offender's contractual agreement and implemented with the offender's consent.' (page 21)

Lipsey (1991, 1995) reports that certain types of intervention with juvenile offenders and delinquents can actually increase the likelihood of negative outcomes. The juvenile justice treatment type which has the most negative effect with juvenile offenders is that of deterrence or 'scared straight' approaches such as shock incarceration. The best known UK study of this type is that of Thornton et al (1984) which reported on the 'short sharp shock' for young offenders in England and Wales detention centres. This found no impact on reconviction.

Bonta and Crowley (1997) describe Gendreau and Goggin's (1996) meta-analysis of criminal justice sanctions, which included fines, shock incarceration, 'scared straight' programmes, intensive probation (i.e. intensive in terms of reporting
conditions and restrictions); drug testing, electronic monitoring and restitution. None achieved significant reductions in recidivism and in fact the first four were associated with increased reoffending.

Gendreau et al (1994) relate findings as to the ineffectiveness of punishment and deterrence in reducing reoffending, to well-known findings from behavioural psychology, that show that punishment is effective in suppressing behaviour only in very limited conditions, where it is clearly and closely linked to that behaviour. To be effective punishment must be immediate, at high intensity, it must be varied so experimental subjects do not get used to it, it must be always contingent upon the target behaviour, and escape from the punishment must be impossible. In the real world of offending, where it is estimated that 95 percent of crime never leads to a criminal sentence of any kind (Barclay, 1993) it is clear these conditions can not be fulfilled, so it should not be surprising that punishments have no impact on reoffending.

Many of the offender treatments included in research reviews represented as concluding 'nothing works' (for example Martinson, 1974, Brody, 1976) were sentences. There are problems in treating sentences as interventions. The content of a given sentence may be obscure, and may include what can be defined as intervention. The findings summarised here imply that sentencing alone is ineffective. An important finding which runs counter to the above is that of Lloyd et al (1994), who found that sentences likely to have the largest degree of intervention input, conditional probation orders, were least effective. Offenders sentenced to conditional probation orders reoffended appreciably more than could be expected on the basis of their age, gender and previous convictions. However reducing reoffending is only one aim of sentencing. Some would argue that punishment has intrinsic merits, relating to justice, and denunciation, regardless of whether or not it deters from further crime.

- programmes that target offender-need factors not predictive of criminal
behaviour;

Gendreau et al (1994) report that interventions which target non-criminogenic factors, such as anxiety, depression, or low self-esteem tend not to reduce offending. Perhaps this should not be surprising as these conditions do not predispose people towards offending. These authors report that approaches derived from subcultural or labelling theories of criminal behaviour based on views of respecting offenders' culture are ineffective; as are approaches which seek to provide legitimate opportunities only (rather than addressing offending), or that rely on incidental learning.

- unfocussed, unstructured, and inconsistently applied interventions;

McIvor (1997) states that interventions which have proved ineffective in reducing the frequency or seriousness of re-offending include those which are unstructured and unfocussed, and which are applied inconsistently. Sheldon (1994) reports similar findings.

- psychotherapeutic models and individual casework counselling;

MacDonald et al (1992) cite a number of significant studies which showed that unstructured psychodynamic counselling in social work generally has proven to be ineffective, despite what they describe as the profession's 'long engagement with psychodynamic theory' (page 616). In the same paper MacDonald et al go on to review a number of studies which indicate that where casework is more effective when it has a greater degree of structure, such as in task-centred approaches. Sheldon (1994) reports similar findings from studies with offenders. Gendreau et al (1994) note that traditional Freudian psychodynamic and Rogerian nondirective therapies tend to be ineffective.
- medical treatment;

McGuire (1995) and Gendreau et al (1994) report that treatments based on administering medication, or dietary change have been found not to be effective. However a significant exception to this finding is substitute prescribing for opiate-using offenders. This is a form of medical treatment, which has been found to have a significant impact on crime, see Killias and Rabasa's (1997) significant paper on heroin prescribing in Switzerland. Hough (1996) provides an extensive review of research in this area, and concludes, significantly, that coerced treatment, delivered via the criminal justice system, appears to be no less effective than voluntary treatment.

- employment related programmes run outside the youth justice system.

Lipsey (1990) found that employment-related programmes for juveniles run outside the juvenile justice system seemed to produce a slight increase in recidivism, whereas employment programmes run within the juvenile justice system produce the most marked reductions in recidivism over controls. However Lipsey (1995) cautions that there are too few studies in this component of his meta-analysis for definitive conclusions, and that these paradoxical findings may be the result of a statistical fluke.

Paradoxically, it is with negative findings such as those summarised above that research is most easily able to support policy-relevant statements, such as 'punishment does not work'. That these findings offer no threat to the continued existence of punishment-based sentences for offenders illustrates the fact that sentences have other objectives than rehabilitation. It also illustrates how research input is only one of a number of conflicting voices in the ear of the policy-maker. But it does show that an individual practitioner who seeks to add to the punishment inherent in a sentence, is very unlikely to help reduce or deter against further offending.
Effective programmes

This section provides a brief overview of the positive recommendations of some of the major research reviews. These are set out uncritically here, though the next chapter calls into question some of the evidence on which they are based. These are reviewed more comprehensively in the final chapter in the light of new findings. A number of authors have set out the characteristics of effective programmes under various sets of headings (see for example McGuire, 1995, Losel, 1996, Andrews et al, 1990, Andrews, 1996, Home Office, 1995b, Vennard et al, 1997.) Effective programmes are discussed here in terms of their:

- theoretical approach;

A number of authors report that more effective interventions are based on a sound theory or conceptual model (for example Andrews, 1990a, 1990b, Losel, 1995, 1996.) However it is often difficult in the literature to separate intervention theories from the methods of intervention employed: programmes based on a cognitive-behavioural theoretical model are likely to use cognitive-behavioural methods of intervention, and vice versa. By 'sound' Andrews et al (1990) and Antonowicz and Ross (1994) seem to mean corresponding to psychological rather than what they describe as sociological theories of crime.

Andrews et al (1990) classify as appropriately theoretically based, programmes with designs that correspond to these authors' principles of risk, criminogenic need, and responsivity. However as Losel (1995), and Lab and Whitehead (1990) note, because classification is necessarily performed post-hoc and is not blinded, there is a danger of circularity of definition. Programmes which achieved positive findings may be more likely to be classified as being based on an appropriate theory.
- selection of offenders;

To many reviewers matching high-risk offenders to high intensity interventions is a crucial element of effective programming. Andrews et al (1990, and see also Andrews, 1995, 1996) offer a simple and influential model of what they term 'appropriate treatment' corresponding to three principles of risk, need and responsivity. The risk principle argues that treatment tends to have greatest impact on higher risk offenders because offenders at lower risk have less to gain. This principle is supported by Lipsey's (1990) findings, based on analysis of over 400 programmes for juvenile offenders. However Antonowicz and Ross (1994) in an analysis of 44 'rigorously controlled studies' found there was no significant difference in the way high and low risk offenders responded to intervention. Further to this Robinson (1995) in a high-quality study of over 2,100 offenders, found that in fact low risk offenders responded more positively than high risk offenders to cognitive skills training.

Lab and Whitehead (1990) note that Andrews et al (1990) fail clearly to define what constitutes high and low risk, and categorise the component studies of their review erratically according to these criteria. Lab and Whitehead also question whether prediction techniques are adequate to make this distinction. This looseness of what is meant by high and low risk enables Robinson (1995), on finding lower risk offenders respond more positively to intervention, to conclude this is because really they are high risk compared to offenders in other studies.

- method of intervention;

It is with method of intervention that consensus between reviewers of the 'what works' literature is at its highest, in the common finding that cognitive-behaviourally based methods are most effective. Cognitive-behavioural methods draw on cognitive and behavioural psychology. However the term 'cognitive-behavioural' can cover a variety of quite different programmes. Vennard et al
(1997) describe how this term covers interventions drawn from the following areas of psychological theory:

- **behaviourism**, which stresses the importance of external rewards or punishments (positive or negative reinforcers) in shaping how people act;

- **cognitive theory**, which emphasises the importance of internal thought processes, which mediate the role of environmental reinforcement. Describing offenders as 'impulsive' (see for example Ross et al 1988), is to point to the importance of thought processes.

- **social learning theory**, which allows that individuals can learn by observing others' behaviour.

Thornton and Hogue (1993) describe a typical cognitive-behavioural programme for imprisoned sex-offenders. This involved 35 to 40 structured group-exercises designed to increase motivation not to offend, and build relapse prevention strategies. Exercises included examining group-members' offences, and hearing their justifications which generally involved distorted thinking. These cognitive distortions could then be challenged by the group tutor, and other offenders in the group. The relapse prevention element involved offenders identifying the sequence of events which preceded their offences, in order this could be avoided in future.

McGuire (1995) offers examples of six types of cognitive-behavioural intervention: *behaviourally-based*, which use reinforcement strategies to reward good behaviour and reliably to punish bad behaviour, as may be the case in token-economies; *relaxation and systematic desensitisation*, which help people to relax in situations which previously aroused tension and aggression; *social skills training*, which teach strategies, for example, to help people to resist peer-pressure to join in situations where offending is likely; *self-instructional training*: whereby offenders learn to monitor their thoughts and practise self-control; *training in moral reasoning*: which seeks to reduce anti-social attitudes. McGuire (1995) also cites *multi-modal*
programmes, which use a variety of techniques to change thinking and behaviour, as offering the strongest prospects of reducing reoffending. Chapter 6 of this thesis describes a multi-modal programme, the 'Reasoning and Rehabilitation' programme (see Ross et al, 1988.)

A number of reviewers find that effective programmes focus on criminogenic need (see for example Andrews et al, 1990, Gendreau et al, 1994, Antonowicz and Ross, 1994.) Criminogenic factors are those which are directly associated with offending. In the case cited above, of Thornton and Hogue's (1993) programme for sex offenders, an example of criminogenic need would be distorted thinking patterns, for example that victims might want to be abused, which are often used to justify this type of offending. A programme which aimed to correct these cognitive distortions might be more likely to reduce their offending than a programme which acted on some less relevant factor, such as self-esteem. With the 'Reasoning and Rehabilitation' programme (Ross et al, 1988), which is aimed at persistent offenders, the criminogenic needs targeted include impulsiveness and lack of self-control, which it is hypothesised lead offenders to continue getting into trouble.

Most reviews emphasise the importance of matching programme content to offenders' learning styles. In most cases this means using methods which actively engage and involve offenders (see for example Andrews et al, 1990, Gendreau et al, 1994, Antonowicz and Ross, 1994.) When a programme involves teaching skills, as in many social-skills training exercises, this means that programme members should actively rehearse and practice the particular skills.

McIvor (1990) reports that approaches which aim to improve non-criminal opportunities in the community show some promise of success. She also reports that effective programmes seem to be those which are responsive to offenders' circumstantial and personal needs.

- amount or intensity of intervention;
In addition to the content of intervention, its intensity also is critical according to most research syntheses. Sheldon (1994) reports that intensive, short-term, time-limited interventions are most effective. In terms of how 'intensiveness' is defined, for Gendreau et al (1994), intensive services, suitable for high risk offenders, are those which occupy 40 to 70 percent of an offender's time, and are of 3 to 9 months in duration. Lipsey (1990) found that the most effective services involved at least 2 contacts per week, for at least 26 weeks.

- implementation, programme integrity and enforcement;

Programme integrity is an important concept, referring to whether an intervention is implemented as intended. McGuire (1995) states that this means ensuring that methods used in a programme support the programme aims. This is an area of programme-design on which most reviewers are agreed. Sheldon (1994) emphasises the importance of programmes being planned, and having clear goals. Good implementation also means ensuring that the programme is adequately resourced and managed, that staff are appropriately trained in delivering the programme, and that the programme content is well-documented in manuals and training materials.

The need for recording and monitoring programme content is partly to guard against what Hollin (1995) terms programme drift, programme reversal and non-compliance. Hollin (1995) cites an example of programme reversal occurring in a programme designed to teach self-control and anger-management techniques. Some programme staff disagreed with the theory on which the programme was based, and felt that what the programme participants needed to do was not learn to techniques of self-control, but to express their pent-up anger. Programme drift and non-compliance tend to occur over time, as programmes are delivered by different staff, and aims shift, or elements of a programme change or are removed.

Lipsey (1990) reports that where a researcher, responsible for evaluating a programme is also involved in the design or delivery of the programme, the
programme is more likely to achieve successful outcomes. Lipsey (1995) attributes this to improved programme integrity resulting from having someone in situ to ensure the intended programme is delivered to the intended client group. However Lipsey (1995) also acknowledges the possibility of bias, which might influence reported results and publication (this matter is explored in the next chapter.)

Staff characteristics, in terms of training, professional orientation, motivation and competence are likely to affect programme delivery very significantly (see for example Andrews et al, 1990a.) However this is not an area where there is yet much evidence. Antonowicz and Ross (1994) included information about staff experience, motivation and training, supervision, and team-working in their meta-analysis of well-controlled programmes, but found no association with effectiveness. Losel (1993) reported his earlier analysis of organisational characteristics, including staffing variables as moderators of programme outcome, but was not able to reach conclusions. Losel (1993, 1995) concludes there is no evidence as yet that staff motivation and competence are clearly related to outcome. However there is work by Stanley and Goldberg (1984) which found that staff characteristics are significant; and by Palmer (1973) which indicates that how clients are matched to staff does affect outcome. Gendreau et al (1994) state that it is important well trained and supervised therapists should relate to offenders in interpersonally sensitive and constructive ways. These authors define trained as having a degree or equivalent, knowledge of theories of criminal behaviour, and between 3 and 6 months of formal training in the application of behavioural interventions.

The way in which a programme is enforced is also a significant contributor to programme integrity. If most offenders on a community-programme attend only half its sessions, it can hardly be said to have been appropriately implemented. Lipsey (1990) emphasises the importance of delivering the intended treatment to each intended client. This means imposing restrictions and controls. McIvor (1990) cites evidence this should only be done when justified by serious risk of reoffending, and where services commensurate with need are also applied. Bonta et al (1994) also emphasise the need for enforcement of programme contingencies and behavioural
strategies.

- programme context.

One of the main contextual issues discussed in the literature is that of whether programmes are most effective when run in the community, rather than in prisons. Most meta-analyses and narrative reviews report that community-based disposals are more effective in reducing reoffending (see e.g. Andrews et al. 1990b, Lipsey, 1990, McIvor, 1990.) However Antonowicz and Ross (1994) in a meta-analysis of well-controlled studies found no difference between community and custodial settings. Losel (1995) attributes less positive findings for prison-based intervention to the negative environment of custody, and to the difficulty of transferring lessons learned in prison into the community. However this does not mean that programmes run in prison cannot be effective, see for example Marshall (1997), Robinson (1995), Thornton and Hogue, (1993.)

An important contextual factor, identified by Andrews et al, (1990), is that programme structure and activities support effectiveness when they disrupt criminal networks. It is easier to see how this can be done in the community than in custody. For example (Chapman, 1995) describes a project for young offenders who take cars, which organises go-carting activities between midnight and 3.00 a.m. on Friday nights. This is precisely the time when the joyriders targeted by the project might otherwise be with their peers, and at risk of offending.

McIvor (1990) warns that programmes should not indirectly increase the risk of offenders going into custody, as can happen via 'net-widening', whereby sentencers may feel they are helping offenders, by sentencing them to disposals which promise to address their needs, but which are at a higher 'tariff' than warranted by the seriousness of the original offence. The danger is that any of these offenders who breached the conditions of a disposal intended for offenders genuinely diverted from custody, could be penalised by themselves being sentenced to custody for
what initially may have been a minor offence (see also Raynor, 1988, and Roberts, 1989 for a discussion of this problem in relation to probation programmes.)

Andrews et al, (1990a) emphasise the importance of professional over-ride, in that the current state of knowledge means that principles of effective programmes are only guidelines, and not all decisions about offenders can be evidence-based.

Table 1.1 summarises the characteristics of effective programmes, identified by meta-analysis and structured review of the effectiveness literature.

**How much agreement is there between research syntheses?**

There is a near consensus regarding what does not work, but far less as to what does. Some reviewers (most notably in terms of influence over the probation service McGuire, 1995) formulate what they claim to be evidence-based principles of effective intervention. However Losel (1995), who gives what is probably the most sophisticated summary to date of 13 meta-analyses, is able to reach only general conclusions. He confirms that most reviews show a positive general effect of treatment, and that cognitive-based programmes, run in the community, seem to be most effective, but goes little further than this.

The apparent high level of agreement in Table 1.1 between authors arises partly because some of the reviews cited are narrative reviews, which include findings from meta-analyses. There is a lower level of agreement across the meta-analyses themselves. Whilst Andrews et al (1990) and Lipsey (1991) find that most programmes work to some extent, Antonowicz and Ross (1994) find to the contrary that:

'only 20 (45%) of the 44 controlled studies were effective. Twenty effective programmes in 21 years indicates that effective programmes are truly exceptional'... (page 98).

Antonowicz and Ross (1994) go on to report, of their meta-analysis, that:
the major finding of this study is that there is very little empirical support for the validity of many of the suggestions or assertions that appear in the literature about the essential characteristics of effective programmes. In fact given the distressingly poor quality of research and reporting of
Table 1.1  What Works? The characteristics of effective intervention, revealed by meta-analysis and structured review

selection of offenders
- match high risk offenders to high intensity services, (Andrews et al, 1990; McIvor, 1990; McGuire, 1994), Gendreau et al, 1994.)

intervention: theories and methods
- training / skills orientated (Lipsey, 1992, Fabiano and Ross, 1987)
- multimodal, multifaceted, structured (Lipsey, 1990, Antonowicz and Ross, 1994)
- improve non-criminal opportunities and facilitate community reintegration (McIvor, 1990).
- are responsive to offenders' circumstantial and personal needs (McIvor, 1990);

dose / amount of intervention
- at least 2 contacts per week for over 26 weeks (Lipsey, 1990)
- intensive, short-term, time-limited (Sheldon, 1994).
- intensive, meaning services occupy 40-70 percent of the offender's time and are of 3 to 9 months duration (Gendreau et al, 1994).

implementation / integrity / enforcement
- monitor, supervise and implement the treatment well. Have a treatment plan and deliver the intended treatment to each intended client (Lipsey, 1990).
- have a researcher involved in the project (Lipsey, 1990)
- well planned, with clear goals (Sheldon, 1994)
- apply restrictions and control only when justified by serious risk of reoffending, and where services commensurate with need are also applied; (McIvor, 1990)
- enforcement of programme contingencies / behavioural strategies to be firm but fair (Bonta et al 1994).
- trained therapists relate to offenders in interpersonally sensitive and constructive ways. (Andrews et al 1990)
- programme structure and activities disrupts criminal networks (Andrews, 1990.)
- staff motivation and competence are not clearly related to outcome (Losel, 1995)

programme context
- allow professional override (Andrews et al, 1990)
- do not indirectly increase the risk of offenders going into custody (McIvor, 1990)
- there is a high level of advocacy and brokerage as long as the community agency offers 'appropriate' services. (Gendreau et al, 1994)

Research-syntheses are also less well founded than sometimes claimed. The most influential research reviews (probably Lipsey, 1990, and Andrews, 1990) rest on the controversial technique of meta-analysis (see Mair, 1994, Mair and Copas, 1998 and the next chapter of this thesis.) The most recent research findings, from large-scale well-controlled studies, (Robinson, 1995, Bottoms, 1995, Raynor and Vanstone, 1996) of programmes which embody many of what McGuire (1995) terms the 'principles of what works' yield only modestly positive findings. It may be premature to claim that the answer to effective rehabilitation has been found, as Glaser (1975), states, 'The history of corrections is paved with punctured panaceas.' However there is a clear consensus amongst reviewers as to which intervention methods do not work. These include punishment, deterrence, and intensive approaches which do not seek to change offenders. Likewise unfocussed and psychodynamic counselling and casework do not work. Disposals which are merely alternatives to custody, with no intervention input, do not work. Disposals which do not focus on and challenge offending, or the factors which lead to offending and anti-social behaviour do not work. Treatment can work, and is probably best when it targets behaviours which lead to offending, and targets the thinking, social perceptions and skills which underlie these behaviours. It is probably most effective when delivered in a community setting, in a structured, explicit, time limited way.

The current research

Reducing offending and rehabilitating offenders continue to be overarching aims of the probation service. There is increasing pressure to demonstrate effectiveness in achieving these aims, and increasing calls for work with offenders to be based on research evidence. However much remains to be discovered as to which principles of intervention are likely to be most effective, in particular circumstances and with particular groups of offenders. Even if knowledge about 'what works' was better
developed, it would still be necessary to monitor individual offender-interventions, to ensure they were succeeding in reducing reoffending. These two questions: 'what works', and 'how can we tell if it's working', are what the research case-studies reported in later chapters seek to address. This research was also carried out in the context of the emerging evidence-base described in the section above.

The first stage of the research (reported in detail in Wilkinson, 1994a, 1994b, 1994c, and described briefly in the next chapter) involved developing suitable outcome measures, using information about reconviction. Having settled on outcome measures and evaluation methods, a set of programme-evaluations was carried out. These were each of mainstream, well-known and influential probation projects, all of which worked with high-risk offenders in the community with the aim of reducing their offending.

The next chapter, Chapter 2, looks at how we can know whether methods claimed to reduce offending actually do so. The first part of this chapter revisits the 'what works' debate, to look more closely at the research methodologies which underpin the research evidence, and in particular at the technique of meta-analysis. The conclusion is that pronouncements based on meta-analytic review have been somewhat overstated.

The second part of Chapter 2 considers how it is possible to evaluate whether offending is being reduced by a given probation programme. This means identifying salient outcome measures, most commonly reoffending as measured by reconviction. This section also considers the use of intermediate outcome measures, which can measure the outcomes of programmes not related to reoffending, and which may provide shorter-term feedback as to whether programmes are likely to be successful in reducing offending.

The remaining chapters apply these evaluation methods to five probation programmes, all of which had the aim of holding offenders in the community without their reoffending, and reducing their offending in the longer term. The
evaluations for the most part cover overlapping time-periods, and are arranged in rough chronological order, and in order of the extent to which programme content was influenced by the progress of debate as to what works, and what constitute effective methods.

Chapter 3 reviews the long term impact on reconvictions of the first Inner London Probation Service Demonstration Unit. This ran between 1981 and 1985, and worked with offenders on two separate offence-specific programmes, one for offenders who had committed burglary; and one for offenders who had taken cars. This was one of the first probation programmes to carry out offence-specific work with offenders, aimed at reducing their offending.

Chapter 4 reviews another offence-specific programme for offenders who take cars: Ilderton Motor Project. This programme still functions in the way described, working with young people who take cars. Chapter 4 reviews the impact on offending of the programme for offenders who passed through the programme between 1990 and 1993.

Chapter 5 evaluates the impact on reconvictions of a probation centre, Sherborne House, which runs a day programme for young male offenders aged between 17 and 20. This works with people who have committed all types of offence (mainly theft, burglary, robbery and violence against the person), but at such a level of seriousness, and with such a degree of previous offending, that offenders would have been sentenced to custody, were it not for the existence of the centre. At the period being reviewed, 1991 and 1992, the Sherborne House programme was starting to take on some of the early messages around what works.

Chapter 6 evaluates what can be considered a fully-blown evidence-based programme for offenders: the Reasoning and Rehabilitation programme as run at Camberwell Probation Centre. The programme is predicated on the notion that many repeat offenders offend because of deficits in their 'social intelligence'. The programme uses cognitive-behavioural and educational methods to rectify these
deficits, and thereby reduce offending. The Reasoning and Rehabilitation Programme has been run and evaluated, with mixed results on a number of sites in the UK and in North America, in particular in the Canadian Prison Service.

A particular point of interest of Chapter 6 is that this evaluation uses a number of intermediate outcome measures, to assess changes in the way offenders think, the problems they face, and their attitudes to crime. Changes in these measures are correlated with changes in reconviction.

The concluding chapter, Chapter 7, draws together the findings of these case-studies, and reviews these findings against the claimed characteristics of effective programmes, as revealed by structured research literature review. This chapter concludes by pulling together what messages the new research might have for telling us what methods are likely to work with 'main-stream' offenders.
Chapter 2

How can we tell if interventions reduce offending?

This chapter takes up the questions 'how do we know "what works?"' and 'how do we know if it is working in practice?'. When reviewing the evidence-base for work with offenders, it is essential to consider in some depth the research methodologies used to produce this evidence. This chapter first looks at the methods of evaluation research, and of research synthesis, from which the 'what works' messages described in Chapter 1 were derived. This chapter also considers the methods whereby it is possible to evaluate whether an intervention is effective in practice. Starting with a broad perspective on methods of research synthesis within the 'what works' literature, this chapter focuses in on methods of evaluating programmes, paying particular attention to the status of reconviction as an indicator of the effectiveness of offender-interventions. The final section of this chapter focuses in on how the intervention case-studies which make up the research-based chapters were chosen, and how they fit in the continuum of probation practice.

Methods of research synthesis

Research synthesis is fundamental to the evidence-based approach. The previous chapter summarised what various reviewers of the 'what works' literature have described as successful ways of working with offenders, based on their reviews of individual research projects. However different authors have derived quite different messages from reviews of largely the same, or overlapping groups, of research studies. One reason for this is that methods of research review are not uncontroversial. In particular the technique of meta-analysis, developed in medical research, has made a significant contribution to debate about what methods are effective in achieving offender-rehabilitation. Even the more critical commentators
(most notably Mair and Copas, 1997) regard this technique as having put paid once and for all to the notion "nothing works" in offender rehabilitation. However the validity of this technique, as it has been applied in the offender field, may be in doubt.

Many reviews and syntheses of research have been carried out to determine the characteristics of effective programmes. These have included qualitative literature reviews, which apply varying degrees of structure (for example Sheldon, 1994, McIvor, 1991, Roberts, 1990). The most influential research syntheses have used quantitative meta-analysis techniques, most widely cited perhaps being Andrews (1990) and Lipsey (1991). There also exist syntheses of and commentaries on these and other research syntheses (Gerandreau et al 1994, Loesel, 1995, McGuire, 1995, Mair, 1994), which with varying degrees of confidence assert the characteristics of successful programmes. However the empirical evidence on which such assertions are founded remains less than conclusive (see for example Mair, 1994, Antonowicz and Ross 1994, Mair and Copas, 1997). Furthermore, the level of consensus as to what constitutes the characteristics of effective programmes is less than sometimes claimed (see for example Loesel, 1995).

Qualitative research synthesis

Evaluative research projects which aim to identify effective methods of working with offenders are usually experimental or quasi-experimental in design. Such studies yield either clearly positive results (e.g. Ross et al, 1988, Bush, 1995), broadly positive results, from which a number of things might be hypothesised, (e.g. Goldberg and Stanley, 1984), mixed results (e.g. Folkard, 1974, 1976), or they may show wholly negative findings. In drawing conclusions about what is effective, the reviewer of the research must select studies and components of studies, attempting to follow themes and build a larger picture. Necessarily there is a significant subjective element to this process of narrative review. It is possible to build very different pictures from the same components. An example of this is
Palmer's (1975) revisiting and drawing very different conclusions from 138 studies included in Martinson's (1974) review.

**Quantitative research synthesis: meta-analysis**

Meta-analysis is a form of quantitative research synthesis, whereby the results of many diverse projects can be combined, in order that trends may be detected which might not be revealed by a single study. The technique of meta-analysis has been said to have settled the debate as to whether or not 'nothing works' (see for example Lipsey and Wilson, 1993, McGuire, 1994, Mair, 1994). However as Mair (1994) has observed in this context, as a technique meta-analysis is far from uncontroversial. Given its continuing contribution to the 'what works' debate (see for example Underdown and Ellis, 1998, McGuire 1997) it is worth scrutinising the technique in some detail.

Meta-analysis was developed in medical research, in order the results of closely related experiments into very similar types of treatment could be combined. It was developed for situations where treatment effects are small, and may be masked by chance variation between studies. If two studies of a medical treatment are each too small individually to reveal a statistically significant effect, using the summary statistics from each study their combined findings can be tested for significance. Thus a number of small and inconclusive studies can be joined to become one large experiment, from which more definitive pronouncements can be made. However the need in medical meta-analyses is to demonstrate that the patients in all studies suffer the same disorder, and are similar in terms of relevant bio-data. There is also a need to demonstrate that treatments are identical (see Thompson and Pocock, 1991, for a discussion of these issues.)

Meta-analysis has been applied more recently in the behavioural sciences (for a review see for example Lipsey and Wilson, 1993), and latterly in fields of offender treatment. Here the technique is applied more loosely. Whilst there is no one agreed
procedure, all meta-analyses involve classifying studies according to a number of dimensions including type of subjects of the research and type of treatment, and type of outcome measure used. The summary statistics from each study are recalibrated to a standard scale, so they may themselves become the subject of analysis. The aggregated data are then analysed quantitatively. It is possible to group projects together along any of the dimensions defined and look at the outcome. Thus, say, all employment projects may be grouped together in one analysis to assess their impact on offending.

Meta-analysis is particularly useful in situations where the effects of intervention may be small, so that findings from individual small-scale research projects tend to be mixed, or not to be statistically significant. The intention behind meta-analysis is that an increased sample size produced by combining individual studies lets genuine effects show through the statistical “noise” of random variation. Thus significant findings can be obtained from groups of projects which would not be statistically significant in all projects individually. The technique is particularly useful in fields of medicine, where relatively small effects may be very important. For example, reducing a death rate from 10 in a thousand to 7 people in a thousand could mean many lives would be saved. In the behavioural sciences meta-analysis may be used because small effect sizes are all that exist. However on the latter point Lipsey (1993) comments:

'in contrast to the previous era of conventional research reviews, meta-analysis has yielded stark, dramatic patterns of evidence for the general efficiency of such [psychological, educational and behavioural] treatment.' Lipsey, 1993, page 1182.

Nevertheless meta-analysis remains controversial.

Problems with interpreting meta-analytic findings

Meta-analysis is not uncontroversial within medical research, where the technique
originated. In this field, meta-analysis is used to synthesise research studies which meet fairly stringent conditions, usually where treatments are carried out on tightly defined populations, who suffer from diagnosed well-defined disorders, and who are administered tightly defined treatments, usually under conditions of random allocation. Studies which do not meet these conditions are excluded from analysis. However, in social science research generally, and certainly as applies to offenders, this tightness of definition across studies is rarely possible. This section sets out some of the main problems with the use of meta-analysis in the criminal justice field.

Sampling issues in meta-analysis

Meta-analyses treat research studies as though they were research subjects, classifying and analysing them to discern information about the wider world. The statistical techniques used to do this (commonly t-tests and standard deviations) rest on certain assumptions, for example that research subjects constitute a random sample of a wider population, and that differences between research subjects on certain indices will be normally distributed. For example, Lipsey (1990), computes a standardised effect size for each separate study included in the review. The statistic Lipsey uses is Cohen’s $d$, computed as the difference between the treatment group mean score, and the control group mean score, divided by the pooled standard deviation of those scores, and assigned a positive or negative value, depending on whether the treated group did better or worse than the control or comparison group. Thus effect size for each study is expressed in standard deviation units. Lipsey (1990) calculates the average treatment effect size for a group of studies as the arithmetic mean of the set of standardised effect sizes, expressing this either as a correlation coefficient ($r$), representing the association between intervention and outcome; or alternatively as the notional reduction in delinquency outcomes (however measured) compared to a control group 50 percent of whom show a negative outcome.
This pooling of effect sizes has implications when interpreting the findings of meta-analysis. Mair and Copas (1997) observe that a meta-analysis of a range of offender programmes which showed a mean effect size of zero, rather than supporting the hypothesis "nothing works", merely would show that equal numbers of treatments work as do not. A salient issue here is that of heterogeneity within groups of experimental subjects and between experiments. Thompson and Pocock (1991), reviewing the use of meta-analysis in medical research, make it clear that for the general effects revealed by meta-analysis to be meaningful, it must be clear that the same treatment is being investigated on the same type of experimental subjects, so that the component studies being meta-analysed can genuinely be seen as parts of a larger experiment. Few would argue that much, if any, criminological research can satisfy these conditions.

Mair and Copas' (1997) main charge against meta-analysis also relates to the issue of using the same statistical techniques on a set of research studies as would be used on a random sample of individuals drawn from a population. This makes the assumption that all research studies are drawn at random from a wider population of research studies. But Mair and Copas (1997) produce evidence that selection of individual research studies for inclusion in meta-analysis may be skewed.

The most likely cause of bias in the selection of research studies for inclusion in meta-analysis, from a hypothetical population of such studies, is that of publication effects. Most meta-analyses are of published studies. It is a near truism that studies with positive or novel findings are more likely to be published. This is sometimes known as the 'desk-drawer phenomenon'. If this is the case it should be no surprise if meta-analysis across a published literature shows positive findings. On this basis Buchanan (1993), agreeing with Oakes (1986), concludes of meta-analyses

'...that these illegitimate offspring of statistical theory should have been stifled at birth' (page 528).

Lipsey and Wilson (1993) reviewed 302 meta-analyses carried out over the fields of psychological, educational and behavioural treatment, and found some evidence of
publication bias. The mean positive effect size (i.e. treatment effect being in the desired direction) for published studies was significantly higher than for unpublished studies (r=0.53 for published studies, r=0.39 for unpublished studies, where r is the correlation coefficient indicating the positive change in standardised outcome measures associated with intervention.)

Losel (1993, 1995) argues that because the largest meta-analysis of offender treatment (Lipsey, 1992) contains unpublished studies it therefore possesses no publication bias. Losel (1995) attributes the relatively low effect size (r=0.103) of this meta-analysis to the presence of these unpublished studies. However, just under 10 percent of the studies in Lipsey (1990) are from the 'grey literature'. Whilst no effect of publication on outcome is reported, the unpublished studies would have to have had generally very negative findings to have much influence on the overall effect size.

Redondo (1996, reported in Vennard et al, 1997) similarly included unpublished studies in his review of European programmes, with the finding that publication did not have a bearing on effect size. However, Mair and Copas (1997) in a significant critique of meta analysis, present evidence that the overall reductions in recidivism achieved by offender programmes, shown in Lipsey (1990), can be entirely explained by publication effects. The basis of their argument is that larger studies (in practice those with over several hundred subjects), are more likely to be published regardless of their findings than smaller studies, say with under a hundred subjects, which are likely to be published only if they report interesting findings, in this case reductions in offending. By this token there should be an inverse correlation between the effect size reported in a study, and the size of the study. This is exactly what is seen, the degree of this correlation explaining entirely the general effect sizes revealed by meta-analysis.

Lipsey (1995) explains the bias toward larger positive effect sizes in smaller studies as being the result of a higher level of monitoring and programme integrity than is likely to be the case with larger scale studies. Another possibility is that larger
studies may imply larger programmes, which may contain more heterogeneous
groups of offenders, and less well controlled treatments - in other words that they
are less likely to be effective than smaller programmes - and that the smaller effects
may be entirely due to this.

Mair and Copas (1997) base some of their critique of meta-analysis on evidence
from the medical field, relating to meta-analyses of the use of intravenous
magnesium in the treatment of suspected myocardial infarction (heart attack). A
number of high quality meta-analyses showed that this form of treatment was
clearly effective, however a "mega-study" (ISIS-4 Collaborative Group, 1995) with
a sample of 58,050 patients showed that magnesium treatment had no effect on
mortality. Hlatky (1995) comments: "ISIS-4 thus shows a clear-cut failure of
meta-analysis that should lead to a re-evaluation of this technique".

Mair and Copas revisit the magnesium meta-analyses, and show that it was the
inclusion of smaller studies that skewed the findings toward being more positive.

Another factor which might reflect publication bias is the strong influence of
researcher involvement. Lipsey (1991) found that whether the people who evaluated
a project were involved in designing and implementing the project was second in
importance only to treatment modality, in determining outcome. Lipsey (1995)
concedes that

'a cynical view might attribute this to some biasing or 'wish fulfilling' influence researchers have on
the outcome of studies they control...' (Page 76.)

before offering the alternative formulation, that researcher involvement is likely to
be associated with smaller studies, and higher programme integrity. However
another cynical view might be that people are more likely to seek to make public
their successes than their failures.

Vennard et al (1997) point out that publication bias also applies to more
conventional types of literature review. However this is to emphasise a weakness of all secondary source-based reviews, rather than to defend meta-analysis.

**Over-generalisation from meta-analytic findings**

Another criticism relates to over-generalisation, and over-confident generalisation from tentative meta-analytic findings. This criticism applies to all forms of literature review, but particularly to meta-analysis, because of the 'black box' nature of this technique, where there seems a greater distance between the findings of a review, and the original studies on which the review is based, which may not be cited individually or referenced (as is the case in Lipsey's (1990) widely cited paper.)

Lipsey (1990) is quite clear the age range included in his study represents: 'a formative period marked by behaviour that will not be continued into adulthood' (page 2). However a danger with well known papers is that headline findings are repeated without such caveats. For example one of this paper's most frequently repeated findings (described in the previous chapter, see for example McGuire, 1994) is that the largest single effect size occurred with employment programmes run in the juvenile justice system. This finding is based on four (probably) North American studies of people aged under 21, who probably never had previous jobs. This should not be taken as evidence that placing previously employed adults in the UK in jobs (if this could be done) would have a similar effect.

A third problem is with the quality and relevance of the original studies taken in by meta-analysis. It is questionable whether findings can be improved by combining many poor quality research studies. For example, only a third of the studies in Lipsey (1990) were rated as being of high quality in terms of their methodology. Some meta-analyses reach back to studies carried out in the 60s and 50s. Much may have changed in the last forty years, in terms of the background to offending, and what is likely to be effective in working with offenders.
A fourth criticism of meta-analysis is that whilst it may be represented as entirely objective, in fact meta-analysis remains **subjective** in the way it categorises component research studies. So they can be subjected to meta-analysis research studies must be classified in terms of key variables, which involves judgement on the part of the researcher carrying out the coding for analysis.

A fifth criticism relates to **circularity of definition** when categorising research studies. For example Loesel (1996) criticises one of the major North American meta-analyses (Andrews et al, 1988) on these grounds. If studies are classified post-hoc, after their findings are known, there is a danger they will be grouped not according to input, but according to success. Indeed a number of reviewers have proceeded in this way: grouping together successful programmes, and looking to see what features they share in common (e.g. Gerandrea and Ross, 1979, Fabiano and Ross, 1986, Antonowicz and Ross, 1994.) Proper scientific procedure is to make a prediction, and see if it is borne out by the evidence, or following Popperian logic, falsified. It is valid to proceed descriptively, however if principles are to be derived that describe the world, these must be tested prospectively.

Two further criticisms relate to the way the findings of meta-analysis are used, for deriving advice as to how effective programmes for offenders should be constituted. The first of these is that meta-analysis is **uninformative**. Being based on studies with experimental or quasi-experimental designs, meta-analysis tells us little about the content of programmes, and still less about their processes, which might inform *why* each may or may not have worked (Loesel, 1995, makes this criticism based on his synthesis of 13 meta-analyses). But this may not be a very fair criticism to make of meta-analysis, which does not seek to answer questions which start with 'how', but to identify when and where 'how' questions should be asked.

This criticism becomes more apt when attempts are made to convert the essentially descriptive findings of meta-analysis into prescriptive methods for working with offenders (see for example McGuire, 1994). Meta-analysis can say whether certain methods 'work', but not how they work. This is a weakness of quasi-experimental
methods in general (see Pawson and Tilley, 1994), which is unlikely to be addressed by syntheses which provide information at a still higher level of generality. As Glaser (1975) notes: 'The greatest gains... have come from answering questions that begin with "Why"'. Those responsible for implementing programmes need prescriptive information covering what needs to be done, when and how. This is better provided by process evaluation.

- **heterogeneity of research studies:**

The category 'offender treatment' as interpreted by Lipsey (1990) and Andrews *et al* (1990) includes a wide range of programmes ranging from psychological treatments which aim to change the way offenders think, to purely penal sanctions which have no therapeutic ingredient, but which aim to change offending behaviour via deterrence. Clearly any analysis which groups together very different programmes, for possibly quite different offenders and using different outcome measures, will not produce meaningful results. It makes sense to analyse categories of programme separately. For example sex-offender programmes might form one coherent group; employment based schemes for young property offenders another. The problem is that when grouping studies in this way, the number of primary studies within a given category may become too small for results to be reliable.

**A critical review of Lipsey's 1991 meta-analysis of treatment programmes for juvenile delinquency.**

To illustrate how these problems apply in practice, it is useful to look in detail at the largest, most cited, and purportedly the most reliable (see e.g. Loesel, 1995) meta-analysis of offender treatment. This was carried out by Lipsey (1990), using information about 443 individual research studies. Below are listed a selection of some of the findings reported in Lipsey (1990), Table 3. Descriptive Data for Major Variables Coded. These are grouped according to the three main areas of problem
which might arise from drawing conclusions from generalised findings from these studies:

the weakness of component studies:

- only a third of included studies were rated high quality in terms of their methodology.
- confidence in treatment integrity was rated low in 44 percent of studies, moderate in 36 percent.
- only 14 percent of studies gave a detailed description of their treatment method.
- 10 different measures of delinquency were used, from 'antisocial behaviour' to 'institutionalisation'. These measures are derived from diverse sources, but for purposes of meta-analysis were recoded as a uni-dimensional measure of success or failure.
- 43 percent of component studies had follow-up periods of six months or under.
- the overlap of the delinquency measure with the content of treatment was rated as below moderate in 64 percent of component studies.

potential for inappropriate generalisation:

- all studies included were of juveniles, aged between 12 and 21.
- 92 percent of included studies were carried out in the USA, 3 percent in Canada.
- 14 percent of included studies were published in the 1960s or earlier.
- two thirds of projects were new (under 2 years old), only 35 percent were established projects.

evidence of publication bias

- larger studies, which are more likely to be published, show smaller effect
sizes (see Mair and Copas, 1997).

The findings listed above are intended to illustrate weaknesses of meta-analyses in general, as applied to offender studies, and to caution against generalising too far on the basis of their findings. The points under the first of the above headings call into question the extent to which it is possible to know what treatment is being evaluated and according to what criteria. These criticisms are not intended entirely to undermine the findings of Lipsey's (1990) meta-analysis which remains pre-eminent amongst those so far carried out. It is only because the author has presented his methods and results in such detail that critical scrutiny on this level is possible. The strongest criticisms relate to the wisdom of other authors generalising too far on the basis of what are said to be the findings of this meta-analysis, and to the influence of publication bias when reporting general effects.

If meta-analysis may not reveal 'what works' what can?

An obvious answer to the problem of small effect sizes is to carry out research on a scale where groups are sufficiently large for statistical tests to have the necessary power to distinguish treatment effects. The review of the effectiveness of sentences carried out by Lloyd et al (1995) is significant. This one study used data produced by the Home Office Offenders' Index as to the offending of over 11,000 individuals. It found that probation, community service and custodial sentences had reconviction rates of the level that might be expected on the basis of offenders' sex, age and previous offending. The main exception was with conditional probation orders, which had, overall, a two-year reconviction rate which was four percent higher than predicted. Most offenders on conditional probation orders take part in some sort of programme designed to reduce offending. It is unfortunate Lloyd et al (1995) did not have access to accurate information about programme content, or to more detailed information about offenders.

One large scale evaluation of a well-controlled evidence-based programme for
offenders has recently been carried out. This was Robinson's (1995) evaluation of
the effects on reconviction of the cognitive skills training programme in the
Canadian Correctional Service. This used a sample of 2125 offenders, divided
between a waiting list comparison group, and the treatment group. Some 24.8
percent of the comparison group were reconvicted within a year compared to 21.3
percent of the treatment group. This modest reduction in reconvictions occurred
because of reduced reconviction on the part of offenders defined as being at lower
risk of reconviction, on the basis of their prior offending career. Of lower-risk
programme participants 14.2 percent were reconvicted compared to 20.2 percent of
lower-risk non-participants. This latter finding was statistically significant. For
higher risk offenders no treatment effect was apparent. This finding is interesting as
it contradicts the 'risk principle', propounded by Andrews et al (1990b). More large
scale studies of this type are needed to test the findings of meta-analysis.

The value of meta-analysis

Criticisms relating to publication bias mean caution should be applied when
interpreting the positive findings of meta-analysis across an entire class of studies.
But the corollary of this is that the negative findings of meta-analysis must stand.
Where a negative finding (such as that punishment, deterrence, psychodynamic
approaches, low intensity work with low risk offenders are not effective in reducing
reoffending) emerges, the chances are this represents a picture uninfluenced by
publication or other bias.

The summary provided by Hlatky (1995) of the standing of meta-analysis in
medical research syntheses probably applies equally well to research around work
with offenders:

The cycle of over-optimism, disappointment, and eventual balanced assessment described with new
drugs and devices also appears to apply to methodological innovations such as meta-analysis.
Although meta-analysis is valuable for providing a quantitative summary of evidence, numbers alone
do not tell the whole story... meta-analysis should be an adjunct to, not a replacement for, expert judgement in weighing evidence.

The main conclusion to be drawn from this section is that less of what is purportedly known about effective methods for reducing offending, is reliably grounded in evidence than advocates of what Bottoms (1995) terms the 'new orthodoxy' sometimes claim. Until more large-scale studies are carried out, what is needed is more high quality project evaluations. The chapters which follow describe a set of evaluation case studies of offender-programmes which aspire to this status. These evaluations are all of programmes carried out with the aim of reducing offending. The next section looks briefly at programme evaluation methods, and how these can be applied to evaluating success in reducing offending.

**Programme evaluation methods**

This section looks at the general issues involved in evaluating programmes which aim to reduce offending. Specific issues relating to individual programmes will be covered in the relevant research-based chapters which follow.

**Assessing the effectiveness of probation programmes**

Macdonald *et al* (1992) in a review of social work effectiveness research define four types of effectiveness study design:

- **experimental design**, characterised by random allocation to experimental and control groups;
- **quasi-experimental design**, where allocation to intervention conditions is non-random;
- **pre-experimental design**, using before/after measures to assess change;
- **client opinion surveys**, which assess how clients felt they benefited.
Design issues are discussed below, with relation to making sense of information about reconvictions. However, the studies reported in later chapters all can be characterised as quasi-experimental, using comparison rather than control groups. Although some of these studies draw also on before/after measures, and on client opinions. However Pawson and Tilley (1994) have criticised quasi-experimental designs as being uninformative. They argue that whilst quasi-experimental methods may reveal whether an intervention has worked, these methods are very unlikely to convey why an intervention has been successful, or what processes brought about desired outcomes. But whilst this is a strong argument in favour of carrying out research into the processes within successful interventions, it does not seem to be an argument against performing outcome-evaluations to identify those successful interventions in the first place. Quasi-experimental methods remain the best available for providing a relatively rapid indication of whether a criminal justice project is achieving its aims of reducing offending. If it is, this might justify research into processes, or the use of a “true' experimental design.

The set of evaluations reported in later chapters can all be characterised as being outcomes/effectiveness focused. The importance of a strong focus on outcomes in monitoring and evaluating all kinds of state-funded programmes is becoming increasingly acknowledged (see for example Williams and Webb, 1992, for an account of the outcomes approach.) The set of evaluations reported later have a number of methodological features in common:

- reduced offending as the intended primary outcome;

All the evaluations set out to answer questions as to whether a set of offender-programmes had achieved what they set out to, in this case reduce offending. The main (but not only) measure used of offending is reconviction. The next section reviews the use of reconviction as an outcome measure in some detail. This can be a matter of some sophistication, when factors such as the nature and frequency of reconviction; seriousness of sentencing outcome; follow-up periods; comparison
group outcomes; and predicted reconviction outcomes are taken into account.

- **intermediate outcome measures;**

Intermediate outcome measures can be thought of as steps on the way to achieving reduced offending. In the programmes evaluated in later chapters the main intermediate outcome indicator is completion of the programme in question. Other intermediate outcome measures used include variously getting a driving licence (for offenders who took cars), reduced impulsiveness (for offenders lacking in self control), and increased social stability, in terms of relationships, housing and employment. Mair (1991) advocates the use of intermediate outcome measures alongside reconviction, as a measure of the success of offender-programmes.

- **programme integrity**

It is important to know what a programme actually involved, what methods were actually used and how they were applied. This relates also to how closely the programme articulated the theory on which it was based, and how well the intervention delivered to clients conformed to specification. Without knowledge of the content of a programme, learning about what has worked is impossible, and it is not possible to replicate successful methods on other sites. Hollin, (1995), offers an account of the importance of programme integrity in offender-interventions. Programme integrity also cuts across trends at the time of the research case studies, to offer more transparent services, more consistent in format, via National Standards for probation work (Home Office, Department of Health and Welsh Office, 1995).

- **service user-satisfaction**

The views of service-users are increasingly being seen as critical in the provision of
services in the area of social care (see for example Department of Health, 1995, for a useful and influential summary of this area.) These moves were influencing the probation service at the time of the research case-studies (see for example Foad, 1993), and are reflected in a recognition that the views of offenders as service users are an intrinsically important outcome of offender-programmes.

Programme evaluation techniques

Programme evaluation is becoming an increasingly well developed discipline. Fitz-Gibbon and Morris (1987) offer an account of programme evaluation theory and of evaluation design. Programme evaluation methods draw on a range of techniques used in social and psychological research. These include analyses of the data drawn from:

- official sources (i.e. for conviction and arrest);
- self-report;
- observation;
- secondary source material, including documents and agency databases;
- structured interview and client opinion survey;
- attitude questionnaires and;
- psychometric testing.

In Chapter 6, an evaluation of a cognitive-behavioural programme, all these methods are brought to bear in a single programme-evaluation. This was partly in order to look at the processes associated with offending outcomes.

Statistical methods

The aim of the research reported in the chapters which follow, is to identify
effective methods of working with offenders. The assumption underlying the evidence-based approach is that where effective methods are identified, they can be replicated elsewhere with similar success. Whilst findings about individual programmes may be of intrinsic interest, what is more interesting is what these findings might tell us about methods which are likely to be generally effective. In other words if similar offenders received similar interventions, what are the chances there would be similar outcomes? Statistical techniques exist that allow inferences to be made about populations based on observations obtained from random samples, or samples in which the characteristics and relationships of interest are independent of the probability of being included in the sample.

The project evaluations described in later chapters have a number of features in common. In each evaluation there are two groups of offenders, one who attended the programme, and a comparison group of similar offenders sentenced to some alternative community or custodial disposal. The offending of both groups is then monitored over a period of some years. Finally the offending careers of the two groups are compared, so as to see whether the intervention group offended less than the comparison group.

**Demonstrating programme impact**

To demonstrate that the intervention reduced reoffending on the part of those offenders who attended, the need is to show that before the programme the two groups are drawn from the same population of offenders. That is, that the two groups are similar, especially in terms of factors likely to be criminogenic. To this end the two groups are compared before entry to the intervention, or alternative sentence, on a number of relevant measures. These include demographic variables, but concentrate on factors related to the nature and seriousness of their previous offending. There is also a need to show the two groups are similar in terms of their future risk of reconviction. Following the intervention and over their subsequent offending (or non-offending) careers, the need is to demonstrate the opposite,
namely that in terms of relevant outcome measures (in this case measures of reoffending), the two groups are now different to the extent they might represent different populations. Where this difference can be demonstrated, it constitutes strong evidence that the intervention has done something to change the offenders, making them less likely to reoffend.

The main measure of programme impact used in the research case-studies described in later chapters is that of the difference in reconviction outcome between intervention and comparison groups. This difference is expressed as the percentage fewer of the intervention group who were reconvicted than were of the comparison group. Thus if 40 percent of an intervention group were to be reconvicted within two years, compared to 50 percent of the comparison group, this would be expressed as a 20 percent reduction in reconviction.

In Chapter 6, an evaluation of the Camberwell Probation Centre, another set of comparisons is made. The group that underwent the intervention were assessed, before and after the programme, using a set of attitude scales and psychometric tests. The hypothesis is that the intervention significantly changes these scores, before and after the programme. So, as above, the need is to compare the two groups of means, to see if they are significantly different. But here, because each individual who underwent the intervention has two sets of scores, the test is carried out on the paired set of scores for each individual, on each of the scales and tests.

The main statistical test used to assess whether there are significant differences between samples, in research designs of this type, is known as the t-test. This tests the significance of the difference between means of independent or paired samples. In each case the t-test tests the null hypothesis, namely that the samples are taken from the same population (in other words that no change has happened), and determines how likely this is to be the case. If after an intervention the chances of two formerly similar samples being from the same population are under one in twenty, it is likely that the intervention has brought about some change in reoffending.
Assumptions \( t \) tests make about populations

\( T \) tests make certain assumptions about the populations, the means of which they are comparing. The main assumption is that the distribution of values will approximate to a normal distribution. Table 2.1, drawn from the group of offenders referred to Sherborne House, the programme evaluated in Chapter 5, shows that this group's previous offending does indeed approximate to a normal distribution. Elsewhere in the research-based chapters which follow, checks were carried out to ensure that where \( t \) tests were used, the underlying assumptions were met. For example the psychometric tests used in Chapter 6, and evaluation of the Camberwell Probation Centre programme, reference was made to the published scores of norm groups, which show that again, scores approximate to a normal distribution. But in fact the \( t \) test is robust, in that moderate violations of this assumption of normality of distribution do not significantly affect results (see Noruis, 1988.) The \( t \) test is more powerful in detecting differences possibly caused by intervention than are the non-parametric tests which could alternatively be used. The assumptions of the \( t \) test, of randomness of selection of samples, and normality of distribution appear broadly to be met.

Table 2.1. Number of previous convictions for offenders targeted for Sherborne House

<table>
<thead>
<tr>
<th>Number of previous convictions</th>
<th>Number of offenders (n=240)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
</tr>
<tr>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
</tr>
</tbody>
</table>

Kurtosis .918  Skewness .787

67
The main outcome measure used in the studies which follow in later chapters is reconviction. The next section reviews the role of reconviction as an indicator of the success of offender-interventions.

Reconviction as an indicator of the success of sentencing

This section considers how reconvictions stand as a measure of the outcome of offender-interventions, and looks at how information about reconviction can be used to determine whether a programme is effective in reducing offending. Reconvictions have been the main indicator used to determine interventions' impact on offending, however there are a number of difficulties with the use of reconvictions for this purpose, which are explored below:

- sentences have aims besides reducing offending;

Within the criminal justice system sentences have many aims besides that of reducing further offending. Classically these objective are retribution, reparation, deterrence and denunciation. The reconviction rate of an individual sentence can do nothing to indicate the extent to which these aims are achieved. However whilst reducing reoffending is not the only aim of a sentence, it is an important aim of the criminal justice agency entrusted with supervising the offender. Reducing reoffending furthermore in the early 1990s became an explicit aim of probation (see for example Home Office Probation Circular CPO 13/92).

Furthermore, whilst reconviction rates may not serve to help evaluate the multiple aims of a sentence, reconviction rates are critical for measuring the success of
programmes and interventions run within a given sentence. Each of the programmes evaluated in later chapters was run for offenders who were on probation. The success of probation as a disposal of the court may need to be evaluated from a number of perspectives, according to the extent to which it serves aims of punishment or deterrence. But it is entirely legitimate to compare interventions within this category of sentence according to their aims, which in this case were largely relate to reducing offending.

Mair (1991) sets out a model of primary and secondary measures of the effectiveness of a sentence. For a community penalty, primary measures might include reconvictions, measures of diversion from custody, financial costs, and sentencer and programme participant satisfaction. Secondary measures would relate to the specific objectives of a given programme, and might include measures of help with employment, accommodation etc.

- reconviction rates may not be seen by probation officers as a legitimate measure of their work;

Some probation officers are reluctant to accept reconvictions as a legitimate measure of the effectiveness of their work (Humphrey and Pease 1992). That those charged with carrying out work with offenders may feel this way is more challenging. Such attitudes may arise from feelings amongst officers that they have no influence over the offending of those under their supervision. These attitudes may be changing as probation officers become more aware of the growing body of evidence that this is not the case (see work described in the previous chapter: for example McIvor 1990; Roberts 1990; Lipsey 1991), and as probation services focus more closely on reducing crime as one of their key aims (see for example ACOP 1997.)

The remaining difficulties with reconviction rates as measures of the success of interventions are more to do with the technicalities of how the information is
produced, and is to be interpreted.

- **reconvictions may not be an accurate measure of offending**;

It can be argued that an individual's criminal convictions may not offer an accurate measure of his or her offending, given that only 3 percent of offences result in a conviction (a finding of the British Crime Survey, reported in Barclay 1993). However a number of studies have been carried out which compare officially recorded delinquency and self-reported offending for the same group. Generally these show a good deal of congruence once methodological, measurement or selection biases have been taken into account. For example Bottoms (1995) carried out the first major offender treatment evaluation in the UK which used both official criminality and self-reported delinquency as criteria for assessing reoffending. He found that, for the most part, very similar results were obtained whether official (reconviction or caution), or self-report measures of offending were used.

There appears to be a consensus that the two approaches are complementary. Tarling (1993) describes Farrington's (1989), comparison of official records and self-reports as follows:

'members of their cohort answered correctly about their contacts with the police, thus providing further confidence that people are prepared to give truthful answers in such studies. The proportion who denied committing the offence of a particular type but who were in fact convicted of such an offence was less than 4 percent.' Tarling 1993: 6.

This implies a high level of congruence between an individual in the criminal justice system's actual, self reported and officially recorded offending.

A further finding from this study was that offenders who admitted committing many offences were more likely to be caught and convicted at some stage in their career. More than half of those who committed burglary or theft of a vehicle received a
conviction for these offences. This dropped to about a quarter for theft from vehicles and further to one in eight for those who committed shoplifting or assault. Low percentages were found for theft from machines, theft from work, drug use (all at 7 percent), vandalism (6 percent) and fraud (3 percent).

The results of these and other comparisons made by Farrington (1989a) led to the conclusion official records and self-reported offending were measuring the same underlying concepts, though with different measurement biases. Furthermore, he suggests they are comparable and complementary measures, both of which help advance knowledge about delinquency and crime.

- **crude reconviction rates do not take account of the seriousness of reoffending;**

Simple measures of reconviction do not take into account seriousness. A bipolar measure of offending treats offences such as shoplifting and armed robbery as equally serious. But reconviction need not mean failure: if someone who frequently commits robbery does not offend for a period, and then shoplifts, this might be regarded as success. However this criticism of bipolar measures is most valid at the level of the individual. With sufficiently large groups, differences in type and seriousness of re-offending can be averaged to a group mean, amenable to statistical tests, to see whether differences exist, or if changes have occurred.

There is evidence that for most offenders, offending is a uni-dimensional construct, in that seriousness and frequency of offending seem to be predicted by the same factors. This is based on the finding from research into criminal careers (see for example Farrington, 1992), that for most offenders (not sex offenders or a small group of offenders who specialise in fraud) the variables that predict serious offending are largely those which predict offending generally. This finding would suggest that whilst it is useful to show the success (or otherwise) of an intervention by reporting the proportion of reconvictions, and number of reconvictions, and also
proportion of reconvictions for certain classes of offence, or at a certain level of seriousness, for the most part it is unlikely much extra information will be conveyed, as these are all measures of the same underlying dimension. In the research-based chapters which follow, a variety of different reconviction measures are used in each case, so this matter will be revisited.

- **reconviction rates may only measure the performance of criminal justice agencies in processing offenders;**

The lengthy process from the detection or reporting of a crime, the decision to prosecute, successful prosecution, conviction and sentencing means changes in reconviction rates may occur as the result of the action any of a number of agencies in the criminal justice system or outside. Reconviction in fact refers to what Martinson (1974) calls 'reprocessing' by the criminal justice system. A number of the different agencies involved in this reprocessing might regard conviction as a performance indicator, contingent on their performance as much as on the behaviour of offenders. Of course whilst this may be the case, reconvictions remain important performance indicators for interventions which seek to reduce offending, a main aim of the probation service.

- **there is no agreement about the length of follow-up periods, when they should start and what they mean;**

There is no standard agreed follow-up period. The majority of studies use a two-year follow-up period for reporting reconvictions, but this is merely a convention which not all studies follow, and which does not have a clear basis. Furthermore it is not clear whether convictions should be counted from time of sentencing, or from end of sentence. This is most important in the case of short custodial sentences where reconvictions can be counted either from time of sentence, or from time of release from prison. Counting reconvictions from time of sentence, including

72
periods of incarceration, gives the best measure of how the public may have been protected from crime. However counting from time of release may give a better measure of rehabilitative impact (or otherwise) of the sentence on the offender. In the studies which make up the later research-based chapters both methods are presented. This matter is returned to below under the heading 'comparison group issues'.

- **crude reconviction rates do not take account of pseudo-reconvictions**;

This is an entirely technical matter. Some convictions which occur during or even after a sentence will be for offences committed before the start of the sentence, due to the speed at which the prosecution and court processes operate. This matter is explored in some detail by Lloyd et al (1994). There are different ways in which it is possible to address the problem of pseudo-reconvictions, which are illustrated by the approaches taken in each of the research studies which make up later chapters:

- looking at outcomes in the long-term, so short-term effects such as pseudo-reconvictions become less important. Chapter 3, looks at long term reconviction rates for an initiative known as the Demonstration Unit, exploring detailed trends in reconvictions over a ten-year period.

- using information about the occasion on which the offence was committed. Chapter 4, looks at the reoffending of probationers who attended a motor project, and uses information about arrests, rather than convictions, and so eliminates the problem altogether. The criterion measure here is the occasion on which an offence is committed, and for which an offender is apprehended (at the time or later), when this was not followed by a not-guilty finding.

- using a standard multiplier for specific classes of community sentences, based on large-sample studies such as Lloyd et al 1994. This is probably the least satisfactory technique, but is the easiest to use, and is applied in Chapter 5, which
reviews Sherborne House.

- taking into account the length of time taken to process reconvictions, and thus discounting sentences passed in magistrates' courts within one month, and in crown courts within three months of a reference conviction. This procedure is used in Chapter 6, a review of an intervention for offenders run at Camberwell probation centre.

Although each of the criticisms set out above can be answered, together they retain force. A simple bipolar measure of reconviction is not an ideal measure of an individual's offending; reconviction rates are not easy to interpret, and are not the only outcome of sentences. However reconvictions are important outcomes which relate directly to the objectives of probation orders (Home Office CPO 13/92). As Mair and Copas (1997) observe, measurement of other important areas of human life and social functioning (poverty, drug use, employment, intelligence, mental health) is equally difficult, and lacking in consensus. Simply that crime is difficult to measure does not invalidate reconviction as a useful measure of crime. Difficulty of measurement is not a reason for not measuring crime. The next section looks at ways of making sense of information about reconvictions.

Making sense of reconviction outcomes

It is not possible to know whether a given reconviction outcome is good or bad without some basis for comparison. This can be provided by the following:

- control groups;

Control groups are characterised by random assignation to experimental or control group. This is the only way of controlling for bias and extraneous influences, and such studies represent the gold standard in effectiveness research. But true random
allocation between types of sentence, or at point of sentence, is rarely feasible in
the UK, especially with regard to community sentences. This is because such
studies would have to intervene in the sentencing process. A few studies which use
an element of random assignment to alternative styles of community supervision
have been carried out within a given type of community sentence (see for example
Folkard, 1976). Random assignment to treatment conditions is generally more
feasible when assigning offenders to programmes within a sentence type (for
example within a prison, assigning offenders to pre-release courses of differing
types.)

In practice random assignment has a tendency to break down. As Robinson (1995)
reports of one of the largest recent studies which attempted to randomly allocate
identified suitable offenders either to a treatment programme, or to a waiting list
control

'the maintenance of the randomised procedure proved to be unpopular in the field. There was a
preference to use other priority criteria (e.g. proximity to release and programme need severity) as a
method of assigning eligible candidates to waiting lists.' (page 10.)

The use of random assignation was discontinued before the end of this important
study.

- comparison or pseudo-control groups;

More common than random allocation is non-random allocation to groups. This is
the main technique used in later research-based chapters. Assessing the impact on
offending of an intervention means knowing how much crime those who attended
would have committed, had they not experienced the intervention. In a quasi-
experimental research design, the offending of the group attending a project is
compared with that of a comparison group. The comparison of pseudo control
group method is not without flaws.
Where groups of attendees and non-attendees are not selected randomly, differences in reoffending outcomes may be due to differences between the groups unrelated to treatment. To demonstrate the effectiveness of an intervention, the need is to show that differences in offending outcomes are not the result of pre-existing differences between the groups. In the studies which follow, the comparison group was made up either of offenders who were referred to the programme, but did not start; or from a matched group of offenders selected for the purposes of making comparisons. The usual reason for failure to start was that the court passed an alternative sentence. Thus there is no element of random assignation to conditions. In each case there is therefore a need to demonstrate the comparison group is similar to the group attending the programme, especially in terms of potentially criminogenic factors. This is so that the subsequent offending career of the comparison group can be taken to indicate what the group attending the programme would have done, had the programme not been there. This also means that the comparison group should have been sentenced for a similar offence, and given a sentence that could constitute a plausible alternative to the programme in question.

Always the need is to demonstrate that comparison and experimental groups are sufficiently similar at the start of the intervention in question for the subsequent offending of the comparison group to be taken as a reasonable illustration of what the experimental group would have done had the intervention not existed.

- **before / after measures;**

Frequency and seriousness of reconviction can be assessed immediately before and immediately after a programme. This can be done using self-report measures of offending (see for example Burnett, 1994, Mair and Nee, 1992), or via officially recorded convictions (see for example Raynor and Vanstone, 1995). This technique is also used in later chapters. The drawback with this approach is that offenders change over time, for example it is also well known that offenders are less likely to commit crime as they get older (see for example Lloyd et al, 1994). So it may not
be straightforward to determine whether changes are related to the intervention in question.

Another problem is related to looking at the nature and seriousness of crime committed before and after an intervention. Criminal careers research shows that most offenders do not appear to specialise in any particular type of offending. For example the best predictor of future violent offending is not previous offences of violence, but simply number of previous offences (see for example Farrington, 1992a). An offender who has committed a violent offences is as likely at their next offence to commit a property offence. However if evaluating a programme for violent offenders, it could on this basis be concluded the programme was successful in reducing violent offending. But this would be to misinterpret the nature of criminal careers. The hypothetical programme which targets offenders immediately after they have committed one type of offence, is very unlikely to see the offenders committing the same type of offence in similar proportions in the future, regardless of intervention. (An exception here may be made for sex offenders, who tend to be much more specialised in their offending careers.)

- **comparison with similar projects, or comparison over time;**

Comparisons with similar projects, or between years for the same project are really only useful where it can be shown that similar offenders attend the programmes, so that variation in outcome can be related to the intervention rather than to differences in the groups of offenders attending programmes. In particular it is necessary to show that offenders are at a similar risk of reconviction across disposals or over time.

- **prediction techniques;**

The use of prediction techniques for evaluation of probation effectiveness is

the use of prediction techniques, with the outcome in terms of offending rates compared with the outcome expected. Further development of such methods may make it possible to apply them locally in a systematic fashion to give feedback to practitioners on a routine basis (Page 4)

Prediction techniques in this context involve using information about the offender to make an assessment of the likelihood of reconviction within a given time period. The predictors used in this paper are the Offender Group Reconviction Score (OGRS, Copas and Marshall, 1998, Home Office Probation Unit, 1996) and the very similar forerunner to OGRS, the National Risk of Reconviction Predictor (Copas et al 1994, Home Office, 1993). The PSR Predictor (Wilkinson, 1994) is also used. All calculate risk of reconviction on the basis of age and previous recorded offending. To this the OGRS and NRRP add gender, the PSR predictor adds employment status.

Prediction techniques allow predicted and actual rates of reconviction to be compared. This makes it possible to compare offenders' actual reconvictions with what would be predicted, on the basis of their gender, age and previous offending (see for example Raynor, 1997, summarising the results of the STOP programme.) Using predicted likelihood of reconviction also strengthens comparison group methodologies, allowing comparisons to be made of offenders at similar risk of reconviction across disposals, and allowing it to be shown whether offenders in a comparison group are more or less likely to be reconvicted.

**Comparison group issues**

Certain issues recur when using comparison groups. Statistical issues are described above. These relate to ensuring and demonstrating that comparison and intervention
groups are drawn from what is essentially the same population, so that outcomes for
the intervention group reflect what would happen to the comparison group if the
intervention were applied to them, and vice versa. However there are issues specific
to offender-programmes, which relate to comparing reconviction outcomes for
intervention and comparison groups.

**Whether to look at reconvictions for everyone who started a programme, or
for completers only?**

It is sometimes argued (for example by Robinson, 1995) that only by looking at
reconviction outcomes for offenders who complete a programme can the impact of
the intervention be assessed, as it is only those who have completed the programme
who can be said to have had the ‘full dose’ of treatment. If the majority of people
who drop-out from a programme do so in its first one or two weeks (as was found
for example by Maitland and Keegan, 1988), then it is understandable why it may
be felt that the subsequent offending of people who only attended a few sessions of
a programme is unlikely to show the value of treatment.

But it is only when looking retrospectively that offenders who drop out in the first
weeks of a programme are identified. If in a prospective study early drop-outs could
not be identified as a separate group, it would be wrong to separate them at the
analysis stage. From the point of view of a sentencer deciding on sentence, where a
treatment intervention is proposed, what matters is an individual's chance of
success, not their chance of success if they make it through the first weeks of a
programme to become members of a sub-group which can only be determined post-
hoc.

If the reconvictions of completers of community programmes are to be analysed as
being those of a separate group, then they should be compared to the reconvictions
of a group which has also completed a community programme, but with different or
no rehabilitative input. For example completers of a rehabilitation programme could
be compared with completers of a period of community service. What is not admissible is when a group of offenders who have completed a programme are compared with a completely heterogeneous group of offenders (as is the case for example in Robinson, 1995.)

A programme could be hypothesised which involves offenders attending a centre, signing in, and two hours later leaving, with no treatment having taken place. Over a, say, three month period a proportion of these offenders would be likely to reoffend and be reconvicted, leaving the 'programme'. So offenders remaining at the end of the programme would already have a lower reconviction rate. Other offenders would simply stop coming at some point over the three months. Either way, offenders still signing in after three months are more likely to be committed to desisting from further offending, than are the offenders who dropped out. The lower offending of completers compared to drop-outs would be a selection rather than a treatment effect. It is only when the offending of the whole group of completers and drop-outs is compared to that of a similar group that the presence or absence of a treatment effect is revealed.

Comparing reconviction rates for community and custodial sentences

An issue when comparing reconviction rates of community and custodial sentences is the point at which the comparison should be made: whether this should date from date of sentence, or date of completion of programme or sentence. For community sentences, date of sentence is most commonly used. Date of completion may be harder to define, in the case of probation orders, and may vary for completers or point of sentence.

For custodial sentences this point may be from time of sentence or from time of release. With longer custodial sentences, as in the case of studies of the effectiveness of parole (see for example Hann et al 1991) it is logical to measure from time of release, as comparison groups will generally also have been in
custody, and be being released to live in the community.

However with shorter custodial sentences it is feasible to look at offending either from time of sentence (see for example Roberts, 1988, Wilkinson, 1997, and Chapters 3 and 4 in this thesis), or from time of release. The justifications for looking at reconvictions from time of sentence are as follows:

- **measure of community safety;**

Comparing reoffending after community and custodial sentences from time of sentence provides a measure of how the community has been protected from offending. Whilst offenders are in custody they cannot offend against the community (though they may commit further criminal offences.) This has been a consideration for policy makers (see most notably Michael Howard's 'Prison Works' speech at the Conservative Party Conference, October 1993.) It is likely to be a consideration for sentencers and for members of the public when comparing custodial and community sentences. Which type of sentence is going to save them most from offending? Counting reconvictions from time of sentence gives custodial sentences a built-in head-start, in terms of measuring community safety. The incarceration period means that for a time the community is protected. However some studies have shown (Roberts, 1988, Wilkinson, 1997, 1994), that once released, offenders sentenced to custody may offend at a faster rate than a similar group treated in the community, so that over a lengthy follow period, the incarceration advantage is eliminated.

- **fair comparison from sentencing decision, within a given seriousness band;**

Following the previous justification, a sentencer choosing how to sentence an individual has as options sentences at roughly equivalent degrees of seriousness
based on seriousness of offence and previous convictions. In the case of a conditional probation order, the alternatives are likely to be community service, combination orders, or short (under a year) custodial sentences. One of the considerations in this decision is which sentence is most likely to save the public from crime. Custody will be seen as achieving this.

- **practicality**;

If accurate (or even estimated) information about date of release is not available, custodial sentences are short, and follow-up periods long (as is the case, for example in Wilkinson, 1997), comparing reconvictions from time of sentence may simply be more practical. Given reconviction is not a sensitive measure of offending, over a long enough follow-up period any significant treatment effect of a community programme should outweigh temporary incarceration effects.

- **offenders on community sentences may have periods of incarceration during a follow-up period**;

The number and seriousness of convictions in the years after a sentence, for the majority of offenders who reoffend, will represent the outcome of many sentences. An offender reoffending whilst on a community sentence may be sentenced to custody, and not offend again in a two year period. Over a follow-up period offenders may have spent time on custodial remand, and be unable to offend in the community, but be found not guilty, or sentenced to a community penalty. These unknown incarceration periods will not be apparent or taken into account in subsequent analysis of offending, so neither should be known incarceration periods.

- **community sentences also contain an element of restriction**;

82
Whilst an offender is attending a programme, which may be for several hours each day or in the evening, they are not available to offend in the community. This restriction on liberty is analogous to that applied on an offender in custody, but not to a prisoner following release.

— comparing like treatments;

An offender leaving prison could be seen as leaving “treatment”, or may actually be leaving a treatment programme carried out in custody. An offender starting a community programme is starting treatment. It is not reasonable to compare an offender leaving treatment with an offender starting treatment.

However there are also arguments for comparing offenders starting community sentences with offenders leaving custodial sentences:

— equivalence of outcome measures;

Offenders in custody may still offend, against the institutional discipline, or against other offenders, however these offences are usually dealt with via prison disciplinary procedures, and do not become reconvictions. In contrast offenders who offend whilst in the community are usually dealt with through the courts, and reconvictions occur. Thus reconviction is not valid as an outcome measure for the period in custody, and only should apply from release.

— measuring treatment effect;

Given that reconvictions as an outcome measure apply only whilst offenders are in the community, to gauge the effect of treatment the measure should only be applied whilst it is valid.
On balance, these arguments may favour measuring from time of release, especially where a quasi-medical model of treatment intervention is used. However if the interests of the wider community are taken into account (and it is these which provide the terms of reference for defining 'offending' and 'rehabilitation' in the first place, see Chapter 1), and the main issue at stake is comparing sentencing alternatives, then it seems apposite to compare offending from time of sentence. Later research-based chapters use both approaches, in order to attempt a balanced assessment of programme impact on offenders, and on the community.

Extending the evidence-base: selecting probation programmes for evaluation

The remaining chapters in this thesis are based on original research, using the methods described above, and carried out to evaluate the effectiveness of a number of probation programmes in reducing offending.

Programmes for offenders evaluated in research-based chapters

The five programmes for offenders that make up the main substance of the following chapters are:

- The first Inner London Probation Service Demonstration Unit;


- The first Inner London Probation Service Demonstration Unit;

84
A driver retraining scheme for offenders convicted of taking cars. (Wilkinson, 1996)

- **Ilderton Motor Project;**


- **Sherborne House Probation Centre for young offenders;**

A day-programme for serious and persistent offenders aged between 17 and 21 years (not previously published.

- **Camberwell Probation Centre;**


**Criteria for selecting programmes for inclusion in research**

The main criteria for selecting these programmes for evaluation and inclusion in this thesis is that all should be:

- **significant;**

All work with significant numbers of offenders, and are relatively well known in probation circles; all are (or in the case of the Demonstration Unit projects described in the first chapter, were) well-resourced.
- mainstream;

All are clearly in the main-stream of probation practice in terms of the nature of the interventions they use, and the offenders they target (see Tables 2.2 and 2.3 below);

- influential;

All provide models of practice which have been taken as models for use elsewhere;

- work with high-risk offenders;

All work with offenders with significant numbers of previous convictions, who are at high risk of committing further offences. This was a main focus of probation activity at the time of the research, which took place at a time when imprisonment rates were starting to rise again, after a sustained period when government policy was deincarceral in intent (see for example Vanstone, 1993). However more recently some of the more serious offenders who attended these projects would have been more likely to be sentenced to custody as a result of 'a general trend toward more punitive sentencing' (Home Office, 1997, page 4 para 25.)

- clear that reducing offending is an immediate and primary aim.

In all cases these are projects which set out to reduce the offending of those who attend.

In addition all are run (or in the case of Ilderton Motor Project funded and steered) by Inner London Probation Service (ILPS). At the time of the field-work the author and main researcher was employed by ILPS.
In the context of debate as to publication effects, at the time of the research, all could be regarded as unpublished, in terms of evaluation reports. Though there have been some subsequent publications resulting from the work reported here, publication was not a criterion for their inclusion in the current review, nor was a successful outcome in terms of crime reduction.

Probation's work with offenders: case-work

Most of the work probation does with offenders is one-to-one case-work. This takes place in office-based meetings between probation officer and offender, the form and frequency of these meetings is set out according to national standards (Home Office, 1995). These meetings may involve advice and information giving, practical help and crisis intervention and may use counselling interventions in more or less structured forms.

The new research set out here did not set out to evaluate this type of mainstream work. Whilst in some ways this represents an omission, there are good reasons for concentrating on group-programmes and centre-based programmes. Firstly these comprise an increasing proportion of probation practice. These programmes are often the testing-ground, whereby innovations (such as cognitive behavioural work with offenders) are first introduced into probation practice.

There are practical reasons why group-programmes are easier to evaluate than one-to-one work. Work in group-programmes tends to be higher-profile, more visible and more transparent than one-to-one casework, which takes place behind closed doors. Evaluations of one-to-one casework, are also complicated by issues of 'matching', whereby the way in which individual probation officers and offenders work together come to the fore (see for example Palmer, 1973.) Such evaluations are also complicated by the fact they may expose differences in the abilities of individual probation officers.
Evaluations of one-to-one case-work are entirely feasible, and when high-quality work of this type is carried out it yields very interesting results (see for example Stanley and Goldberg's 1984, evaluation of task-centred case-work in probation practice.) As some of the principles of cognitive-behavioural work are introduced into one-to-one case-work (see for example Priestley's Reducing Reoffending Programme, described in Underdown, 1998), the need for further research of this type increases.

Probation Interventions

Information on the number and nature of probation projects is not readily available. The best available source is the Index of Probation Projects (NPRIE, 1994), a database which encompasses all types of project-work with offenders. Table 2.2 shows the results of a search for the main keywords used in the database, omitting keywords occurring in 9 or fewer projects. This gives some illustration of the range and number of projects, though it is important to note that key-word categories are not mutually exclusive. Thus details of 319 group-work programmes were entered into the index, which probably included details of 125 offending behaviour programmes, and may or may not have included details of 59 auto-crime projects.

As might be expected, most of these projects have a more or less direct goal of reducing offending. Many are specific to certain types or classes of offence. Table 2.2 shows that the five Inner London Probation Service projects reviewed in later chapters fall very much within the mainstream of probation projects. A better picture may be provided by Hedderman and Sugg (1997), who surveyed all probation areas for details of any programmes for offenders which contained a cognitive skills component. Their initial results are set out in Table 2.3, together with a summary of the programmes included in the current research, and evaluated in later chapters. This survey too shows that the projects included in the current research are very much in the mainstream of probation interventions.
Table 2.2. Selected extract from the Index of Probation Projects (NPRIE, 1994) Note: key-word categories are not mutually exclusive.

<table>
<thead>
<tr>
<th>key-word</th>
<th>number of entries</th>
<th>programme element included in evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>abuse</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td>accommodation</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>activity groups / centres</td>
<td>28</td>
<td>yes</td>
</tr>
<tr>
<td>alcohol</td>
<td>128</td>
<td></td>
</tr>
<tr>
<td>anger management</td>
<td>49</td>
<td>yes</td>
</tr>
<tr>
<td>assertiveness</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>auto-crime / motor projects (motoring also gets 45 entries)</td>
<td>59</td>
<td>yes</td>
</tr>
<tr>
<td>bail</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>basic education</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>burglary</td>
<td>13</td>
<td>yes</td>
</tr>
<tr>
<td>centres, including...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>probation centres</td>
<td>36</td>
<td>yes</td>
</tr>
<tr>
<td>day-centres</td>
<td>36</td>
<td>yes</td>
</tr>
<tr>
<td>counselling</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>crime prevention</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>drama</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>drug abuse / addiction</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>drink-driving</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>education</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>employment</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>group-work</td>
<td>319</td>
<td>yes</td>
</tr>
<tr>
<td>intensive programmes</td>
<td>10</td>
<td>yes</td>
</tr>
<tr>
<td>money management</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>mentally disordered offenders</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>offending behaviour</td>
<td>125</td>
<td>yes</td>
</tr>
<tr>
<td>sex offenders</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>social skills</td>
<td>12</td>
<td>yes</td>
</tr>
<tr>
<td>victims</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>violence</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>women offenders</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>young offenders</td>
<td>25</td>
<td>yes</td>
</tr>
</tbody>
</table>

Table 2.3. Probation-programmes for offenders which contain a cognitive component, February 1996
(adapted from Hedderman and Sugg, 1997, Table 2.1)

<table>
<thead>
<tr>
<th>Programme Type</th>
<th>No. of programmes</th>
<th>Included in research</th>
</tr>
</thead>
<tbody>
<tr>
<td>self control (inc. R&amp;R)</td>
<td>34</td>
<td>Chapter 6: CPC</td>
</tr>
<tr>
<td>offending behaviour</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>sex offending</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>substance abuse</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>motoring</td>
<td>37</td>
<td>Chapter 3: Demo Unit, Chapter 4:</td>
</tr>
<tr>
<td>Illderton</td>
<td></td>
<td></td>
</tr>
<tr>
<td>domestic violence</td>
<td>7</td>
<td>Chapter 3: Demo Unit</td>
</tr>
<tr>
<td>burglary / theft</td>
<td>2</td>
<td>Chapter 5: Sherborne House.</td>
</tr>
<tr>
<td>probation centre programmes</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>women offenders</td>
<td>7</td>
<td>Chapter 3: Demo Unit</td>
</tr>
<tr>
<td>other</td>
<td>14</td>
<td>Chapter 5: Sherborne House.</td>
</tr>
<tr>
<td>total</td>
<td>191</td>
<td>5</td>
</tr>
</tbody>
</table>
Evaluation of probation programmes

Underdown and Ellis (1998) surveyed all evaluated probation projects, with the aim of identifying examples of good evaluation practice. The survey was carried out for HM Inspectorate of Probation, based in the Home Office, via letters to Chief Probation Officers. Perhaps not surprisingly given this body’s regulatory function over the service, there was a very high return rate (96 percent) of questionnaires. This survey can be regarded as offering a definitive statement as to the standing of evaluative research in probation.

Underdown and Ellis received replies relating to 267 programmes. 30 sex offender programmes were excluded, being deemed to have been adequately evaluated elsewhere (see Hedderman and Sugg, 1996). A further 27 programmes were excluded because they did not supply start or finish dates, or were not programmes for offenders. However of the remaining 210 evaluated offender programmes, under 16 percent (33 programmes) of returns showed any evidence that programmes were being evaluated in terms of measurable change, or effectiveness. Amongst these programmes only four instances were found of ‘well evaluated programmes which provided some evidence of success’ (Underdown and Ellis, 1998, page 108, para 9.4.3). Three of these four instances of good practice make up chapters 3, 4 and 6 of this thesis. (The one other instance of good evaluation practice in this Inspectorate report is provided by the STOP Programme Evaluation, see for example Raynor and Vanstone, 1996.)

The next chapter presents the first case study which describes the evaluation, in the long-term, of the effectiveness of two linked programmes which carried out offence-focused work with offenders designed to bring about reductions in offending.
Chapter 3.

A long term follow-up study of the first Inner London Probation Service Demonstration Unit.

This chapter examines the long-term effect on reconviction of the first Inner London Probation Service Demonstration Unit. This unit, in effect an action research project, was set up in 1981, and ran until 1985. It consisted of a specially convened team of six probation officers, who carried out two separate community-based offender-interventions.

A main aim of the Demonstration Unit was to show that 'offenders could be held safely in the community' without committing further offences (Harraway et al, 1985). This was at a time when the probation service was still gripped by the notion that it was not possible to reduce offending via work with offenders. This idea had arisen from the negative research findings summed up under the banner 'nothing works', described in Chapter 1. It was partly in reaction to the 'nothing works' position that the Demonstration Unit was set up.

As well as contributing to the 'what works' debate, the Demonstration Unit is of intrinsic interest. Much of the work of the Unit was at the time considered highly innovatory, but would in the mid-90s be located within the mainstream of probation practice. Its main innovations are as follows:

- pioneering the offence focussed approach, whereby work done with offenders is related to their offending. In the early 80s more usual probation practice was to provide social work help focusing more on offenders problems, ususally related only indirectly to their offending;

- targeting offenders at relatively high risk of reoffending;
- targeting offenders who were otherwise likely to be sentenced by way of custody;

- using social-skills training techniques amongst its methods of intervention.

This review of the impact of the Demonstration Unit represents the first long-term study of the effect on reconviction of these ways of working.

The inception of the Demonstration Unit

The context of the setting up of the demonstration unit was one of perceived problems with the probation order:

The probation service is once again in a mood for self-doubt and self-questioning. Probation orders are declining as a proportion of all court sentences; and probationers constitute a minority among probation officer's caseloads...Most importantly of all, the belief in "treatment" and "rehabilitation" as the centre of the Service's activity has begun to erode in the face of empirical demonstrations of treatment effectiveness and theoretical critiques of the clinical model on which probation casework has largely been based in the post-war period. (Howard Journal, 1978)

The demonstration unit started with "a general idea about doing something with probation orders" (Harraway et al 1986). This was resolved into aims which addressed both the concerns outlined in the above passage from the Howard Journal: addressing decline in the proportionate use of probation and demonstrating its effectiveness in containing offending.

Reasons for the decline in use of probation

The Demonstration Unit team reviewed what they termed the "official view" of the decline of the probation order, which was that there had been a loss of faith in probation. Probation officers had lost credibility with the courts because of probation officers' confusion as to their role (Bryant et al 1978). Linked to this was
the perceived doubtful caliber of new recruits to the service, trained as social workers and preoccupied with social work ideas (Central Council of Probation and Aftercare Committees, reported in NAPO News, 1978). The view of the Demonstration Unit was that these arguments were not supported by other than anecdotal evidence. They attributed the decline in probation in part to legislative changes (the Children and Young Person's Act 1969; Criminal Justice Act, 1972). These had decreased the proportion of those for whom probation would be a sentencing option; and changed the balance of the services' workload, so that more time was being spent with offenders released from custody and preparing reports for court.

However the Demonstration Unit team identified a deeper and more intractable problem. This was that there had been a loss of confidence amongst probation officers in their ability to secure rehabilitation of offenders, following research findings appearing to show that probation was not effective at securing rehabilitation. These studies most notably included the Impact Experiment, (Folkard, 1976) which showed mixed findings for offenders given more intensive probation supervision, and what was strongly presented as the "nothing works" message based on extensive literature review (most widely cited being Martinson, 1974, and Brody, 1976).

**Addressing decline in the use of the probation order: the operational objectives of the Demonstration Unit.**

The Demonstration Unit team set out to address the decline in use of the probation order. The relevant operational goal was therefore to get more offenders on probation orders, supervise them effectively and to demonstrate that they had achieved each of these things. The Demonstration Unit team stated that the decline in use of the probation order could in part be addressed by changing the place of probation in the sentencing market, and expanding its use.
The Demonstration Unit officers did not agree that the research even then available meant that 'nothing worked' in reducing offending, as examples of positive research findings a reinterpretation of the impact study (Folkard, 1976), which did show promising findings for certain groups of offenders. They also cited as positive research carried out in Inner London Probation Service research into task-centred case-work (Stanley and Goldberg, 1984). The Demonstration Unit team took the view that effective supervision would need to focus on the offence and its meaning, rather than try to change everything in an offender's life.

Demonstrating that they had achieved their goals of increased use of probation and effective supervision meant that all relevant aspects of the unit's work had to be monitored. This meant employing researchers to advise on what and how to monitor, and how to interpret the results of monitoring. It also entailed buying and using computers and acquiring expertise in data-analysis. Finally it meant publicising their work and disseminating findings through publications, and through PR and training events.

**Providing an alternative to custody**

A key aim of the Demonstration Unit was to 'raise probation higher up the sentencing ladder' (Harraway et al, 1985, page 147), so as to offer an alternative sentence to imprisonment. At this time, offering alternatives to custody was being proposed as a main role for the probation service (see for example Pointing, 1986). A common view in probation at the time was regardless of whether probation could be effective in rehabilitating offenders, it was infinitely preferable to prison, which harmed offenders whilst being expensive to the public purse (see for example Willis, 1986).

**Identifying opportunities for increasing the use of probation**
The first part of the unit's work was to find opportunities for expanding use of the probation order. The Demonstration Unit Team set about this task by looking for groups of offenders, defined by type of offence and by number of previous convictions, for whom probation might have been under-used. The idea was that having identified such a group of offenders, they could then be targeted in order for the use of probation to be expanded.

They were assisted by information from a then recent survey of social inquiry reports and sentencing outcomes (Stanley and Murphy, 1984), and by information on sentencing practice obtained from further surveying of local magistrates' courts (reported in Harraway et al 1985). Equipped with this information the Demonstration Unit officers were able to search for groups of offenders for whom probation was less likely to be recommended; or where when recommended by probation officers, was less likely to be accepted by the courts as a sentencing option.

Looking at what was termed "the sentencing market" by category of offence, two offence types stood out as providing opportunities for increasing proposals for probation: Taking and Driving Away a Motor Vehicle (TDA) and Burglary. The decision was made to target offenders who had committed offences in these categories for the following reasons:

- offences of TDA and burglary netted fewer probation orders and attracted more high tariff sentences than did theft, the most common offence category.

- these categories together comprised a reasonably large proportion of all offences sentenced in the magistrates' court (they accounted for just over a fifth of all offences, not counting breaches in 1981, see Harraway et al 1985, page 6). Thus it seemed very likely they could be relied upon to generate a reasonable number of SIRs.

- this type of offending was in the mainstream of probation work (rather than
in a specialist area as would be for example work with sex offenders);

these were reasonably homogenous offence types. The terms TDA and burglary apply to recognisably similar types of offence. This was not the case with theft.

The further decision was taken only to target those offenders who already had at least two prior convictions, one of which must have occurred in the previous two years. This was in order to keep the number of offenders to be dealt with down to a manageable level. It also had the positive effect of preventing first time offenders from being brought inappropriately into the system. Furthermore there were indications from the survey of social inquiry reports (Stanley and Murphy 1992) that as number of prior convictions increased, probation officers became less likely to recommend probation. However whether a court accepted a recommendation for a probation order did not seem to be affected by the number of previous convictions. This gave the Demonstration Unit officers reason for optimism that the addition of previous convictions to the DU's targeting criteria would not lead to a reduced concordance between recommendations for probation and eventual sentence.

So the category of offenders for whom probation was deemed to be under-used, and for whom the attempt was to be made to expand the use of probation, was defined as being offenders:

- for whom a social inquiry report had been requested;

- charged with burglary or TDA, including attempts;

- with at least two previous convictions for any offence, one in the past two years;

- living in Islington, Hackney or Tower Hamlets;
These criteria were applied strictly. All offenders who satisfied these conditions were referred to the Demonstration Unit, which was then responsible for preparing the social inquiry report. All social inquiry reports were written focussing on the offence, rather than on unrelated background factors in the offender's life, which was then the common style of social inquiry reports. Wherever possible the social inquiry report recommended a probation order; only not so doing when this was counter to the wishes of the offender, or was not possible for practical reasons. Courts were also told about what was planned for supervision, the subject of the next section.

**Developing effective methods of supervision: the offence-focussed approach**

The Demonstration Unit set out to carry out what was known as "offence focussed work". This was at a time when the emphasis in conventional probation practice was on the provision of social work help without necessarily making links with offending. By the mid-nineteen nineties focussing on offending in probation practice has become entirely mainstream, and indeed near mandatory. However in the early eighties this was regarded as innovatory.

One thing that both the "traditional view" and the Demonstration Unit explanations of decline in use of probation had in common was that all was not well with supervision. The view of the Demonstration Unit was that a stronger focus was needed on offending:

"we knew that looking at the offence was not a new idea, but we all felt from past experience and from looking through literally hundreds of social enquiry reports... that a focus on the offence was sadly lacking in probation practice...Whatever else the client might present, the offence was something which was present in every case. There was also a view that probation recommendations
tended to depend upon whether or not the client displayed social work needs/problems conceived in a fairly traditional way. If the client had "no need" for probation, despite the fact that he/she might be on their third conviction in the past eighteen months, then a recommendation for probation might be absent.

"We... came to a view that the offending itself... was a problem, regardless of the existence of other non-criminal problems and difficulties... offending was possibly the cause of other types of problems and social distress....

"Finally if our efforts were to yield results of a useful nature to our colleagues, we should concentrate on ordinary mainstream offences. We should avoid over-ambitious experimentation and remain in charge of our own work, which would attempt to meet the modest goal of demonstrating what we did, rather than proving the validity of any one theory or method of treatment."

This concept of working with offenders on factors related to their offending was regarded as a significant innovation, at a time when the most usual form of probation practice was to carry out social work related to offenders' problems in general. Offending, whilst it gained offenders access to this social work, was not necessarily the focus of any work itself.

**Summary of Demonstration Unit aims and methods**

The Demonstration Unit team of six officers, defined their objectives as being to increase use of the probation order, and to demonstrate the effectiveness of the probation order for holding offenders safely in the community. They functioned as a conventional probation team: writing social inquiry reports and supervising offenders on probation orders, whilst monitoring as many aspects of their work as was feasible. The Unit targetted offenders convicted of TDA or burglary, who had two or more previous convictions. The referral process contained no discretionary element: all offenders who satisfied the criteria of offence type and previous convictions were to be referred to the Demonstration Unit.
Methods of supervision

The Demonstration Unit was not wedded to any particular theory of social work, but set out to do work in the mainstream of that encountered by probation officers.

The Demonstration Unit team were supported in their belief they could influence offenders away from offending by research findings which, contradicted the then near consensus in criminology by suggesting work with offenders could have positive outcomes. The view of the Demonstration Unit was that to be effective, supervision needed to focus on offending and on reasons for offending, and this was the focus for all offenders supervised by the Demonstration Unit, (though more 'traditional' help with social problems was also offered.)

Focussing on offending for the Demonstration Unit meant concentrating efforts on the specific type of offence committed. The content of supervision was thus quite different for burglary and TDA offenders. Burglary offenders were supervised mainly via one-to-one work with a probation officer, with about a third of offenders also taking part in group-work. Initial uncertainties about how the burglars' groups should operate, and about the content of one-to-one supervision for burglars were gradually resolved as the unit progressed. The main aims of supervision for burglars were to show that burglary is an unreliable way of obtaining money and to show that offending is not a matter of impulse but is something that can be avoided. They also encouraged the burglary offenders to see themselves as other than offenders. The methods used to do this included social skills exercises, based on advice from two psychologists: Philip Priestly and James McGuire (see Priestly and McGuire, 1978, for examples of exercises similar to those they suggested).

There was more clarity as how the TDA offenders should be supervised; all were to attend the Driver Retraining Scheme (see Haraway et al, 1986, and Harraway, 1986, for a fuller account of this scheme). This scheme was rooted in the idea that most TDA offenders offend out of a desire to drive cars, so were they able to drive legally they would be less likely to offend. The focus was then on legitimatising a
formerly illegal activity. The course lasted 6 months, consisting of between 12 and 16 sessions. These covered the practicalities and responsibilities of becoming a legal motorist. Attending the scheme generally entailed taking driving lessons, for which the offender had to pay. The cost to individuals was up to £100, in a few cases this was covered by loans provided by the Demonstration Unit. Attendance and enthusiasm levels on the Driver Retraining Scheme were high, the Demonstration Unit team attributed this to the clear focus and practical content of the scheme, which proved highly engaging to the probationers.

Evaluation methods

Using a comparison group to help make sense of reconviction rates

The key question in assessing the impact on offending of any project is that of how much offending would have occurred had those who attended the project not done so. A way of answering this is by using a comparison group. In the case of Demonstration Unit this should be a group of offenders similar in as many respects as possible to those who attended the Unit. If the Demonstration Unit does help to reduce reoffending, those who attend should be reconvicted less after attending than offenders in the comparison group over a similar period. An obvious comparison group exists in the form of those targeted by the Demonstration Unit, but who for a variety of reasons did not receive probation orders. However the decision to send an offender to the Demonstration Unit was not made at random. The decision to allocate an offender to the Demonstration Unit group or, by default to the comparison group, was a combination of social inquiry report recommendation and sentencer's decision as to how the offender should be sentenced. For the comparison to have value, there is a need to demonstrate that allocation was not on the basis of factors likely to be associated with offending. The need is to demonstrate the groups are sufficiently similar at the point they were targeted by the Unit, for the subsequent offending of the comparison group to indicate how the Demonstration Unit group would have offended had they not attended.
Sources of information

Findings in this section are based on documentary analysis of material produced by the Demonstration Unit; on re-analysis of data collected by the Demonstration Unit; and on information about reconvictions obtained from the Home Office National Identification Bureau. Information taken from back-up tapes of the Demonstration Unit's own computerised monitoring information was recovered, and made available for re-analysis for 256 of the 269 offenders targeted by the unit.

Information about reconvictions of those targeted by the Demonstration Unit was obtained from the National Identification Bureau of the Home Office, covering all convictions occurring before March 1994. Information could be obtained for only 230 of those targeted by the unit, 86 percent of those for whom SIRs were prepared. For seven of these offenders original monitoring data was not obtainable, meaning that the comparison of those who were supervised by the Demonstration Unit with those who were not can be made for only 223 offenders. However the missing cases seem to be distributed at random, and it seems very unlikely that this attrition from the Demonstration Unit and comparison groups has bearing on the main findings of this study.

Results and discussion

Comparing those who attended with those who did not

There are some differences between burglary and TDA offenders (see Table 3.1). The burglars seem more confirmed as criminals, tending to be older; with more convictions; less likely to be employed; more likely to have a problem with drugs; and at greater risk of reconviction. Risk of reconviction is calculated using the National Risk of Reconviction Predictor (NRRP, see Copas 1994). This shows offenders targeted by the Unit are generally at high risk of reconviction, with little
Table 3.1  Comparing offenders who did and did not receive probation orders

<table>
<thead>
<tr>
<th></th>
<th>burglary offenders</th>
<th>TDA offenders</th>
<th>TDA aged under 23</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>prob. n=51  comp. n=61</td>
<td>prob. n=69  comp. n=42</td>
<td>prob. n=62  comp. n=29</td>
</tr>
<tr>
<td><strong>Average...</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>age at sent</td>
<td>23.7</td>
<td>19.2</td>
<td>18.4</td>
</tr>
<tr>
<td>age 1st sent.</td>
<td>14.4</td>
<td>15.6</td>
<td>15.5</td>
</tr>
<tr>
<td>housing moves</td>
<td>4.4</td>
<td>2.9</td>
<td>2.6</td>
</tr>
<tr>
<td><strong>Percent...</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mainly in work</td>
<td>29 33</td>
<td>52 41</td>
<td>53 48</td>
</tr>
<tr>
<td>drug problem</td>
<td>10 11</td>
<td>3 6</td>
<td>2 7</td>
</tr>
<tr>
<td>prev. custody</td>
<td>53 66</td>
<td>33 52</td>
<td>33 41</td>
</tr>
<tr>
<td><strong>Average no. previous...</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>convictions</td>
<td>6.5 6.8</td>
<td>3.0 4.9</td>
<td>3.1 3.6</td>
</tr>
<tr>
<td>burglaries</td>
<td>2.3 2.6</td>
<td>1.4 1.4</td>
<td>1.3 0.8</td>
</tr>
<tr>
<td>TDAs</td>
<td>1.4 0.9</td>
<td>1.0 1.6</td>
<td>0.8 1.3</td>
</tr>
<tr>
<td>NRRP* score</td>
<td>68 71</td>
<td>64 64</td>
<td>64 64</td>
</tr>
</tbody>
</table>

* Percentage chance of the offender being reconvicted within 2 years, as calculated by the National Risk of Reconviction Predictor (Copas 1994).

Table 3.2  Sentencing for offenders attending the Demonstration Unit or in the comparison group.

<table>
<thead>
<tr>
<th></th>
<th>burglary offenders</th>
<th>TDA offenders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>prob. n=51  comp. n=61</td>
<td>prob. n=69  comp. n=42</td>
</tr>
<tr>
<td><strong>Percent sentenced to...</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Custody</td>
<td>- 26</td>
<td>- 29</td>
</tr>
<tr>
<td>Susp. se.</td>
<td>- 18</td>
<td>- 19</td>
</tr>
<tr>
<td>Sec. 38</td>
<td>- 11</td>
<td>- 5</td>
</tr>
<tr>
<td>Prob.+ Demo Unit</td>
<td>100 -</td>
<td>100 -</td>
</tr>
<tr>
<td>CSO</td>
<td>- 23</td>
<td>- 14</td>
</tr>
<tr>
<td>Fine</td>
<td>- 8</td>
<td>- 26</td>
</tr>
<tr>
<td>Discharge</td>
<td>- 3</td>
<td>- 5</td>
</tr>
<tr>
<td>Def. sen</td>
<td>- 2</td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>- 8</td>
<td>- 2</td>
</tr>
</tbody>
</table>

difference between those who did and did not receive probation orders.

For burglars there are generally no differences between those who attended the
Demonstration Unit (the probation or prob. group in Table 3.1) and those who did not (the comp. group). The main difference between TDA offenders who attended the unit and those who did not is that the latter group is on average two and a half years older. This difference is largely due to a small number of offenders in the comparison group aged up to forty. When offenders older than 22 are excluded, the difference in ages falls to eight months and other differences between the Demonstration Unit and comparison groups diminish. So whatever the grounds for the sentencing decision, they do not appear to have included the criminogenic factors monitored by the Demonstration Unit, or risk of reconviction.

Sentences received by offenders not sentenced to probation (see Table 3.2) range in severity from fines to custody. TDA offenders who did not receive probation orders are more likely to receive a fine or discharge (31 percent compared to 11 percent of burglars). Burglary offenders were more likely to be sentenced to community service (23 percent compared to 14 percent of TDA offenders). A similar proportion of each received custodial or suspended custodial sentences (44 percent of burglars, 48 percent of TDA offenders). As a group the burglary offenders who did not receive probation were sentenced slightly more seriously than the TDA offenders. However the general spread of sentences, and the high level of use of custody, implies that in general these were serious offenders.

The variety of sentences received by the comparison group does pose a problem for the comparison group methodology, in that there is no one "control" condition. The fairly even spread across different disposals means there are not sufficient numbers receiving one particular sentence for it to be regarded as a control condition. However in this research impact on offending is assessed against what would be likely to have happened had those who attended the Demonstration Unit not done so. Had the Demonstration Unit not existed the likelihood is they would have received a similar diversity of sentences. What can also be said of these alternative sentences is that all are imposed on justice grounds alone, none contains any element of social work intervention. Given the close similarities between those supervised by the Demonstration Unit and those not, it seems reasonable to expect
that if the Demonstration Unit does reduce offending, those who attended should be reconvicted at a lower rate than those who did not.

**Impact on reconviction**

Looking first at the burglary offenders, the year after being targeted for the Demonstration Unit fewer who attend the Demonstration Unit are reconvicted than are offenders who do not: 29 percent compared to 47 percent, see Table 3.3. This is equivalent to a 38 percent reduction in the offending of the Demonstration Unit group compared to that of the comparison group. This is statistically significant at p<0.01 in a 2-tailed t-test. This is despite two thirds of the comparison group spending at least some time in custody during this year. Over a two year follow-up period the impact on reconvictions becomes less pronounced, the apparent treatment effect decreasing after about a year. This is a clear representation of a strong though temporary treatment effect.

**Table 3.4. Percentage of offenders reconvicted within one and two years of sentence.**

<table>
<thead>
<tr>
<th>Course</th>
<th>predicted 2 year rate</th>
<th>percent reconvicted within...</th>
<th>1 year</th>
<th>2 years</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Burglary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>probation (n=51)</td>
<td>67</td>
<td>29</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>comparison (n=61)</td>
<td>71</td>
<td>47</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td><strong>TDA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>probation (n=69)</td>
<td>64</td>
<td>48</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>comparison (n=42)</td>
<td>64</td>
<td>36</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td><strong>TDA&lt;23</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>probation (n=62)</td>
<td>65</td>
<td>47</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>comparison (n=35)</td>
<td>65</td>
<td>34</td>
<td>57</td>
<td></td>
</tr>
</tbody>
</table>

The story is very different for the TDA offenders. Within a year of commencing supervision more offenders who attended the Demonstration Unit are reconvicted than are those who did not (48 percent compared to 36 percent, see Table 3.3).
Findings are almost identical when offenders aged over 22 are excluded. This apparent reverse treatment effect continues to apply for all age groups for two years after offenders were targeted for the Demonstration Unit (all these findings are significant at the p<0.05 level using a 2-tailed matched groups t-test). After three years the situation reverts to that which applied before the Demonstration Unit offence, with more of those who do not attend being reconvicted.

It might seem the lower offending amongst those who do not receive probation orders could be the result of an incarceration effect, given 47 percent of offenders (40 percent of those aged 22 and under) receive a custodial sentence and would not be at liberty to offend in the year immediately following the Demonstration Unit offence. However all but one custodial sentence is for less than 6 months whilst lower offending is maintained over two years following the Demonstration Unit offence, so it is unlikely an incarceration effect is responsible for this reduction.

To summarise, burglary offenders who attend the Demonstration Unit offend more immediately before attending, less immediately after and the same amount in the long term as those who do not attend. TDA offenders offend less before, more immediately after and less in the long term than their comparison group. If the comparison groups are valid, the obvious conclusion is that the TDA scheme has at best no impact and at worst may increase offending. In contrast the course for burglary offenders has a strong impact on offending over a one year period. In order to look at why this may be so, it is necessary to look at trends in offending in the longer term.

**Trends in reconviction in the longer term**

This research provides a rare opportunity to look at trends in offending over a ten year period, these trends are represented in Tables 3.4 and 3.5.

**Table 3.4**: the proportion of all offenders convicted of any offence each year for
4 years before and 6 years after the offence which led to their being targeted for the Demonstration Unit.

- the average number of sentencing occasions per offender per year, for 4 years before and 6 years after the Demonstration Unit Offence where the main sentence was a community or custodial sentence. Omitting less serious reconvictions has no bearing on any of the trends reported below.

*Table 3.5:* for burglary offenders only, the cumulative number of community or custodial sentences received by each offender for 10 years after the Demonstration Unit Offence.

- for burglary offenders only, the cumulative number of custodial sentences received by each offender for 10 years after the Demonstration Unit Offence.

**Burglary offenders:**

The trends in reconvictions for offenders convicted of burglary are as follows:

- more than two years before the Demonstration Unit Offence those later to become the probation and comparison groups offend at a similar rate, in terms of proportion convicted each year and average number of community and custodial sentences received by each offender each year.

- in the two years immediately before the Demonstration Unit Offence the probation group offend more. A higher proportion are convicted each year of any offence, and they each receive on average more community and custodial sentences each year.

- at the Demonstration Unit Offence itself by definition all offenders are
convicted causing the peak in offending shown in the tables which is an artefact of this way of looking at aggregated offending careers. The peak in reconviction indicates that at this point in their offending career all offenders are convicted and targeted by the Demonstration Unit. There is no reason why the rate of offending (as opposed to conviction) in this year should be much different to the years' immediately preceding.

- in the year immediately after the Demonstration Unit offence, which includes the period spent on probation attending the Demonstration Unit, a significantly lower proportion of those who attended the unit offend compared to those who did not attend. Significantly fewer offences are committed by the probation group which lead to a community or custodial sentence.

- over the following two years the proportion offending and the number of serious reconvictions per year for the probation group climb to the same level as those of the comparison group.

- Looking at the cumulative number of offences sentenced by way of community and custodial sentences it is apparent the reduced offending at one year and the smaller reduction two years after the Demonstration Unit offence results in a difference in overall number of offences committed still present five years after the Demonstration Unit offence, when those who attended had still on average committed 0.38 fewer offences overall.

- Over the five year period that then follows the gap in offending between the two groups is slowly closed as the probation group offends at a slightly higher rate (as it did before attending the Demonstration Unit)

- The trends in the above two paragraphs apply equally to the average number of custodial sentences received by each offender.
TDA offenders

The trends in reconviction shown in Tables 3.4 and 3.5, are quite different to those of the burglary offenders, and are as follows:

- two years and more before the Demonstration Unit Offence those later to become the probation groups offends less than the comparison group, in terms of proportion convicted each year and average number of community and custodial sentences received by each offender each year.

- in the year immediately before the Demonstration Unit Offence the probation and comparison groups are convicted roughly to the same degree, a similar proportion being convicted each year of any offence, and each receiving on average a similar number of community and custodial sentences each year.

- at the Demonstration Unit Offence all offenders are convicted as with the burglary offenders.

- in the two years immediately after being targeted by the Demonstration Unit those who were sentenced to probation offended more than those who received other sentences. More offences are committed by the probation group leading to a community or custodial sentence.

- however 3 years after the being targeted by the Demonstration Unit, the comparison group starts to offend at a higher rate, a larger proportion offending each year and each offender receiving on average more community and custodial sentences per year. The proportion of TDA those in the probation group offending each year gradually falls, as does the number of community and custodial sentences received by each offender.

- looking at the average number of community and custodial sentences cumulatively, at first those sentenced to probation receive more community and
custodial sentences. However by the three year point those in the comparison group have caught the probation group up, meaning they have started to offend more. Over the next 7 years those in the comparison group offend more, so that 10 years after being targeted by the Demonstration Unit they have each committed on average 0.83 more community and custodial sentences, and 0.8 more custodial sentences.

### Table 3.4  Trends in offending before and after the Demonstration Unit offence

<table>
<thead>
<tr>
<th>years before and after Demo Unit offence</th>
<th>percent reconvicted</th>
<th>ave. no. of community or custodial sentences per offender per year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>burglary</td>
<td>TDA</td>
</tr>
<tr>
<td></td>
<td>prob. (n=51)</td>
<td>comp. (n=61)</td>
</tr>
<tr>
<td>-4</td>
<td>33</td>
<td>36</td>
</tr>
<tr>
<td>-3</td>
<td>43</td>
<td>40</td>
</tr>
<tr>
<td>-2</td>
<td>47</td>
<td>39</td>
</tr>
<tr>
<td>-1</td>
<td>69</td>
<td>42</td>
</tr>
<tr>
<td>DUO</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>1</td>
<td>29</td>
<td>47</td>
</tr>
<tr>
<td>2</td>
<td>37</td>
<td>39</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
<td>39</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>29</td>
</tr>
<tr>
<td>5</td>
<td>33</td>
<td>32</td>
</tr>
<tr>
<td>6</td>
<td>33</td>
<td>27</td>
</tr>
</tbody>
</table>

### Table 3.5  Community and custodial sentences after the Demonstration Unit, cumulative

<table>
<thead>
<tr>
<th>years after Demo Unit Offence</th>
<th>average no. of community and custodial sentences per offender</th>
<th>average number of cust. sentences per offender</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>burglary TDA prob. (n=51) comp. (n=61) prob. (n=69) comp. (n=42)</td>
<td>burglary TDA prob. (n=51) comp. (n=61) prob. (n=69) comp. (n=42)</td>
</tr>
<tr>
<td>0</td>
<td>0.00 0.00 0.00 0.00</td>
<td>0.00 0.00 0.00 0.00</td>
</tr>
<tr>
<td>1</td>
<td>0.45 0.71 0.79 0.56</td>
<td>0.25 0.42 0.28 0.25</td>
</tr>
<tr>
<td>2</td>
<td>0.98 1.40 1.26 0.98</td>
<td>0.55 0.73 0.45 0.42</td>
</tr>
<tr>
<td>3</td>
<td>1.61 1.98 1.68 1.58</td>
<td>0.86 1.15 0.67 0.83</td>
</tr>
<tr>
<td>4</td>
<td>2.04 2.40 2.07 2.22</td>
<td>1.16 1.35 0.80 1.17</td>
</tr>
<tr>
<td>5</td>
<td>2.49 2.87 2.33 2.52</td>
<td>1.39 1.60 0.97 1.44</td>
</tr>
<tr>
<td>6</td>
<td>3.03 3.23 2.55 2.94</td>
<td>1.73 1.79 1.09 1.58</td>
</tr>
<tr>
<td>8</td>
<td>3.71 3.82 3.04 3.75</td>
<td>2.00 2.08 1.26 1.96</td>
</tr>
<tr>
<td>10</td>
<td>4.27 4.23 3.36 4.19</td>
<td>2.24 2.23 1.33 2.13</td>
</tr>
</tbody>
</table>

109
Interpretation of trends in offending.

For burglary offenders the offending careers of the probation and comparison groups are very similar. The point where they diverge is that where the probation group attended the Demonstration Unit. In the year after this point burglary offenders on probation show a significant reduction in reconvictions, compared to those in the comparison group. Burglary offenders sentenced to probation offend slightly more before attending the Demonstration Unit, less immediately after, and about the same amount in the long term as those who do not attend. This is strong evidence of a treatment effect brought about by the work of the Demonstration Unit.

The reverse is true for TDA offenders. Those sentenced to probation offend less before attending the Demonstration Unit, more immediately after, and less in the long term than those who do not attend. It is very unlikely the lower offending of TDA offenders in the probation group is the result of a delayed treatment effect, it is hard to conceive of a treatment that would only start to take effect over four years after it was administered. Furthermore the TDA offenders who received probation orders were generally lower risk and offending at a lower rate than the comparison group before the Demonstration Unit. It seems most probable they returned to the rate of offending to be expected of these offenders, following a temporary increase around the time they attended the Demonstration Unit. Whilst it is probably too much to say the Demonstration Unit increased the offending of TDA offenders; it seems clear the programme for burglars was effective in reducing offending, the programme for TDA offenders was not.

Offence specific effects

Offenders targeted by the Demonstration Unit were defined as burglary or TDA offenders on the basis of their most recent offence, and the content of supervision focussed on that type of offence. However there is little evidence of offence
specialisation prior to these offenders being targeted by the Demonstration Unit. The previous offending careers of the burglary offenders contained as much TDA as those of the TDA offenders (on average 1.2 TDA offences each, see Table 3.4). Similarly the TDA offenders had more burglary than TDA in their previous records (on average 1.4 burglaries each compared to 1.2 TDAs).

An obvious question is whether there is evidence of the Demonstration Unit having any special impact on the offences of burglary and TDA. Table 3.6 shows the proportion of offences of the TDA and burglary offenders made up of burglary or TDA, and shows the change in this proportion before and after these offenders were targeted by the Demonstration Unit. For example, of TDA offenders who received probation orders in the year before they attended the Unit, 41 percent of convictions were for TDA related offences. In the year immediately after attending only 24 percent of their reconvictions were for TDA related offences, represented in Table 3.6 as a change of -17 percent in TDA as a proportion of all offending. In isolation, this might appear evidence for an offence specific effect. However a similar trend is apparent for the comparison group. Offenders in this group show a 21 percent reduction in TDA as a proportion of all offending despite the absence of any offence specific intervention. There is no evidence of any offence specific effect taking place as a result of intervention of the Demonstration Unit.

<table>
<thead>
<tr>
<th></th>
<th>1 year before DU offence</th>
<th>1 year after DU offence</th>
<th>Change in offending</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>offs.</td>
<td>TDAs</td>
<td>% TDA</td>
</tr>
<tr>
<td>TDA prob</td>
<td>32</td>
<td>13</td>
<td>41</td>
</tr>
<tr>
<td>TDA comp</td>
<td>23</td>
<td>9</td>
<td>39</td>
</tr>
<tr>
<td>TDA&lt;22 prob</td>
<td>30</td>
<td>11</td>
<td>37</td>
</tr>
<tr>
<td>TDA&lt;22 cmp</td>
<td>16</td>
<td>6</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>offs.</td>
<td>burgs.</td>
<td>% burg.</td>
</tr>
<tr>
<td>burg prob</td>
<td>35</td>
<td>17</td>
<td>49</td>
</tr>
<tr>
<td>burg comp</td>
<td>29</td>
<td>7</td>
<td>27</td>
</tr>
</tbody>
</table>

111
Burglary actually increases as a proportion of offending after the Demonstration Unit, from 49 to 60 percent of all offending (see Table 3.6). However again a similar trend occurs with the comparison group.

Looking at offending in the longer term (see Table 3.7) there is no evidence of any offence specific treatment effects. For burglary offenders reconvicted of burglary annually, Table 3.7 shows there are no clear differences between the probation and comparison groups. This is the case until 8 years after targeting for the Demonstration Unit, when the proportion of the comparison group convicted per year of burglary begins to diminish.

For TDA offenders the amount of TDA committed rapidly tails off, with only a handful of convictions for TDA each year. The tailing off occurs a few years sooner for the probation group, but with such a small amount of offending there is no point in looking too hard for trends. For example a relative peak at 6 years for TDA offending by the comparison group is the result of 4 TDA convictions which occurred over 2 years.

<table>
<thead>
<tr>
<th>years after Demonstration Unit offence</th>
<th>burglary offender, pet reconvicted of burglary</th>
<th>TDA offenders, pet reconvicted of TDA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>prob. ((n=51))</td>
<td>comp. ((n=61))</td>
</tr>
<tr>
<td>1</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>3</td>
</tr>
</tbody>
</table>

112
Did the Demonstration Unit programmes reduce reoffending?

The obvious conclusion of this research is that work with burglars, which seeks to challenge offending and teach skills whereby offending may be avoided, reduces the rate at which they are reconvicted for at least a year. However work with TDA offenders, which seeks to legitimatize their offending, is associated with an increase in reconviction rates.

The 38 percent lower rate of offending by the burglary offenders in the first year after attending the Demonstration Unit, whilst large is in the mainstream of what can be expected of projects which aim to combat offending. For example in his analysis of 397 projects for juvenile delinquents Lipsey (1995) reports that the best treatment types, when delivered in 'low dose' and monitored by a researcher involved in delivering the project (as were the Demonstration Unit interventions), show an average 23 percent change from a 50 percent recidivism control. This would be equivalent to a 46 percent reduction in offending. The findings for TDA offenders are disappointing in this context, but might be explained by more recent theories as to what works in reducing offending.

What does the Demonstration Unit tell us about offence specificity?

The courses for burglars and TDA offenders were predicated on an idea of offence specificity: that those who commit burglary or TDA related offences specialise in this type of crime. Harraway (1985) describes the typical TDA offender as "... familiar to probation officers as the 'TDA merchant' or 'car freak', whose criminal activities usually, if not exclusively, centre on cars..." (page 55). The existence of this type of offender, and of similar offenders specialising in burglary and by extension the need to "treat" these types of offending, was the main reason for setting up the Demonstration Unit in the form that it took. However whilst these
types of offender may seem familiar to practitioners, research exists which calls into question the notion of offence specificity.

Large-scale research into criminal careers shows specialisation in a particular type of offence occurs only at a low level, other than with sex offenders and those committing fraud over a long period, (see for example Farrington, 1992; Stander et al 1989). From a review of 33 small-scale projects Klein (1984), concludes

"The evidence is very supportive of a general delinquency factor and extremely weak for offence specialisation..." (page 191).

However this latter point has been criticised as being an artefact of using too few categories to represent delinquent behaviour (Loeber and Waller, 1988). This matter will be returned to in the next chapter, which evaluates another project for TDA offenders: the Ilderton Motor Project.

It would be too much to conclude from the findings of the Demonstration Unit that work focussed on a particular type of offending is or is not effective. This is because most of the offenders who attended the Demonstration Unit were not specialists in a particular type of offending. Prior to the offence which led to their being targeted by the Demonstration Unit, the TDA offenders had committed more burglary than TDA, and the burglary offenders had committed as much TDA as the TDA offenders. All that can be concluded from the findings of the Demonstration Unit is that specialist burglary and TDA offenders are relatively rare, and likely to be a minority of those who commit these types of offences.

Conclusion

The main message that emerges from this first evaluation case-study is that some offender-interventions work, but others do not. This study clearly shows it is only
possible to determine how effectively an intervention reduces offending by using an objective measure of reoffending, and using an appropriate comparator. Reconvictions are an important bottom line for programmes that aim to reduce offending. Mair (1992) argues that a range of indicators are needed for evaluating the success of community programmes. But for programmes which aim to reduce offending, reconviction will remain an outcome indicator of overarching importance. In the case of the Demonstration Unit, it is only when reconvictions enter the equation that the success of the programme for burglary offenders becomes apparent.

Findings based only on those who attended a project can be confusing. The Demonstration Unit team concluded their project had been a success (see Harraway et al, 1986), which in terms of its objectives it had been even with the TDA offenders. The Demonstration Unit expanded the use of probation amongst its target groups of offenders, and held them in the community without markedly worse consequences in terms of reconviction than for standard probation clients. Only when a comparison group is used, and reconvictions are reviewed over some years, is it apparent that work with burglars succeeded in temporarily but significantly reducing offending, whilst that with TDA offenders did not.

Of the two projects which constituted the Demonstration Unit, the TDA project was felt to have been the more successful. This research shows that a project may generate enthusiasm and enjoyment amongst those who attend, and be experienced as a successful project by those running it, but this does not mean that the project will necessarily reduce offending. Equally this research shows that just because those running a project are less than confident about what they are doing, this does not necessarily mean that offending will not be reduced.

Reliable measures of intermediate outcomes might have provided more rapid feedback as to the success or otherwise of a project. Alarm bells should have been set ringing by the fact only 12 of the 54 TDA offenders who satisfactorily completed the scheme were known to have passed a driving test, the original aim
of the scheme.

Finally if programmes are intended for offenders who specialise in one type of crime (other than sex offenders), offenders entering the programme should be defined as specialising in that type of crime by more than just their most recent offence. This point is reinforced by the next chapter, which also consists of an evaluation case-study of a programme for offenders who take cars, but who show a significant interest in cars, and who have an extensive history of committing car crime.
Chapter 4

Offence specific work with offenders who take cars: Ilderton Motor Project

This chapter, in common with the previous, focuses on offence-specific work and evaluates an intervention for young offenders who take cars, albeit in this case a more intensive and practically focussed project, which works with offenders who are more clearly specialised in car crime.

Ilderton Motor Project, which is based in South London, is one of the UK's best known probation motor projects. The first part of this chapter reviews research into car crime and the criminal careers of offenders who commit these types of offence, to determine whether there is evidence to confirm whether the type of specialist offenders these programmes are aimed at actually exist. Interventions such as motor projects are predicated in a notion of offence specificity: that specialised offenders exist who commit mainly one type of crime, in this case car-crime. This idea of offence-specialised offenders is explored in some depth. The second part of the chapter reviews the impact of Ilderton Motor Project on the offending of a group of probationers who attended the project, comparing this with the offending of a matched comparison group. The project's impact on offending generally, and on car crime in particular is reviewed.

Ilderton Motor Project, actually predates the Demonstration Unit programme for TDA offenders which was described in the previous chapter, having been set up in 1974. However the period over which offenders in this study attended Ilderton was the end of the 1980s and early 90s, after the Demonstration Unit project had concluded. Since this time, (as shown in Chapter 2, Tables 2 and 3) programmes for motor offenders have become widespread in probation practice. The Index of Probation Projects (NPRIE, 1994) lists 59 of these projects, in a survey Hedderman and Sugg (1997) discovered 37 car-crime projects, and in a census Martin and

117

Motor projects are the probation service's main response to rising car crime (see Whitehead, 1993, for a discussion of policy in this area). From the mid-1980s they have been the main way in which the probation service has been was active in crime prevention (see Laycock and Pease 1985). This is likely to have become still more the case by the mid-90s. In their census, half the motor projects catalogued by Martin and Webb (1994), had come into existence since 1990. However as yet little evidence exists as to the effectiveness of motor projects in in reducing motor vehicle-related offending.

Responses to car crime

In recent years there has been an increase in all types of car crime. In 1992 over half a million thefts of cars in England and Wales were recorded by the police, 80 percent more than in 1982 (see Mayhew, 1992, Webb and Laycock, 1992, and NACRO, 1994, for reviews of trends in car crime). Responses to this increase have been more in crime prevention than offender treatment, and this is reflected in two strands of research. The first looks at cars and the places from which they are stolen. This covers the practical issues of the makes of car most likely to be taken (Houghton, 1992); and the local impact on car crime of innovations such as vehicle watch schemes (Honess et al, 1993), or video-monitoring car-parks (Webb and Laycock, 1992).

A second strand of research is based on offenders’ experiences, looking at who commits car crime and why (see for example Light et al, 1993, McGillivary, 1993; Spencer, 1992; Briggs, 1991). Common findings of these studies reveal typical offenders who takes cars to be male, becoming involved in car crime in their early to mid-teens. Their education is characterised by high levels of truancy and low levels of attainment. They become involved in car crime initially out of a desire for
excitement, backed by peer pressure. From this follows enhanced status amongst peers. Those more advanced in car crime careers may become motivated by financial rewards provided by markets in car parts and radio-cassettes.

There is scant evidence as to which ways of working with offenders or potential offenders are effective in bringing about crime reduction. Harraway et al (1985 and see chapter three), reviewing a scheme for offenders who take cars, reports a two-year reconviction rate of 39 percent. Davies (1993) reports a two-year reconviction rate of 72 percent for similar offenders starting motor-offender programmes, and of 54 percent for those who complete these programmes. In both cases these reconviction rates were reviewed against reconviction rates for other groups of young offenders (for example released from custody), and were interpreted as favourable. However neither study incorporated a control or matched comparison group, so it is not possible to know with any degree of confidence likely reconviction rates had offenders not attended each scheme.

Motor projects and crime reduction

Motor projects seek to reduce crime in two ways, by preventative work with young people at risk of offending; and by what may be construed as treatment: work with young people involved with the criminal justice system, and referred to a project because of their offending. The idea behind crime prevention is that a relatively small number of individuals may be responsible for a large amount of motor crime. Whilst occupied at a project they are not offending. Another aim is to divert young people at risk of involvement in auto-crime into constructive activities before their offending escalates. Whilst Ilderton undertakes such preventative work, the emphasis with those on probation is on helping those already committing car crime to change from a way of life bound up with offending.

Ilderton has provided the model for many motor projects. The work of Ilderton Motor Project has been described in detail by Pearce and Thornton (1980), and
more recently by Martin and Webb (1994). It gives young people who have become involved in car crime a chance to pursue an enthusiasm for cars in a constructive and responsible fashion. The aim is this should take place in a disciplined environment which challenges their offending. Ilderton is effectively a large garage and workshop where young people can work on cars and learn about how they function and are maintained. Most of the work is done in the project where the young people may work on their own cars, or restore old cars for use in 'banger' racing. Around 20 people may attend at any one time, being required to attend at least 8 times per month. Offenders supervised by the probation service may attend the project as the condition of a probation order or on a voluntary basis, though most who attend Ilderton are not on probation. Martin and Webster (1994) estimate on average 15 probationers a year attend Ilderton. There is no prescribed 'course', attendance may be open-ended. Those included in this research who attended Ilderton were in contact with the project for on average 14 months.

**Car crime and offence specialisation**

The work of motor projects is based more or less explicitly on an idea of offence specificity: that those who take cars specialise in this type of crime. Harraway (1985) describes these offenders as

'... familiar to probation officers as the ... "car freak", whose criminal activities usually, if not exclusively, centre on cars...' (page 55).

The existence of this type of offender and need to 'treat' this type of offending is a justification for the existence of motor projects. But though such offenders may seem familiar to practitioners, as was described in the previous chapter, research into criminal careers calls into question the notion of offence specificity.

Kempf (1986) reviewing research into criminal careers, and analysing the 27,160 strong Philadelphia birth cohort, reports 'the specialist career pattern is uncommon'
(page 189). More recent and UK-based criminal careers research shows for most offence types, exceptions being sex offences and fraud, specialisation is barely detectable (see for example Farrington, 1992a; Stander et al 1989).

Motoring offences show a low degree of specialisation. For example a study of 43,400 offenders showed those previously convicted of motoring offences on reconviction are only very slightly more likely than others to be reconvicted for offences in the same category (22 percent, Home Office, 1985, reported in Tarling 1993, pp 125). In this case motoring offences include all offences of taking cars, but also all other driving related offences. A similar point emerges from research into offending amongst juveniles. Klein (1984), reviewing 33 studies of delinquency amongst juveniles concludes

'The evidence is very supportive of a general delinquency factor and extremely weak for offence specialisation...' (page 191).

Though the latter point has been criticised on methodological grounds by as being an artefact of using too few categories to represent delinquent behaviour (Loeber and Waller, 1988).

Studies based on interviews with offenders produce rather different findings. For example Nee (1993), interviewing offenders who took cars reports that about half regarded themselves as specialists in this type of crime, with those more interested in cars reporting longer offending careers in car crime.

Evaluating the effectiveness of Ilderton means addressing the matter of offence specialisation. Do specialist motor-vehicle offenders attend Ilderton? And are there specific effects on car crime compared to other types of offending? If specialist car offenders do not exist in significant numbers, the thinking behind motor projects is flawed. However if Ilderton is attended by 'car freaks', who subsequently commit less of this sort of crime, this would support the idea of motor projects generally.
Making sense of Ilderton's impact on reconviction

This research was carried out to assess whether for probationers attending Ilderton Motor Project is associated with their committing less crime, and in particular less car crime.

Assessing the impact on offending of a project means knowing how much crime those who attended would have committed had they not done so. This is commonly done using a quasi-experimental research design, whereby the offending of the group attending a project is compared with that of a comparison group. For Ilderton this would be a group of offenders similar in as many respects as possible to those who went to Ilderton, but who did not attend. The need is to show the two groups are sufficiently similar at the point the Ilderton group started at Ilderton, for the subsequent offending of the comparison group to indicate how the Ilderton group might have acted had they not attended.

This comparison group method has flaws. Where groups of attenders and non-attenders are not selected randomly, changes in offending may be due to differences between the groups unrelated to attending Ilderton. Pawson and Tilley (1994) have criticised pseudo-experimental evaluation methods for saying nothing about the processes whereby changes might have been achieved. However this method remains the best available for providing a rapid indication of whether Ilderton is achieving its aims. If it is this might justify research into processes, or the use of a "true" experimental design.

Gathering information about the offenders and their offending

Information was obtained from Ilderton's records about all offenders supervised by ILPS who attended Ilderton between January 1987 and March 1994. Information about offending, including all arrests, charges and convictions, was supplied via the Metropolitan Police Youth Affairs Branch.
In this study the term 'offence' refers to an occasion on which one or more offences took place resulting in an arrest. Offences where the defendant was later tried and found not guilty are excluded; but offences where proceedings were discontinued are included. Where a suspect was apprehended some time after the possible commission of an offence the date of the offence is used. Thus in this study the indicator of offending is not conviction or sentencing occasion, as is common in UK studies. Rather it is the actions leading to an arrest recorded by the police. For evaluation purposes recorded arrests give a better estimate of offending activity than convictions, omitting 'noise' introduced into the data by prosecution and sentencing practice. For example one incident occasioning an arrest can result in a number of charges and eventual convictions for a number of offences. Equally several incidents occasioning arrests over a period, can be dealt with in one court appearance. So neither recorded convictions nor sentencing occasions give a very close measure of actual offending (see Tarling 1993, pages 39-40 for a discussion of this problem).

Information from police records also prevents what Lloyd et al (1994) term 'pseudo reconvictions' from clouding the data. Pseudo reconvictions occur when offenders are convicted during an existing sentence for offences committed before the sentence started. These can not be regarded as an outcome of the sentence. All offence dates used in the current study can be regarded with some confidence as being the actual dates when an offence was committed.

Information about the offenders and their offending was used to calculate using a reconviction predictor (the PSR Predictor, Wilkinson, 1994) the chance of each offender being reconvicted of any offence within two years of starting at Ilderton.

**Interpreting police records**

Similar behaviour on the part of an offender may be interpreted in different ways leading to differing charges. A person apprehended trying to break into a car may
have been trying to take the car or to steal goods from it, and could be charged with attempted Taking and Driving Away (TDA) or with theft from a motor vehicle. Similarly an offender apprehended driving, or being driven in, a stolen vehicle could be charged with TDA, taking without consent (TWOC), or with being allowed to be carried in a motor vehicle (ATBC). For this reason in this study motor vehicle crimes are grouped into two categories. The first is 'motor-vehicle related offences'. This includes all car crime: all offences relating to theft from, taking of, or driving motor vehicles. The second category is a subset of the first: 'taking cars'. This includes all offences that relate to taking cars, overwhelmingly TDA and some TWOC.

Matching the offending careers of the Ilderton and comparison groups: assigning reference offences

A good comparison group would contain offenders with similar offending histories to the Ilderton group up to the point when the latter started at Ilderton. Those in the comparison group should have committed an offence which could have occasioned their attendance at Ilderton, at an equivalent point in their offending careers to the offence which actually resulted in the Ilderton group's attending Ilderton. This allows offending before this point to be compared to the offending of the Ilderton group before they attended; offending after this point can be compared with the offending of the Ilderton group after they attended. The offence which constitutes the point at which the offending careers of both groups can be compared is termed the 'reference offence'.

For the Ilderton group the reference offence is the last vehicle related offence (including thefts from as well as of cars) committed before starting at Ilderton. For the comparison group the reference offence is a similar offence committed at an equivalent point in their offending career. For most offenders in the comparison group two reference offences are assigned. The first is the offence which led to the preparation of the pre-sentence report present in ILPS' monitoring system.
Information collected at this point provided the basis for selecting the comparison group, matching it to the Ilderton group. Thus at reference offence 1 the comparison group was selected 'blind', no information being available about subsequent reconvictions.

At reference offence 1, the comparison group is on average a year younger than the Ilderton group (see Table 4.1). This difference could be significant in the context of a motor-vehicle offending career, these offenders being relatively young. The average age of offenders convicted of TDA for whom Inner London Probation Service prepares pre-sentence reports is 22.5 years, 7 years younger than the average for offenders on whom pre-sentence reports are prepared¹.

Because car crime is committed by younger offenders, the older offenders in the Ilderton group could be closer to 'growing out' of committing motor vehicle crime. For this reason a second reference offence was assigned for the comparison group, so as to produce a closer fit between the ages and offending careers of those in the reference and comparison groups. This was done using the detailed information about offending careers provided by the police. For 25 members of the comparison group it was possible to select a second reference offence. This was a motor-vehicle related offence committed about a year later than the first reference offence and which could also have resulted in the comparison group member attending Ilderton. This reduced the age difference to under 5 months, a period unlikely to be significant in the context of an offending career.

Are the Ilderton and comparison groups similar?

Table 4.1 shows the sentences which were received for the reference offences. These are a mixture of fines, probation orders, community service orders and short custodial sentences. These sentences are at a similar level of seriousness to the probation orders to which ILPS' clients who attend Ilderton are subject. This suggests the reference offences are at a similar level of seriousness for Ilderton and

125
comparison groups.

At reference offence 1 those in the comparison group are on average a year younger than offenders who attended Ilderton. They have an offending career on average 8 months shorter than that of the Ilderton group. The comparison group have committed fewer offences on average (5.4 compared to 6.6) and correspondingly fewer vehicle related offences (2.5 compared to 4.9) and offences of taking cars (1.2 compared to 1.8). These measures suggest that at reference offence 1 those in the comparison group might be at an earlier point in their offending careers than offenders in the Ilderton group. The offending careers seem to fit more closely at reference offence 2. Offenders in the comparison group are on average only five months younger than those in the Ilderton group; they have committed slightly more offences (7.2 compared to 6.6), though still slightly fewer vehicle offences (3.6 compared to 4.9) and offences of taking cars (1.6 compared to 1.8).

On average 74 percent of previous offences committed by the Ilderton group are vehicle offences, and 38 percent are offences of taking cars. Thus these seem to be specialised offenders. For the comparison group these figures are 46 percent and 29 percent at reference offence 1; 50 percent and 30 percent at reference offence 2. This too seems a relatively high degree of specialisation in car crime, especially given the comparison group was selected on the basis of one offence without information about their previous offending.

The two groups are similar in terms of the proportion previously sentenced to custody. 26 percent of the Ilderton group had received custodial sentences in the 2 years immediately prior to the reference offence, compared to 25 percent of the comparison group at reference offence 1, and 38 percent at reference offence 2.

It might seem a reasonable consequence of the generally closer fit of offending careers at reference offence 2 only to carry out further analysis of the comparison group at this point. However at reference offence 1 the comparison group was
selected 'blind'. Information about previous offending only was available from ILPS' pre-sentence report monitoring database, there could be no knowledge of this group's subsequent offending. There may be a less close match of offending careers but there is no possibility of researcher bias. Reference offence 2 was selected in the light of the comparison group's entire offending career to June 1994, when the research was carried out. Whilst reference offence 2 was selected strictly in accordance with criteria given above, the availability of information about the entire offending career eliminates any prospective element. For this reason findings are presented for both reference offences. Reference offence 2 provides a more exacting test of whether Ilderton reduces offending. By definition most of the comparison group (the 25 for whom a second reference offence could be identified) have committed one more vehicle offence at reference offence 2 compared to at reference offence 1, and accordingly will have one fewer motor vehicle offence included in the follow-up period from this offence.

The Ilderton group is not a cohort, so the start and end dates of follow-up periods are different for different offenders. Table 4.1 shows the average length of follow-up for each of the groups: 47 months for the Ilderton group, 48 for the comparison group at reference offence 1, and 41 months at reference offence 2. There are wide variations around these averages, follow-up periods ranged between 10 and 106 months for the Ilderton group, and between 14 and 97 months for the comparison group. This means that when follow-up periods of over 1 year are assigned, a substantial number of offenders are excluded from analysis (see Table 4.2).

**Analysing information about offending**

For each offender the reference offence was taken as time point zero. Offences were aggregated into six month periods before and after each reference offence. The follow-up period for both groups extended to 1 June 1994, the most recent date at the time of the research to which information could be obtained. The follow-up period was calculated as the time between the reference offence and this date. Where there were two reference offences this process was repeated for each
### Table 4.1. Comparing the Ilderton and comparison groups

<table>
<thead>
<tr>
<th>comp. grp at</th>
<th>Ilderton group</th>
<th>comp. group at</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=35)</td>
<td>ref. off 1 (n=40)</td>
</tr>
<tr>
<td><strong>percent...</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>black/other race</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>employed/education</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td><strong>average...</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>time from 1st off. to ref. off.</td>
<td>34.7 months</td>
<td>26.3 months</td>
</tr>
<tr>
<td>length of contact with Ilderton</td>
<td>14 months</td>
<td>N/A</td>
</tr>
<tr>
<td>length of follow-up period</td>
<td>46.6 months</td>
<td>48 months</td>
</tr>
<tr>
<td>age at ref. off.</td>
<td>19.4 years</td>
<td>18.4 years</td>
</tr>
<tr>
<td>percent likelihood reconvict. in 2 yrs</td>
<td>75</td>
<td>73</td>
</tr>
<tr>
<td>no. offxs. before ref. off.</td>
<td>6.6</td>
<td>5.4</td>
</tr>
<tr>
<td>no. m/v related offences before ref. off.*</td>
<td>4.9</td>
<td>2.5</td>
</tr>
<tr>
<td>percent offences m/v related</td>
<td>74</td>
<td>46</td>
</tr>
<tr>
<td>no. off's taking cars (in 2 yrs before ref. off).*</td>
<td>1.8</td>
<td>1.2</td>
</tr>
<tr>
<td>percent offences which involve taking cars</td>
<td>38</td>
<td>29</td>
</tr>
<tr>
<td>no. custodials in the 2 years before the ref. off.</td>
<td>0.46</td>
<td>0.38</td>
</tr>
<tr>
<td>pct. sent. to custody in 2 yrs before ref. off.</td>
<td>25</td>
<td>25</td>
</tr>
</tbody>
</table>

**sentence: percent...**

- financial penalty / bind over / discharge | - | 35 | 38 |
- Probation / CS<=120 hrs/attendance. centre | - | 40 | 25 |
- Probation plus Ilderton Motor Project | 100 | - | - |
- PO+cond./CS>120 hrs/comb. order | - | 5 | 8 |
- Custody<=6 mths/suspended sent | - | 20 | 25 |
- Custody>6 mths | - | - | 5 |

* Motor vehicle (m/v) related offences include all offences of taking cars (see below), plus all thefts and attempted thefts from cars.

** Offences of taking cars include Taking and Driving Away (TDA), Attempted TDA, Taking Without Consent (TWOC), Allowing to be Carried in a Motor Vehicle (ATBC)

### Table 4.2. Percent offending and average number of all offences since reference offence

<table>
<thead>
<tr>
<th>time after ref. offence</th>
<th>Ilderton group</th>
<th>comparison group at reference offence 1</th>
<th>comparison group at reference offence 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pct. offend.</td>
<td>average no. offences.</td>
<td>pct. offend.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>pct. offend.</td>
</tr>
<tr>
<td>1 year</td>
<td>53</td>
<td>1.2 (n=34)</td>
<td>80 *</td>
</tr>
<tr>
<td>2 years</td>
<td>65</td>
<td>2.3 (n=26)</td>
<td>91 *</td>
</tr>
<tr>
<td>3 years</td>
<td>62</td>
<td>2.5 (n=21)</td>
<td>100 **</td>
</tr>
</tbody>
</table>

* More offending than the Ilderton Group: statistically significant at the \(p<0.05\) level (matched groups t-test, 2 tailed.)

** More offending than the Ilderton Group: statistically significant at the \(p<0.01\) level (matched groups t-test, 2 tailed.)
reference offence.

To allow evaluation of any specific impact on motor-vehicle crime, offences were defined as described above, either as "all offences", "motor-vehicle related offences", or "offences of taking cars". Each category is a subset of the former.

Reoffending outcomes

The groups seem sufficiently similar at the point the Ilderton group commenced at Ilderton, for the subsequent offending of the comparison group to indicate how the Ilderton group might have offended had they not attended Ilderton. Table 4.2 shows the proportion who offend, and average number of offences committed one, two and three years after the reference offence.

Reoffending rates are high. Almost two thirds of the Ilderton group reoffend within two years, but this compares to over 90 per cent of the comparison group reoffending within 2 years of either reference offence. Looking at the comparison group from reference offence 1, the level of offending is considerably higher than for the Ilderton group, both in the proportion who reoffend and in the average number of offences committed by each offender. For example after a year 53 percent of the Ilderton group had reoffended, committing on average 1.2 offences each. This compares to 80 percent of the comparison group reoffending within a year of reference offence 1, committing an average of 2.0 offences each.

At reference offence 2 less reoffending should be expected from the comparison group, given that for most of this group one of the offences recorded as 'reoffending' from the point of view of reference offence 1 will have been classed as 'previous offending'. However offenders in the comparison group still reoffend more than those who went to Ilderton. This applies for all time periods, for all offences, and for motor vehicle related offences. For example 91 percent of the
comparison group reoffend within 2 years of reference offence 2, compared to 65 percent of the Ilderton group; 100 percent of the comparison group reoffend within 3 years compared to 62 percent of those members of the Ilderton group for whom a three year follow-up period existed. These findings are statistically significant at the \( p<0.05 \) and \( p<0.01 \) levels respectively in a two-tailed t-test (see Tables 4.2 and 4.3).

It is clear the Ilderton group reoffends significantly less than the comparison group, though the magnitude of the difference in reoffending varies depending on which reference offence is used as a measuring point.

**Offence specific effects**

The Ilderton group also commits less motor-vehicle related offending. Within a year of reference offence 1, 29 percent of the Ilderton group had been arrested for a motor vehicle related offence, compared to 53 percent of the comparison group (see Table 4.3). All differences between the groups are in the direction of less car crime from those who attended Ilderton.

The offending of both Ilderton and comparison groups, prior to Ilderton or the reference offence, shows a high degree of specialisation. Almost three quarters of the Ilderton and approaching a half of the comparison group's previous offences were motor vehicle related (see Tables 4.1 and 4.3.) These continue to constitute a high proportion of offending over the follow-up period for both groups (see Table 3). However in the two years following the offence which led to them attending Ilderton only 43 percent of the Ilderton group’s offences were motor vehicle related, compared to 74 percent of their earlier offences. For the comparison group at reference offence 1, the proportion of offences which were motor vehicle related remains the same before and after the reference offence, at 46 percent. For the comparison group at reference offence 2, the proportion of vehicle-related offences falls somewhat from 50 to 37 percent of all offences, but this apparent
Table 4.3. Motor-vehicle related offending: percent offending and average number of motor vehicle offences committed since reference offence

<table>
<thead>
<tr>
<th></th>
<th>Ilderton group</th>
<th>comparison group</th>
<th>comparison group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% m-v ave. no. % of off's</td>
<td>m-v ave. no. % of off's</td>
<td>% m-v ave. no. % offs</td>
</tr>
<tr>
<td>off'd offs. m-v related off'd offs. m-v related off'd offs. m-v related</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>before ref off.</td>
<td>100 4.9 74 (n=35)</td>
<td>100 2.5 46 (n=40)</td>
<td>100 3.6 50 (n=40)</td>
</tr>
<tr>
<td>1 year after ref. off.</td>
<td>29 0.6 50 (n=34)</td>
<td>53 * 1.0 50 (n=40)</td>
<td>48 0.5 36 (n=38)</td>
</tr>
<tr>
<td>2 years after ref. off.</td>
<td>42 1.0 43 (n=26)</td>
<td>69 1.7 46 (n=35)</td>
<td>50 1.1 37 (n=32)</td>
</tr>
<tr>
<td>3 years after ref. off.</td>
<td>43 1.2 48 (n=21)</td>
<td>80 * 2.4 53 (n=30)</td>
<td>71 # 1.8</td>
</tr>
</tbody>
</table>

* Statistically significant at the p<0.05 level in a two-tailed matched groups t-test.
# Statistically significant, p<0.05 in a one-tailed t-test. A one-tailed test is justified because the hypothesis tested is that offenders who attend Ilderton will reoffend less. “m-v off’d” refers to any motor vehicle related offence; % of off’s m-v related refers to the proportion of all offences committed which involved motor vehicles.

Table 4.4. Offences of taking cars: percent offending and average number of offences of taking cars since the reference offence.

<table>
<thead>
<tr>
<th></th>
<th>Ilderton group</th>
<th>comparison group</th>
<th>comparison group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% took ave. no % of off's</td>
<td>% took ave. no % of off's</td>
<td>% took ave. no % offs</td>
</tr>
<tr>
<td>cars offs. taking cars cars offs. taking cars cars offs. taking cars</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>before ref off.</td>
<td>74 1.8 38 (n=35)</td>
<td>58 1.2 29 (n=40)</td>
<td>73 1.6 30 (n=40)</td>
</tr>
<tr>
<td>1 yr after ref off.</td>
<td>18 0.2 17 (n=34)</td>
<td>33 0.6* 30 (n=40)</td>
<td>16 0.2 14 (n=38)</td>
</tr>
<tr>
<td>2 yr after ref off.</td>
<td>27 0.5 22 (n=26)</td>
<td>58 * 1.0# 27 (n=35)</td>
<td>39 0.7 23 (n=32)</td>
</tr>
<tr>
<td>3 yrs after ref off.</td>
<td>32 0.6 24 (n=21)</td>
<td>70 ** 1.6* 31 (n=30)</td>
<td>48 0.9 20 (n=21)</td>
</tr>
</tbody>
</table>

# Statistically significant at the p<0.05 level in a one-tailed matched groups t-test. The justification for a one tailed test is that the hypothesis being tested is the Ilderton Group will commit significantly fewer offences of taking cars.
* Statistically significant at the p<0.05 in a two-tailed matched groups t-test.
** Statistically significant at the p<0.01 level in a two-tailed matched groups t-test.
reduction is partly an artefact of shifting one motor vehicle offence to before the reference offence. This is evidence in favour of an offence specific effect. The lower arrest rate amongst the Ilderton group is disproportionately the result of a decrease in vehicle-related offending.

Looking at offences of taking cars, there are fewer statistically significant findings. This category is a sub-set of motor-vehicle offences, so with fewer offences statistical significance is less easily attained. However most of the differences are in the direction of fewer offences of taking cars by the Ilderton group (see Table 4.4). For example two years after the Ilderton offence, 27 percent of the Ilderton group had committed offences of taking cars. For the comparison group at reference offence 1 this figure was 58 percent (significant at p<0.05 in a two-tailed matched-groups t-test). At reference offence 2 this figure was 39 percent.

The proportion of offending that involves taking cars also falls more for the Ilderton group than for the comparison group (see Table 4.4). In the 2 years before the Ilderton group's reference offence, 38 percent of this group's offences were of taking cars, compared to 22 percent of their offences in the two years after: a 40 percent reduction. For the comparison group the equivalent reductions are smaller: 7 percent at reference offence 1, and 23 percent at reference offence 2.

Age and offending

The reason for allocating a second reference offence was the link between age and car crime, this generally being committed by younger offenders. However the average ages of the Ilderton group (19.4 years), and of the comparison group at reference offence 1 (18.4 years), are lower than the average age of offenders for whom PSRs are prepared following a TDA offence, which is 22.6 years. It may be that at their relatively young age the Ilderton group is not significantly closer than the comparison group to growing out of car crime over the three years of follow-up.

132
Table 4.5. Percentages of offenders arrested for all offences and for motor-vehicle offences since reference offence 1.

<table>
<thead>
<tr>
<th>Time since ref. off.</th>
<th>Ilderton group</th>
<th>Comparison group</th>
<th></th>
<th>TDA offences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>aged under</td>
<td>aged over</td>
<td>aged under</td>
<td>aged over</td>
</tr>
<tr>
<td>1 year</td>
<td>60 (n=15)</td>
<td>47 (n=19)</td>
<td>68 (n=22)</td>
<td>94 (n=18)</td>
</tr>
<tr>
<td>2 years</td>
<td>71 (n=14)</td>
<td>62 (n=13)</td>
<td>86 (n=21)</td>
<td>100 (n=14)</td>
</tr>
<tr>
<td>3 years</td>
<td>70 (n=10)</td>
<td>60 (n=10)</td>
<td>100 (n=18)</td>
<td>100 (n=12)</td>
</tr>
</tbody>
</table>

Table 4.6. Percentage and number sentenced to custody since reference offence (excluding sentencing for reference offence).

<table>
<thead>
<tr>
<th>Time since ref. offence</th>
<th>Ilderton Group</th>
<th>Comparison group at ref. off. 1</th>
<th>Comparison group at ref. off. 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% cust. ave. no. custs.</td>
<td>% cust. ave no. custs</td>
<td>% cust. ave no. custs</td>
</tr>
<tr>
<td>1 year</td>
<td>12 0.1 (n=34)</td>
<td>30 0.6* (n=40)</td>
<td>24 0.3 (n=38)</td>
</tr>
<tr>
<td>2 years</td>
<td>15 0.3 (n=26)</td>
<td>46 ** 0.9* (n=35)</td>
<td>34 0.6 (n=33)</td>
</tr>
<tr>
<td>3 years</td>
<td>10 0.1 (n=21)</td>
<td>57 ** 1.4* (n=30)</td>
<td>52 ** 0.9** (n=21)</td>
</tr>
</tbody>
</table>

* Statistically significant at the p<0.05 level (matched groups t-test, 2 tailed.)

** Statistically significant at the p<0.01 level (matched groups t-test, 2 tailed.)

133
Offenders in each group can be banded according to whether they are younger or older than the median age of the combined group which is 18.5 years. Table 4.5 shows the proportion who reoffended by age and by whether the offender attended Ilderton. Older offenders in the comparison group in general do not reoffend less than younger offenders. Though a slightly smaller proportion of older offenders in the comparison group commit TDA related offences, the reverse is the case for all offences. Because of the small numbers at this level of crosstabulation, these observations are not statistically significant. However it does seem age is not a significant factor in the offending of this group, at least over the 3 year period under scrutiny. This finding supports the validity of using reference offence 1 as a point for comparing the offending careers of the two groups.

**Seriousness of reoffending: offences sentenced by way of custody**

A good indication of the seriousness of an offence is the way it is sentenced, the more serious generally being sentenced by way of custody. Table 4.1 shows 26 percent of the Ilderton group previously had been sentenced to custody compared to 25 percent of the comparison group at reference offence 1, 38 percent at reference offence 2. Table 4.6 shows the average number and proportion of offenders sentenced to custody since the reference offence, not including the sentence for the reference offence itself. Those who attend Ilderton are significantly less likely to be sentenced to custody whichever reference point is taken. After two years 15 percent of those who attended Ilderton had been sentenced to custody, compared to 34 percent or 46 percent for the comparison group. After 3 years the difference becomes still more pronounced: 10 percent of the Ilderton group followed for 3 years had been sentenced to custody, compared to over half those in the comparison group (57 percent measuring from reference offence 1, 52 percent from reference offence 2, see Table 4.6.)

Previous convictions are a factor in sentencing, but because the Ilderton and comparison groups are in terms of offending careers this is unlikely to explain any
differences in sentencing here. Whilst it is possible attendance at Ilderton might sway a court to impose a lower sentence, far more likely is that the Ilderton group, when they do offend, commit less serious offences.

Conclusions

The main finding of this research is that offenders who attend Ilderton reoffend less than similar offenders who do not, evidence that Ilderton reduces offending. Furthermore those who attend Ilderton commit significantly less motor vehicle crime, including fewer offences of taking cars, than a comparable group of offenders who did not attend Ilderton. The differences in offending between the two groups are minimal before one group attended Ilderton, and are substantial afterward, especially for motor-vehicle crime. This applies whichever reference point is taken, for example 65 percent of those who attend Ilderton reoffend after 2 years compared to 91 percent of the comparison group: a 29 percent reduction. There are even larger reductions in the use of custody, offenders in the Ilderton group being less than half as likely to be sentenced to custody within 2 years. These are substantial effects in the context of projects that aim to reduce offending, though in line with those achieved by the more successful types of treatment (see for example Lipsey, 1995.)

The reduced offending of those who attend Ilderton seems disproportionately to be the result of reduction in motor-vehicle related offending. The degree of specialisation in this type of offending, revealed in this study seems at odds with findings from criminal careers research. This may be because specialist offenders of the type targeted by Ilderton are only a small proportion of offenders. It may also indicate that the broad categories for classifying offence type used in criminal careers research fail to reveal groups of specialist offenders where they do exist. For example in some criminal careers studies offences of taking cars have been grouped with theft (e.g. Philpotts and Lancucki, 1979), in others they have been classified with other motoring offences (Home Office, 1985). In either case
offenders who specialise in taking cars would be concealed.

A weakness of the comparison group methodology used in this study is that it is not possible to know conclusively whether differences in the offending patterns of the two groups are due to intervention, or are due to pre-existing differences. However the positive findings of this study would support looking in more detail at how Ilderton works to reduce offending. There seems little doubt the offenders who attend Ilderton specialise in car crime. Equally there seems little doubt that attending Ilderton is associated with lower re-arrest rates, especially for motor-vehicle offences, and offences of taking cars. This is consistent with Ilderton providing a useful response to car crime.

Ilderton motor project was aimed at a relatively narrow category of specialist car-crime offender. Similarly the Demonstration Unit programmes, described in the previous chapter each focused on offenders convicted of car crime or burglary. However most offenders do not appear to specialise in one type of crime in this way (see for example Farrington, 1992a), and similarly a large proportion of probation programmes are aimed at the generality of offenders. Most offenders supervised by the probation service have committed a range of property crimes, and crimes against the person. This is true of offenders who attend centre-based probation programmes, which specialise in working with offenders who are at high risk of reoffending. The next two chapters describe two such programmes, the first of these is the Sherborne House programme for young offenders, aged between 17 and 21 years.

---

Notes

1. In a sample of 5936 pre-sentence reports prepared for court appearances between March 1993 and April 1994, the average of all offenders was twenty nine years, eight months. The average age of offenders who had committed TDA offences was twenty two years, seven months. The influence of age on TDA offending is explored in a later section.
Chapter 5

Probation centres' impact on offending: an evaluation of the Sherborne House programme for young offenders

This chapter describes an evaluation of the impact on reconviction of Sherborne House, a probation centre for young offenders aged 17 to 20 years. The programmes described in earlier chapters were each aimed at offenders who had committed certain types of offence. However the Sherborne House programme is typical of many probation centre programmes, in that it is aimed at serious and persistent offenders, rather than at offenders convicted of a particular category of crime. Offenders are deemed suitable for Sherborne House because of the seriousness of their offending, and their high risk of reoffending. The Sherborne House programme is typical of many centre-based programmes aimed at the most serious offenders likely to be sentenced in the community.

The influence of probation centres on probation practice is disproportionate to the number of offenders who actually pass through these centres. This is because probation centres are often where new ideas for working with offenders are tested, and where probation practice is at its most visible. This chapter starts with a brief review of the place of probation centres in probation practice. There then follows a description of the Sherborne House programme, as it existed at the time of this research. The evaluation method employed is similar to that described in previous chapters. Offenders who attended Sherborne House are compared to a group of offenders for whom the centre was proposed as a sentence, but who received alternative (mainly custodial) sentences. The offending careers of the two groups are compared before and after the point at which Sherborne House was proposed as
This evaluation is the most narrowly focussed of the set of evaluation case studies reported in this thesis, in terms of the coverage of the data, which was almost entirely obtained from official sources. However information about reconvictions is here analyzed in a more sophisticated way than in previous chapters, to take into account incarceration effects and to correct for convictions registered after the start of the programme which were for offences committed some time earlier. Reconviction outcomes also are analysed in terms of their seriousness, and by level of risk of reconviction. This illustrates the range of outcome measures which can be drawn from apparently straightforward information about reconvictions.

**Probation centres in probation practice**

Probation centres represent a significant area of the probation service's work. They are a sector in which innovations and new ideas for working with offenders are often first tried out. Probation centres are useful sites for developing new models of practice because, unlike as with casework, far less work happens behind closed doors. This allows practice more easily to be documented, and allows officers to learn from each other. Mair and Nee (1993) comment that probation centres are a:

'major development in mainstream probation work - and may well be the most significant - since the introduction of the community service order...' (p329).

Probation centres have been in existence since the early 1970s, when the first four experimental 'Day Training Centres' as they were then known, were set up in Liverpool, London (now the Camberwell Probation Centre, the subject of the next chapter) Pontypridd and Sheffield. The centres operated in a legal framework set out in Section 4 of the Criminal Court Act 1973. Vanstone (1993) reviews the
influence of these centres on the probation service. Following the 1982 Criminal Justice Act many similar probation day centres were opened. Mair and Nee (1993) report the existence of between 80 and 100 probation centres. The 1991 Criminal Justice Act changed the name of these centres to 'probation centres', and guidance to this act set out that probation centre sentences should be reserved for those whose offences were the most serious that could be sentenced by way of a community disposal, and who were at high risk of reoffending.

An important factor in the development of probation centres was that over the 1980s a main policy objective of government was to reduce the use of custody for offenders in general and particularly for younger offenders (for accounts see for example Allen, 1991, Sabol, 1990, Mott, 1985.) The Parliamentary All-Party Penal Affairs Group, (1980), in a report on ways of reducing the prison population, proposed increasing day centre provision. Thus the expansion of probation centre provision was linked to efforts on the part of the probation service in the early 1980s to work with more serious offenders, by providing sentences which could be used by the courts as alternatives to custody.

As McGuire et al (1992) note, probation centres have formed a cornerstone of the probation service's diversionary strategies. As probation centres sought to work with persistent offenders, and with offenders convicted of serious offences, there was perceived to be a need to satisfy the courts that they could safely sentence offenders to this type of disposal. This was interpreted as a need to demonstrate the rigour of these sentences, in terms of the way their conditions were enforced, and offenders were controlled. For this reason, through the 1980s, probation centres were at the centre of a debate about the perceived dichotomy of care versus control, which was then taking place in the probation service. (See for example James, 1985, for an account of the role of probation centres in this debate; Raynor, 1984 and Fielding, 1984 for more general discussions of the care versus control debate in probation work.) Probation centres offered a regime, in conjunction with a
probation order, which was intended to function as a non-custodial disposal which could hold in the community offenders who would otherwise have been sentenced to custody.

The number of offenders sentenced to a probation order with a condition of attending a probation centre has remained fairly constant over the 1990s. In 1992, the year in which most of the offenders included in this follow-up study attended Sherborne House, 2902 offenders nationally attended a similar probation centre course. In 1996, the most recent year for which statistics are available, this figure was 2821, including offenders on combination orders (Home Office, 1997). The offenders included in the current research attended Sherborne House around the time of enactment of the 1991 Criminal Justice Act. At this time a substantial proportion of offenders sentenced to a probation order with a condition of attending a probation centre might otherwise have received a substantial prison sentence. However since this time the use of custody by sentencers has increased very substantially, (reflected in a rapidly rising prison population, see for example Guardian, 10 April 1998). This has had a significant impact on the type of offenders sentenced to probation centres. Whilst probation centres continue to see offenders at a similar high risk of reoffending, probation centre staff report that in the late 1990's they are working with offenders convicted of somewhat less serious crimes than in previous years, less likely to warrant a significant custodial sentence.

Probation centre programmes

The framework within which probation centre programmes operate is set out in legislation. The 1982 Criminal Justice Act created probation day centres, and under this act courts can order attendance for up to 60 days following consultation with the probation officer, arrangements being made with the centre, and with the client's
consent. Clients must comply with instructions given at the centre. In order to attend a probation centre, an offender must first be assessed, and the sentence proposed in a pre-sentence report to the court, which then decides on the sentence.

James (1985) reported that initially the probation day training centres were aimed at 'socially inadequate offenders', whose repeated offending was supposedly 'linked with an inability to cope with everyday pressures. James (1985) also noted that day centres initially were criticised for preventing offenders from taking up employment, and saw the rapid spread of these centres as being linked to the advent of mass unemployment, and the need to provide offenders with something to occupy their time during the day (Hil, 1986, describes such a centre in some detail). James (1985) states that the centres operated along therapeutic lines, until social skills courses became popular.

It is important to distinguish between some of these day-programmes for offenders which became popular in the early 1980s, intended to occupy the time of 'inadequate' or unemployed offenders; and probation centres programmes, such as Sherborne House, which explicitly aimed to offer an alternative to custody for most serious offenders likely to be sentenced in the community. By the time of the current research in the early 1990s, probation centres were very much aimed at this latter group of offenders.

Probation centres programmes now have a central aim of reducing offending. The interventions they offer are intensive, being relatively short in duration, but occupying a substantial proportion of offenders' time. These programmes generally are reserved for offenders at the highest risk of reconviction. The interventions make use of group-work techniques, and are generally focused on offending and multi-modal. Increasingly these programmes are likely to use cognitive-behavioural methods (see for example Hedderman and Sugg, 1997.) For these reasons these programmes incorporate many of the principles of 'what works' (see Chapter 1 of
this thesis, or McGuire, 1995), and thus might be expected to achieve positive results in terms of reducing reconvictions.

**Probation Centres' impact on offending**

Probation centres are known to have high reconviction rates, in the context of other community programmes. However it has been argued, (for example by Mair and Nee, 1988), that this is only to be expected of the high-risk offenders who attend these centres. Lloyd et al (1994) showed that 63 percent of offenders sentenced to probation orders with special conditions, (usually including some type of intervention, often at a probation centre) were reconvicted within 2 years of sentence. This is a significantly higher rate than the next highest risk group in this study, offenders sentenced to custody, of whom only 54 percent were reconvicted within 2 years of release.

An important finding of Lloyd et al (1994) was that the reconviction rate for conditional probation orders is slightly higher than would be predicted by the offenders' age, gender and previous convictions. These would predict that 60 percent of offenders would be reconvicted within 2 years, whereas the actual rate is 3 percent higher. However individual evaluations of the impact of probation centre programmes have been carried out which report more positive findings. Raynor (1988) reported that a day-centre programme for offenders aged 17 to 20 was successful in diverting relatively serious offenders from custody, with reconviction rates no worse, and in some respects better, than those of young offenders released from custody. This project appeared to be most effective in reducing reconviction with older offenders (meaning those aged 19 and 20), and those who had the largest number of previous convictions. Roberts (1989) reviewed a similar programme for offenders aged 17 to 20, finding that the programme successfully targeted offenders who otherwise would have been sentenced to custody. Using a quasi-
experimental methodology, Roberts found that the project group were reconvicted at a lower rate than a similar group of young offenders sentenced to custody. Maitland and Keegan (1988) carried out similar research on offenders attending Sherborne House, and reported positive findings in terms of a reconviction rate comparable or lower than rates for comparable disposals. Maitland and Keegan's (1988) research is described in more detail below.

Sherborne House: background and earlier research

Sherborne House is a probation centre for young offenders, located in South London. It was set up in 1983, in response to the Criminal Justice Act 1982. Sherborne House is in many ways typical of probation centres, as described above, having the twin aims of diverting offenders from custody, and reducing their offending. Because of its client group and central London location, Sherborne House has become well known. The stories of several young men passing though the centres are told by Graef (1994), in the form of detailed case studies. What is striking from these accounts is the fact that Sherborne House is a relatively small part of these young people's lives, an obvious fact, but one that programme-centred (rather than offender-centred) research may often neglect.

Maitland and Keegan (1988) conducted an evaluation of Sherborne House, and reported that the programme was successful in diverting offenders from custody, developing their capacity to make choices and to communicate, helping them to take advantage of opportunities to move toward employment and education, and helping them to reduce their offending. In terms of reconviction, Maitland and Keegan reported that of 35 people who started the programme, 57 percent were reconvicted within a year, the reconviction rate was slightly higher, at 60 percent for all young people sentenced to the programme. This discrepancy is because not everyone who was sentenced to Sherborne House actually started the programme.
Maitland and Keegan (1988) reported that this rate compared favourably to 2-year reconviction rates of up to 82 percent for the highest risk young offenders released from youth custody.

Maitland and Keegan's (1988) research was particularly interesting in that it used repertory grid technique (see for example Fransella and Bannister, 1977) to assess how the young offenders saw themselves in relation to their social world, presenting a number of detailed case-studies. Most offenders they assessed, who completed the programme, subsequently reported their social world to be a more complex place, containing fewer absolutes. Offenders who went on to be reconvicted were initially more likely to see themselves as being similar to their 'ideal selves'; whereas those who were not reconvicted were more likely to initially report perceiving a need to change, their 'ideal selves' being quite different to how they perceived themselves before the programme. In other words, the offenders for whom the programme 'worked', in these case-studies, may have been those who wanted, and were most ready, to change.

The Sherborne House programme: aims and methods.

Theory

The Sherborne House programme was not based on one explicit theory of offending, however it is possible to distinguish four key ideas around which the programme was designed. Firstly, was the notion that young repeat-offenders, who despite their young age already have significant track-records of offending, continue to offend partly because they evaluate the short-term benefits of crime as outweighing the long-term costs, if they think about costs at all. Secondly, an idea
was that these young people offend because they are denied legitimate opportunities for excitement, for attaining status in society and with their peers, and for progressing in life. Thirdly, underlying the programme was the idea that these offenders experience 'poor self-esteem', never seeing themselves as having been successful at anything, and so seeing crime as a route to attaining some status. The Sherborne House programme was designed to address these three perceived reasons for offending.

A fourth, and key idea underlying the design of the programme was that in order to challenge thinking about the benefits of crime and to show the costs, and in order to connect young people with legitimate opportunities, it is first necessary to engage with them. To this end the programme set out to gain their attention and enthusiasm, and draw them into the programme. The intention was that this could be achieved by offering opportunities for the young people to try out new activities they might be good at; to make them feel better about themselves. It also meant encouraging them to make the decision that crime is a negative thing in their lives. And it meant trying to give the young people the skills (including social skills) and opportunities (including opportunities for getting work, education and training) to help them to move on from a life bound up with offending.

Aims

At the time offenders included in this research attended the programme, Sherborne House defined five main aims for its work with young offenders, these were to:

- provide a credible alternative to custody for persistent or serious offenders;
- help reduce their offending;
- develop their capacity to make choices and take responsibility for their actions;
- improve their communication skills in dealing with officialdom and personal relationships;
- link them with community resources, especially in relation to employment and education.

Format and schedule of the programme

Young offenders attended the programme in groups of up to thirteen. Usually two groups ran simultaneously in parallel but largely separate programmes. Each programme ran for 10 weeks, with the young people being required to attend 4.5 days per week, with sessions running from 10.00 to 16.00. The centre also provided breakfast, and access to probation officers for discussing any problems, personal or practical, between 9.15 and 10.00 a.m.

Staffing Sherborne House

Sherborne House had a significant staffing complement. At the time offenders included in this research attended the programme, Sherborne House had 15 staff: a senior probation officer, 6 probation officers, a 'futures' post (a person who helped make links with community resources, especially in relation to employment and education), 4 craft instructors, a cook, and 2 administrative support staff. In addition other workers with specialist knowledge or skills came into the centre on a sessional basis.

Targeting offenders

The programme was aimed at offenders aged over 16 and under 21, at serious risk
of custody for a current offence because of its seriousness, or because of the number of their previous convictions. A substantial number of young offenders in the current research had committed serious offences, and if they had not attended Sherborne, could otherwise have expected a substantial prison sentence. Additional criteria were: an address in the London area, reasonable physical and mental health, and no current addiction to drink or drugs (though most reported some drink or drug use, sometimes verging on the problematic.) Also, to be suitable for the programme the young people had to be willing to address their offending, and to agree to take part in the programme.

Programme content

The content of the programme has evolved somewhat since the time the offenders included in this research attended Sherborne House. It now consists much more of structured group-work designed explicitly to tackle offending. However, at the time offenders followed up as part of this research attended the programme, the main programme elements were as follows:

- offending behaviour groups;

Offending behaviour work took the form of discussion groups, led by two probation officers, taking place four times each week. The way the groups were run was based in part on social skills training methods described by Priestley and McGuire (1985). The discussions would be initiated as general discussions about crime, and experiences of crime in general. The focus would then be moved to the participants, each setting out their previous convictions, and describing them to the group in a 'hot-seating' exercise. During this the young offenders would be challenged by probation officers, and by other group members as to whether any genuine benefits
were realised by their offending. The discussions also took in the impact of crime on victims, and also the group-members' experiences as victims of crime.

The group-work culminated in a setting out of the group members' main options for the future, showing that offending may generate income, status and thrills, but also leads to prison; whilst 'going straight' equals poverty and unemployment. The probation officers at this point would point to the benefits of other ways of living, particularly in the long term. This message would usually be reinforced by other group members.

- men and offending groups;

'Men and offending' explored links between gender and offending, based on the idea that many men who offend confuse being 'manly' with being physically aggressive.

- futures and education;

All offenders were assessed in the first week of the programme by a tutor from a local Adult Education Institute, with the option of attending educational courses, or courses in computer skills. Offenders were also supported in applying for training courses, or for starting employment. This help included assistance in applying for charitable trust money to buy equipment needed to start a business, or for work clothes.

- black empowerment / being white;
For the 'black empowerment / being white' part of the programme, offenders were divided into groups according to race. The black offenders received a 'black empowerment course' which aimed to make them feel proud of being black, by teaching black history, and the history of black people in the UK, and by covering how to deal with discrimination. The white offenders in the group were encouraged correspondingly to reflect in a positive way on what it means to be white in a multiethnic society.

- advice sessions and living sessions:

These sessions covered legal rights; drugs; alcohol; safer sex; nutrition / diet; parenting; leaving care; accommodation; mental health; DIY; budgeting / finance; coping with the system and the law. They were essentially information giving sessions which aimed to offer hard information in a dispassionate way, inviting credible outside experts (for example from drug treatment agencies) to speak to the group of young people.

- drama workshop plus two feedback sessions:

Drama was offered as another way of engaging the young people, and helping them to express themselves and discover potential in themselves of which they may not have been aware.

- multi-media design workshops:

As with the drama sessions, design workshops were offered as other ways of engaging with the young offenders, offering them samples of constructive activities
they might enjoy and be good at, which they could later follow up elsewhere. Workshops were offered in carpentry; metal work; music; photography; video making; art and pottery.

- **outdoor and adventure activities:**

Outdoor and adventure activities were offered again as ways of capturing the interest of the young people, and offering positive ways of spending time and finding excitement. Activities offered included multi-sports; orienteering; horse-riding; climbing; assault course; narrow-boat trip; carting; bowling; ice-skating; water sports; museum and exhibition visits.

- **moving on.**

All programme participants also had a field probation officer, who would take over their supervision on completion of the programme. Reports for this probation officer were prepared mid-way through the programme and on completion. A final assessment and hand-over meeting included the offender, field probation officer, and the Sherborne House-based probation officer. The aim was to provide a focus for supervision of the remainder of the probation order, with progress made at Sherborne House being incorporated in the supervision plan.

**Enforcement**

All offenders were required to agree to a contract which set out their responsibilities for following the centre's rules about time-keeping, not using drink or drugs, and respecting others. A points system was used to enforce attendance,
points being lost for lateness. Failure to keep to the contract, or persistent lateness resulted in warnings, and eventual court action.

Research design

The key question addressed by the current research is whether, as an intervention, Sherborne House is successful in reducing reoffending. In practical terms, because Sherborne House exists as one of many possible sentencing disposals, this question really means: "would the offenders who attended Sherborne have offended more or less if the centre had not existed?" To know whether a given reconviction rate is good or bad, it is necessary to have some basis for comparison, as was discussed in general terms in Chapter 2. The 'gold-standard' in research is the randomly allocated control trial. However to seek to allocate offenders at random to different community sentences would be to intervene in the sentencing process, with sentencers unlikely to cooperate. The question of how the Sherborne House programme-members would have offended if they had not attended Sherborne can best be answered by studying the offending of a comparison group, composed of offenders who were deemed suitable for the Sherborne House programme, but who were given alternative sentences. As in previous chapters, an obvious comparison group exists in the form of the group of offenders for whom Sherborne House was proposed to the court as a sentence, but who did not attend.

Although the comparison group was not selected at random, there seem good grounds to expect that these offenders might be similar to those who attended Sherborne, as all had fulfilled Sherborne's selection criteria. All offenders for whom Sherborne House was proposed as a sentencing option had first to be assessed by a probation officer, as part of the procedure for preparation of a pre-sentence report (or social inquiry report in the case of offenders referred before October 1992) for court. Offenders deemed suitable for Sherborne House, (on the grounds of
seriousness and persistence of offending, and concomitant risk of custody, coupled with the likelihood they would benefit from the programme) had then to be assessed separately by a Sherborne House probation officer. In order to be assessed, offenders were required to visit the centre and meet with a probation officer. A separate report was then prepared on the basis of this assessment, which went to the sentencer with the pre-sentence report proposing Sherborne House as a sentencing option.

Following sentence, only about half the offenders for whom Sherborne was proposed as a sentencing option were actually sentenced to a conditional probation order with a condition of attending Sherborne. The remaining offenders were, for a variety of reasons (explored below), sentenced to a different disposal. In about two-thirds of cases this was a custodial sentence (see Table 5.2). It is this latter group of offenders who form the comparison group.

**Period covered by the current research**

In this research the group who attended Sherborne House did so between April 1991 and January 1993, with 77 percent being sentenced in 1992. The comparison group was made up of all offenders referred to Sherborne over 1992 (in fact 2 percent were referred in December 1991 or January 1993), but who did not start the course.

**Results and discussion**

*Is the comparison group similar to the group of offenders who attended Sherborne House?*
When using a comparison group methodology to assess the success of an intervention, the need is to demonstrate that at the point when sentence was passed, offenders who went on to make up the comparison group and those who attended Sherborne House, are drawn from the same population of offenders. For this comparison to be valid, the comparison group must be similar to the group which went to Sherborne House. This comparison is made in Table 5.1.

Table 5.1 shows that the Sherborne and comparison groups are very similar in terms of factors relating to their offending careers. There were no statistically significant differences between the two groups. Almost all were young men, aged 18 years or thereabouts, all were unemployed. The two groups were also very similar in terms of the extent and seriousness of their previous offending. On average offenders who attended Sherborne had registered 7.1 previous convictions each, 1.2 of which on average had been dealt with by way of a custodial sentence. For the comparison group these figures were 5.4 and 1.5. The profile of previous offences was also very similar.

The Sherborne group on average had more previous convictions than the comparison group. This could be because some offenders in the comparison group met admission criteria for Sherborne House more because of the seriousness of their current offence, than because of the extent of their previous offending. Table 5.2, showing the main offence, also provides evidence that this might well be the case. This table shows the main offence for which offenders were convicted, resulting in their being proposed to the court as being suitable for Sherborne House. The profiles are very similar for the comparison group and for the group who were actually sentenced to a probation order with a condition of attending Sherborne House (referred to as the Sherborne group.) The main differences between the Sherborne and comparison groups are that a higher proportion of the comparison group had committed offences of violence against the person (12 percent compared to 6 percent of the Sherborne group); and offences of robbery (18 percent compared...
### Table 5.1. Comparison of those attending Sherborne House with those receiving alternative sentence, where there was a minimum of a one-year follow-up period.

<table>
<thead>
<tr>
<th></th>
<th>Sherborne (n=122)</th>
<th>Comparison (n=99)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at reference conviction</td>
<td>18.4 yrs</td>
<td>18.3 yrs</td>
</tr>
<tr>
<td>Percent women</td>
<td>1.5</td>
<td>1.0</td>
</tr>
<tr>
<td>Percent unemployed</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Percent risk of reconviction in 2 yrs (NRRP)</td>
<td>79</td>
<td>73</td>
</tr>
<tr>
<td>No. of previous convictions</td>
<td>7.1</td>
<td>5.4</td>
</tr>
<tr>
<td>No. of previous community sentences</td>
<td>3.8</td>
<td>2.4</td>
</tr>
<tr>
<td>No. of previous custodial sentences</td>
<td>1.2</td>
<td>1.5</td>
</tr>
<tr>
<td>No. of previous offences of...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>violence against the person</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td>burglary</td>
<td>1.5</td>
<td>1.1</td>
</tr>
<tr>
<td>sex offences</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>robbery</td>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td>taking motor vehicles</td>
<td>1.5</td>
<td>1.1</td>
</tr>
<tr>
<td>theft</td>
<td>1.8</td>
<td>1.3</td>
</tr>
<tr>
<td>fraud</td>
<td>0.1</td>
<td>0</td>
</tr>
<tr>
<td>criminal damage</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>other motor-vehicle/driving related</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>drugs</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Time from first conviction to reference</td>
<td>4.1 yrs</td>
<td>3.4 yrs</td>
</tr>
<tr>
<td>Possible follow-up period (to 30.09.94)</td>
<td>30.7 months</td>
<td>27.5 months</td>
</tr>
</tbody>
</table>

### Table 5.2. Percent committing main offence category at sentence for comparison group and for all offenders referred to Sherborne House.

#### - offenders followed up for one year

<table>
<thead>
<tr>
<th>main offence</th>
<th>comp group (n=99)</th>
<th>Sherborne House (n=122)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violence against the person</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Sexual offences</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Burglary</td>
<td>39</td>
<td>50</td>
</tr>
<tr>
<td>Robbery</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td>Taking cars</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>Theft/handling</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Fraud/forgery</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Criminal damage/arson</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>Motoring offences</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Drug offences</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>Other offences</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

#### - offenders followed up for 2 years

<table>
<thead>
<tr>
<th>main offence</th>
<th>comp. group (n=58)</th>
<th>Sherborne House(n=113)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violence against the person</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Burglary</td>
<td>41</td>
<td>50</td>
</tr>
<tr>
<td>Robbery</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>Taking cars</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Theft/handling</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>Fraud/forgery</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Criminal damage/arson</td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td>Motoring offences</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Other offences</td>
<td>2</td>
<td>-</td>
</tr>
</tbody>
</table>

154
to 9 percent of the Sherborne group.)

People convicted of violence or robbery are very likely to receive lengthy custodial sentences. Offenders in the Sherborne group were correspondingly more likely to have committed offences of taking cars (mainly Taking and Driving Away, 15 percent compared to 8 percent of the comparison group); and of burglary (50 percent compared to 39 percent of the comparison group.) However offenders in the comparison group for whom there was a two year post-release follow-up period (i.e. those who largely received short custodial sentences of around a year), had an almost identical main offence profile to the Sherborne Group. This is because the offenders in the comparison group who had committed the most serious offences, including offences against the person such as robbery or violence, were likely to have received lengthy custodial sentences, and therefore not to have been available in the community for a two year follow-up period. It may be that the comparison group subset followed up for two years, which had committed somewhat less serious offences, are most similar to the Sherborne group, and constitute the best comparison group.

Table 5.3: Alternative sentences received by comparison group

<table>
<thead>
<tr>
<th>Sentence</th>
<th>Percent (n=108)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine/discharge/bind over</td>
<td>9</td>
</tr>
<tr>
<td>Probation/supervision order</td>
<td>13</td>
</tr>
<tr>
<td>CS (up to 100 hrs)</td>
<td>6</td>
</tr>
<tr>
<td>Attendance centre</td>
<td>1</td>
</tr>
<tr>
<td>CS (over 100 hrs)</td>
<td>4</td>
</tr>
<tr>
<td>YOI (up to 6 months)</td>
<td>22</td>
</tr>
<tr>
<td>YOI (over 6 months)</td>
<td>44</td>
</tr>
<tr>
<td>Imprisonment (over 6 months)</td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
</tr>
</tbody>
</table>

* The base figure for this table includes 9 offenders for whom there was under a year of follow-up period.

Table 5.3 shows the distribution of sentences received by all offenders in the
comparison group, two thirds of alternative sentences being custodial, in all cases this being in a Young Offender Institution. It seems very likely that Sherborne House was succeeding in diverting at least a substantial proportion of its clients from custody, given the type of offences they had committed, and given the number and seriousness of their previous convictions.

Completing the programme

Of the 122 offenders who started the Sherborne House programme, 63 percent completed the course. This compares favourably to rates found in other community programmes (see for example Shorter, 1986.) Most of those who dropped out tended to do so in the first two or three weeks of the programme. The reasons for this included repeated infractions of Sherborne House's rules, leading to breach action, rearrest, or offenders simply dropping out of contact and failing to attend the programme. Whether or not offenders complete the programme is significant. Offenders who experience only a small amount of an intervention are unlikely to realise its intended benefits. Programmes which have high drop-out rates must in some way be failing to engage offenders.

Table 5.4. Comparison of those attending Sherborne House with those receiving alternative sentence, where there was a minimum of a one-year follow-up period.

<table>
<thead>
<tr>
<th>Mean....</th>
<th>Completers (n=77)</th>
<th>Non-completers (n=45)</th>
<th>Sig. in matched groups t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at reference conviction</td>
<td>18.4</td>
<td>18.6</td>
<td>-</td>
</tr>
<tr>
<td>Percent who were women</td>
<td>3</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Percent unemployed</td>
<td>100</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>Percent risk of reconviction (NRRP)</td>
<td>80</td>
<td>88</td>
<td>p&lt;0.02</td>
</tr>
<tr>
<td>No. of previous convictions</td>
<td>6.5</td>
<td>7.9</td>
<td>p&lt;0.05</td>
</tr>
<tr>
<td>No. of previous custodial sentences</td>
<td>1.0</td>
<td>1.5</td>
<td>p=0.05</td>
</tr>
</tbody>
</table>

Table 5.4 compares the group of offenders who completed the programme with

156
those who dropped out. Offenders who failed to complete tended to be at significantly higher risk of reconviction, to have had significantly more previous convictions, and on average to have been sentenced to custody on more occasions. Perhaps surprisingly, age did not seem to be associated with whether an offender completed the programme. This is contrary to Maitland and Keegan's (1988) finding, that older offenders (i.e. those aged 19 and 20 years) were significantly more likely to complete the programme.

Reconviction outcomes

This next section compares reconviction outcomes for the Sherborne and comparison groups. The main issues raised by using comparison groups to make sense of reconviction rates have been described in some detail in Chapter 2. Briefly, the most salient of these are:

- whether to look at reconviction from time of sentence or time of release for comparison group members sentenced to custody;

- whether to look at reconviction outcomes for all who started the Sherborne House programme, or for programme completers only.

Each way of evaluating reconviction outcomes has its pros and cons, and so all are used here, though at some point a decision must be made as to which is most fitting.

The comparison group is divided into people who received custodial sentences and those who received other sentences, fines and community sentences, which meant they remained in the community. The salient difference between the two groups is the period of incarceration undergone by the custody group. This meant these
offenders could not for a time register further reconvictions. Information about the reconvictions of the offenders who received custodial sentences is shown both from point of sentence, and after release. The first of these options represents the public's experience of crime from these offenders over a standard period, but the latter option probably is more useful for illustrating any treatment effect. In the tables below, where follow-up period for offenders in the comparison group sentenced to custody dates from time of release, date of release was taken as the earliest possible, calculated as half the length of the custodial sentence. The Sherborne group are divided into those who completed the programme, and those who failed to complete.

Table 5.5 shows the proportion of offenders reconvicted, and the proportion who were reconvicted and sentenced to custody, within 1 and 2 years of being sentenced to Sherborne or to an alternative sentence. For the comparison group, Table 5.5 shows reconviction outcomes for offenders sentenced to custody, and those given community sentences. From the point of view of Sherborne House, these findings do not at first seem encouraging. Offenders who attended Sherborne House were reconvicted in similar proportions to those given other sentences. However there are some positive findings. For example, after a 2-year follow-up period almost three quarters (73 percent) of offenders who attended Sherborne House had been reconvicted, but this compares to 81 percent of offenders sentenced to custody (registering reconviction from time of release for the latter group.) Furthermore, of offenders who completed the Sherborne programme, only 65 percent were reconvicted within two years. An advocate for Sherborne House could argue that a 65 percent reconviction rate, compared to an 81 percent reconviction rate represents a 20 percent reduction in reconviction. Even though this does not attain statistical significance, it is still a significant reduction in reoffending. However, the one-year reconviction outcomes, and the proportion of Sherborne offenders sentenced to custody on reconviction, are much less favourable. The verdict which is reached depends largely on which outcome measure is chosen.
Table 5.5. Reconviction outcomes by reference sentence.

<table>
<thead>
<tr>
<th></th>
<th>Offending within 1 year:</th>
<th>Offending within 2 years:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>percent reconvicted</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sentenced to cust.</td>
</tr>
<tr>
<td>Sherborne completers</td>
<td>77</td>
<td>57 17</td>
</tr>
<tr>
<td>Sherborne non-comp.</td>
<td>45</td>
<td>67 38</td>
</tr>
<tr>
<td>Sherborne total</td>
<td>122</td>
<td>63 26</td>
</tr>
<tr>
<td>Custody from sentence</td>
<td>72</td>
<td>36 14</td>
</tr>
<tr>
<td>Custody post-release</td>
<td>62</td>
<td>40 15</td>
</tr>
<tr>
<td>Other community sents.</td>
<td>37</td>
<td>62 27</td>
</tr>
<tr>
<td>(probation /CS /comb.)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The first year reconviction figures shown in Table 5.5 do not seem to show any reduction of offending resulting from the Sherborne House programme. For example, 57 percent of offenders who completed the programme reoffended within a year of sentence, compared to only 40 percent of offenders sentenced instead to custody being reconvicted within a year of release. It could be the relatively low first-year reconviction figures for the custody group post-release are due to an underestimate of the average incarceration period, as this was calculated as the minimum time which would be spent in custody, given the length of the original sentence passed.

Table 5.6 gives a more detailed breakdown of reconviction into 6 month bands, and by two levels of seriousness: all convictions, and those sentenced by way of custody. This further illustrates the difficulty in giving a simple verdict on reconviction outcomes. Looking at two-year outcomes, 73 percent of the Sherborne group were reconvicted compared to 81 percent of the group sentenced to custody, followed-up from time of release. This would represent a 10 percent reduction in offending. However 55 percent of the Sherborne group were sentenced to custody within 2 years of starting at Sherborne, compared to only 29 percent of the group who received other non-custodial sentences, and compared to 44 percent of the custody group following release. This represents a 25 percent increase in custodial sentencing of the Sherborne Group compared to the custody group. It can be argued
Table 5.6. Recoviction outcomes in cumulative six-month bands, showing average number of convictions, and percent offenders reconvicted

<table>
<thead>
<tr>
<th></th>
<th>all recovictions</th>
<th>months after sentence / release</th>
<th>Sherborne: completer</th>
<th>Sherborne: non-complete</th>
<th>Sherborne - total</th>
<th>Community sentence</th>
<th>Custody - from sentence</th>
<th>Custody - from release</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(n=77)</td>
<td>(n=77)</td>
<td>(n=45)</td>
<td>(n=37)</td>
<td>(n=72)</td>
<td>(n=69)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>36</td>
<td>57</td>
<td>65</td>
<td>65</td>
<td>76</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.51</td>
<td>1</td>
<td>1.42</td>
<td>1.94</td>
<td>2.63</td>
<td>2.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(n=66)</td>
<td>(n=46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>38</td>
<td>67</td>
<td>73</td>
<td>82</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.47</td>
<td>1.09</td>
<td>1.6</td>
<td>2.08</td>
<td>2.86</td>
<td>2.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(n=122)</td>
<td>(n=122)</td>
<td>(n=122)</td>
<td>(n=104)</td>
<td>(n=68)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>37</td>
<td>61</td>
<td>68</td>
<td>73</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.49</td>
<td>1.04</td>
<td>1.49</td>
<td>2.0</td>
<td>2.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(n=122)</td>
<td>(n=122)</td>
<td>(n=104)</td>
<td>(n=68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>43</td>
<td>62</td>
<td>73</td>
<td>74</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.65</td>
<td>1.27</td>
<td>1.54</td>
<td>1.65</td>
<td>1.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(n=37)</td>
<td>(n=37)</td>
<td>(n=31)</td>
<td>(n=12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24</td>
<td>36</td>
<td>50</td>
<td>63</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.33</td>
<td>0.6</td>
<td>1</td>
<td>1.37</td>
<td>1.96</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(n=72)</td>
<td>(n=72)</td>
<td>(n=59)</td>
<td>(n=23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td>40</td>
<td>61</td>
<td>81</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.35</td>
<td>0.68</td>
<td>1.31</td>
<td>2.07</td>
<td>3.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>custodially sentenced reconvictions</td>
<td>numbers of offenders in each outcome category are the same as for all recovictions.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>17</td>
<td>30</td>
<td>44</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.04</td>
<td>0.21</td>
<td>0.38</td>
<td>0.61</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>38</td>
<td>56</td>
<td>68</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.18</td>
<td>0.49</td>
<td>0.76</td>
<td>1.0</td>
<td>1.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>25</td>
<td>40</td>
<td>53</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.09</td>
<td>0.31</td>
<td>0.52</td>
<td>0.75</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>27</td>
<td>32</td>
<td>29</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.19</td>
<td>0.46</td>
<td>0.51</td>
<td>0.52</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td>15</td>
<td>27</td>
<td>44</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.13</td>
<td>0.21</td>
<td>0.43</td>
<td>0.67</td>
<td>1.5</td>
<td></td>
</tr>
</tbody>
</table>

160
this might be expected, on the basis that the Sherborne Group having been 'given a chance' by the court when sentenced to Sherborne, might subsequently be dealt with more harshly than would otherwise have been the case.

Looking at the cumulative number of offences committed, rather than at the proportion reconvicted, shows the cumulative impact of successive sentences. For example after 18 months, 40 percent of the Sherborne group have been sentenced to custody, so the figure for the average number of offences committed by the Sherborne group after 2 years (2 each on average) also reflects the impact of periods in incarceration resulting from these custodial sentences.

Table 5.7 adds a further degree of perspective to the reconviction outcomes, by showing relative risk of reconviction across sentence groups, calculated using the National Risk of Reconviction Predictor (Copas, 1994). This shows that compared to the custody group post release, not only were the Sherborne group reconvicted less over two years than the comparison group (73 percent compared to 81 percent of those sentenced to custody, from time of release.) But the proportion of the Sherborne groups who offended was also 10 percent less than would be predicted, on the basis of their age and previous offending, whereas 5 percent more of the group sentenced to custody was reconvicted than would be predicted on this basis.

Table 5.7. Offending within 2 years of reference sentence: comparing predicted with actual

<table>
<thead>
<tr>
<th>Percent offending within 2 years of Sherborne</th>
<th>OGRS score predicted</th>
<th>actual</th>
<th>percent less offending than predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sherborne completers (n=66)</td>
<td>77</td>
<td>65</td>
<td>16</td>
</tr>
<tr>
<td>Sherborne non-completers (n=38)</td>
<td>82</td>
<td>82</td>
<td>0</td>
</tr>
<tr>
<td>Sherborne total (n=104)</td>
<td>79</td>
<td>73</td>
<td>10</td>
</tr>
<tr>
<td>Custody from sentence (n=59)</td>
<td>-</td>
<td>63</td>
<td>-</td>
</tr>
<tr>
<td>Custody post-release (n=27)</td>
<td>77</td>
<td>81</td>
<td>-5</td>
</tr>
<tr>
<td>Other community sents. (n=31)</td>
<td>74</td>
<td>74</td>
<td>0</td>
</tr>
</tbody>
</table>

161
The reconviction predictor helps to show the Sherborne House group are some of the highest risk offenders to be found. It also helps to put what seems like an almost disastrous reconviction rate (almost three quarters of those who attended being reconvicted within two years) into some sort of perspective. Had they not gone to Sherborne House, still more might have been reconvicted.

Table 5.8 adds another layer to the analysis, by correcting for the likely level of pseudo-reconvictions. Because of the time taken by the prosecution and sentencing process, some convictions may be registered after the start of a sentence, for offences committed some months before the start of the sentence. Clearly, these convictions cannot be considered to be outcomes of that sentence. In this case, pseudo-reconvictions are those for offences which were committed before the offender was sentenced to the Sherborne House or comparison group sentence. (Evaluation issues arising from pseudo-reconvictions are discussed in Chapter 2.)

In addition to reconviction rates for the Sherborne and comparison groups, Table 5.8 also shows the two-year reconviction rate for conditional probation orders nationally, including adult as well as young offenders. The rate of pseudo-reconvictions for 4A and 4B orders nationally was determined to be 5 percent (Lloyd et al 1994). Subtracting this proportion to allow for pseudo-reconvictions gives the adjusted percentage reconvicted, likely to be closest to the true figure for reconviction outcomes. For the national sample this was 3 percent higher than predicted. For Sherborne House it is 15 percent lower.

Because the correction factor for pseudo-reconvictions is derived from national statistics, it may not be the most accurate estimate of the level of pseudo-reconvictions amongst this group. However it is more likely to be an underestimate than an overestimate of the true level, given that the unadjusted reconviction rate of the Sherborne group is slightly higher than that of the national sample. Also, as offenders who live in inner London, the Sherborne group could be being processed
simultaneously by a larger number of Crown and magistrates' courts than offenders elsewhere in the country. This can only add to the likelihood of more than one set of prosecutions being 'in the system' at any one time.

Table 5.8. Comparing predicted and actual two year reconviction rates for 4A/4B orders and for Sherborne House (national reconviction figures taken from Lloyd et al 1995 Table 1, page ix).

<table>
<thead>
<tr>
<th>sentence group</th>
<th>raw % recon.</th>
<th>adjusted % recon.</th>
<th>predicted % recon</th>
<th>diff. between pred. and actual % recon.</th>
</tr>
</thead>
<tbody>
<tr>
<td>national 4A/4B (n=3354)</td>
<td>68</td>
<td>63</td>
<td>60</td>
<td>5% more offending</td>
</tr>
<tr>
<td>Sherborne House (n=104)</td>
<td>71</td>
<td>66</td>
<td>79</td>
<td>20% less offending</td>
</tr>
<tr>
<td>comp. grp post release (n=58)</td>
<td>78</td>
<td>75*</td>
<td>73</td>
<td>3% more offending</td>
</tr>
<tr>
<td>custody post release (n=40)</td>
<td>75</td>
<td>73</td>
<td>73</td>
<td>no difference</td>
</tr>
</tbody>
</table>

* This is an estimate, calculated by adjusting the raw proportion reconvicted according to the percentage of these offenders who were sentenced to community or custodial sentences.

Correcting for pseudo-reconvictions favours the Sherborne group further over the comparison group, and over the national comparator, in terms of the programme's success in reducing reconvictions. The Sherborne House reconviction rate is not broken down into completers versus non-completers in this table, as there is no similar breakdown for the national 4A/4B comparison group.

The predictor is strikingly accurate for the comparison group: for the custody group exactly 73 percent were reconvicted within 2 years, as predicted. Slightly more of the non-custody comparison group were reconvicted than predicted (75 percent, where 73 percent was the prediction.)

Slightly more of the national comparator group were reconvicted than would be predicted on the basis of their age, sex and previous offending (63 percent, compared to a prediction of 60 percent.) However the Sherborne group were reconvicted considerably less than predicted, the adjusted figure being 66 percent,
compared to a predicted figure of 79 percent to be reconvicted within 2 years. Another way of understanding this comparison is that of 3,354 offenders sentenced to conditional probation orders, the subset who had a similar age, sex and history of offending as the Sherborne House group were more likely to be reconvicted.

Table 5.9 explores the relationship between reconviction outcome, and risk of reconviction in more detail. This table returns to using unadjusted reconviction data, dividing the Sherborne and comparison groups into three bands, according to risk of reconviction. Though these findings are not statistically significant, some interesting observations can be made.

As was apparent in Table 5.8, the reconviction predictor predicts reasonably well for the combined Sherborne and comparison groups. Offenders in the lowest likelihood of reconviction band are indeed less likely to be reconvicted than those in the medium likelihood band, who in turn are less likely to be reconvicted than those in the highest band.

Differences between the "low risk" and "high risk" bands are greater for more serious reconviction outcomes sentenced by way of custody, where the proportion in the highest band sentenced to custody is double that in the lowest band. This is really an illustration of how high a risk group this is in general. With all offenders tending to the high end of the risk distribution, the 'proportion reconvicted over two-years' sets too low a threshold to discriminate between these risk-groups. For example, where 67 percent of the comparison group in the lowest risk band are convicted of any offence within two years, compared to 94 percent of the highest risk band, this is less of a contrast than 22 percent of the lowest risk band being sentenced to custody within two years, compared to 61 percent of the highest risk band.

What is clear is that the predictor discriminates better between risk bands of the
Table 5.9. Reconviction outcomes by risk of reconviction. *(n.b. because of the small numbers in each group, no adjustment is made for pseudo-reconstructions in this table.)*

<table>
<thead>
<tr>
<th>Percent reconvicted within 12 months of sentence or release</th>
<th>Low risk</th>
<th>Med. risk</th>
<th>High risk</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed Sherborne</td>
<td>42 (n=26)</td>
<td>74 (n=23)</td>
<td>57 (n=28)</td>
<td>.61</td>
</tr>
<tr>
<td>Did not complete Sherborne</td>
<td>100 (n=5)</td>
<td>50 (n=20)</td>
<td>75 (n=20)</td>
<td></td>
</tr>
<tr>
<td>Sherborne total</td>
<td>52 (n=31)</td>
<td>63 (n=43)</td>
<td>62 (n=48)</td>
<td>.83</td>
</tr>
<tr>
<td>Custody instead of Sherborne</td>
<td>19 (n=27)</td>
<td>59 (n=17)</td>
<td>56 (n=18)</td>
<td></td>
</tr>
<tr>
<td>Non-custodial instead of Sherborne</td>
<td>59 (n=17)</td>
<td>46 (n=13)</td>
<td>100 (n=7)</td>
<td></td>
</tr>
<tr>
<td>Comparison total</td>
<td>34 (n=44)</td>
<td>53 (n=30)</td>
<td>68 (n=25)</td>
<td>.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percent reconvicted within 24 months of sentence or release</th>
<th>Low risk</th>
<th>Med. risk</th>
<th>High risk</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed Sherborne</td>
<td>48 (n=23)</td>
<td>75 (n=16)</td>
<td>74 (n=27)</td>
<td>.85</td>
</tr>
<tr>
<td>Did not complete Sherborne</td>
<td>100 (n=4)</td>
<td>69 (n=16)</td>
<td>89 (n=18)</td>
<td></td>
</tr>
<tr>
<td>Sherborne total</td>
<td>56 (n=27)</td>
<td>72 (n=32)</td>
<td>80 (n=45)</td>
<td>.98</td>
</tr>
<tr>
<td>Custody instead of Sherborne</td>
<td>70 (n=10)</td>
<td>80 (n=5)</td>
<td>92 (n=12)</td>
<td></td>
</tr>
<tr>
<td>Non-custodial instead of Sherborne</td>
<td>65 (n=17)</td>
<td>75 (n=8)</td>
<td>100 (n=6)</td>
<td></td>
</tr>
<tr>
<td>Comparison total</td>
<td>67 (n=27)</td>
<td>77 (n=13)</td>
<td>94 (n=18)</td>
<td>.98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percent sentenced by way of custody within 12 months of sentence or release (n values are the same as for all reconstructions within 12 months, as above.)</th>
<th>Low risk</th>
<th>Med. risk</th>
<th>High risk</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed Sherborne</td>
<td>12</td>
<td>22</td>
<td>18</td>
<td>.61</td>
</tr>
<tr>
<td>Did not complete Sherborne</td>
<td>40</td>
<td>35</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Sherborne total</td>
<td>16</td>
<td>28</td>
<td>27</td>
<td>.83</td>
</tr>
<tr>
<td>Custody instead of Sherborne</td>
<td>7</td>
<td>12</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Non-custodial instead of Sherborne</td>
<td>18</td>
<td>23</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>Comparison total</td>
<td>11</td>
<td>17</td>
<td>36</td>
<td>.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percent sentenced by way of custody within 24 months of sentence or release (n values are the same as for all reconstructions within 24 months, as above.)</th>
<th>Low risk</th>
<th>Med. risk</th>
<th>High risk</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed Sherborne</td>
<td>30</td>
<td>56</td>
<td>48</td>
<td>.67</td>
</tr>
<tr>
<td>Did not complete Sherborne</td>
<td>75</td>
<td>63</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>Sherborne total</td>
<td>37</td>
<td>59</td>
<td>58</td>
<td>.83</td>
</tr>
<tr>
<td>Custody instead of Sherborne</td>
<td>30</td>
<td>20</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Non-custodial instead of Sherborne</td>
<td>18</td>
<td>38</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Comparison total</td>
<td>22</td>
<td>31</td>
<td>61</td>
<td>.95</td>
</tr>
</tbody>
</table>

165
comparison group than of the Sherborne group. For most outcome measures the difference between the proportion reconvicted in the lowest and highest risk bands is greater for the comparison group than for the Sherborne group. For example, 11 percent of the comparison group in the lowest risk band are sentenced to custody within one year of their original sentence, compared to 36 percent of comparison group members in the highest risk band: more than three times the proportion in the lowest band. For the Sherborne group the equivalent figures are 16 percent in the lowest risk band, and 27 percent in the highest: less than twice the proportion in the lowest band.

The Pearson's R correlation coefficient between risk band and proportion reconvicted is shown in Table 5.9 to indicate how effective the predictor is for each group. A better measure might have been Mean Cost Rating (Inciardi et al, 1976), which takes account of the differences in the proportions in each prediction group referred to in the previous paragraph. However, Pearson's R as a measure of linearity is commonly used to express relationships between variables, and is effective in showing the key trends here. The higher the correlation, the better the predictor 'works'. What the correlations show is that correlation with risk band is lowest for the Sherborne House group, (looking either at completers, or at the total group: the group of non-completers is really too small for conclusions to be drawn from statistical observations.)

It is clear in Table 5.9 that the Pearson's R correlations are higher for the comparison groups than for the Sherborne Group. This means that the predictor 'works' better for the comparison group. This is because in the comparison group, more 'high risk' offenders are reconvicted than 'medium risk' offenders, of whom in turn more are reconvicted than 'low risk' offenders. The reason correlations are lower with the Sherborne group is that 'high risk' offenders are generally not reconvicted in larger proportions than the 'medium risk' group. This illustrates the fact that Sherborne House seems to be most successful with the highest risk
offenders. The Sherborne House reconviction rate is lower than predicted because the highest risk offenders are reconvicted less than predicted. Furthermore 'high risk' offenders who attended Sherborne offend less than 'high risk' offenders in the comparison group.

Summary of main findings

The main findings of this research are as follows:

- Sherborne House did appear to be functioning as a genuine alternative to a custodial sentence for a substantial proportion of offenders sentenced to Sherborne, in terms of the type of offences they had committed, and the nature and seriousness of their prior convictions;

- reconviction rates are high (71 percent of offenders being reconvicted over 2 years), but somewhat lower than is predicted by offenders' gender, age and offending antecedents

- reconviction rates are slightly lower than those of a comparable group of offenders sentenced to custody, following release (81 percent reconvicted over 2 years, slightly more than predicted by gender, age and previous convictions.)

- the Sherborne House programme compares still more favourably to the post-release comparison group, over two years, when pseudo reconvictions are taken into account.

- Sherborne House's raw one-year reconviction rate compares unfavourably to that of the comparison group, especially to that of comparison group members sentenced to custody.
offenders who completed the Sherborne House programme were less likely to be reconvicted than people who did not complete the programme (65 percent compared to 82 percent over 2 years.)

the above observations also largely apply to the more serious reconviction outcomes, sentenced by way of community or custodial sentences.

the Sherborne House programme seems to make most difference to the offending of the highest risk offenders.

These findings are very much in line with earlier research. The one-year reconviction rate of 63 percent for all offenders sentenced to Sherborne, is very close to the 60 percent figure reported by Maitland and Keegan, (1988.) However, whilst it is clear Sherborne House is no miracle cure for offending, these findings do provide some positive indications that the programme is reducing offending. This is particularly the case amongst offenders at the highest risk of reconviction, and amongst offenders who complete the programme.

Conclusions

Sherborne House in the context of 'what works'

A number of elements of the Sherborne House programme are consistent with what have been put forward as 'the principles of what works' (see e.g. Gerandreau et al, 1994). These are as follows:

the programme targets high risk offenders for high intensity services. The fact that the programme is most successful with the highest risk offenders is concordant with this 'Risk Principle' (see also Andrews' et al, 1990). (However it is
also the case that lower risk offenders in this study could well be deemed high risk, in the context of most offenders sentenced in the community.) The Sherborne House programme, unlike those described in earlier chapters, meets Gerandreau et al's (1994) definition of an intensive service, occupying 40 to 70 percent of offenders' time over 3 to 9 months. To speculate, it may be that the highest risk offenders who are ready to change need a programme like that of Sherborne House to help them, whereas lower risk offenders are able to change without help.

- the programme employs social learning strategies that employ modelling and reinforcement of alternatives to antisocial styles of acting.

- the programme aims to be highly responsive to the young offenders' learning styles, employing many different mechanisms to engage the programme members, consistent with Gerandreau et al's 'Responsivity' principle.

- the programme is enforced, and internal controls are applied in a fair manner, though these may be negative more than positive reinforcers (i.e. de facto penalty points.)

- the probation staff relate to offenders in interpersonally sensitive and constructive ways, and are trained and supervised appropriately, all being trained and supervised probation officers, who have opted to work at Sherborne House.

- the programme structure and activities disrupt the criminal network: though the extent to which this is the case is arguable, in that offenders spend much of their time with other offenders, however attending the programme means they may be spending less time with their usual friends and associates.

- there is a high degree of advocacy and brokerage: this being a significant part of the role of the Sherborne House probation officers, and the offender's field
probation officer.

- the goal of the programme is to reduce criminogenic needs (i.e. to work on the needs which have led to offending): though it is not clear the extent to which this is actually the case. Many of the needs the programme seeks to address, for example low self-esteem, lack of educational, training or employment opportunities, or opportunities for legitimate adventure activities, may not be closely linked to offending. On the other hand, some exercises, such as 'hot-seating' are clearly likely to get the young people thinking about why they offend.

Overall, it seems the Sherborne House programme possesses many of the characteristics claimed to be associated with successful programmes. However the main findings, whilst very much open to positive interpretation, fail to provide a strong endorsement of these principles in practice.

The extent to which the Sherborne House programme is deemed to be successful depends very much on the outcome measure chosen. If raw reconviction information is taken for all offenders starting Sherborne House and compared to similar information for all offenders sentenced to custody, with the proportion reconvicted over one year recorded from time of sentence, the Sherborne House programme appears very ineffective: 61 percent of offenders being reconvicted compared to 36 percent of the comparison group. Alternatively, only 65 percent of programme-completers were reconvicted after 2 years, compared to 81 percent of offenders released from custody. This represents a 20 percent reduction in offending, notwithstanding the fact that the Sherborne group is at higher risk of reconviction.

This highlights a tension between community safety and offender-rehabilitation. For example, as Table 5.5 shows, 70 percent of the Sherborne Group have been reconvicted within two years of sentence, each member of the Sherborne group
having been convicted twice on average. For the group sentenced to custody, these figures are 63 percent reconvicted and 1.4 convictions respectively. So custody appears to a small extent to have protected the public from crime. In fact if the Sherborne group had instead been sentenced to custody in the same way as the comparison group, the public would have been spared 77 offences leading to conviction over the following two years. But on the other hand, Sherborne House offers offenders a better chance of rehabilitation, given that over 80 percent of similar offenders sentenced to custody reoffend within two years.

A further point arises from the methodology employed to reveal these positive findings, that of using a slightly lower-risk comparison group sentenced mainly to custody. The extent of number crunching required to allow proper comparisons to be made tends to render eventual findings somewhat opaque.

All of the above illustrates the difficulty of using apparently straightforward outcome measures to deliver a simple verdict on the success of a programme. This sort of evaluation also is limited, in that it does not reveal which offenders (other than those at highest risk of reconviction) are most likely to benefit from the programme. Nor can it show how offenders have changed as a result of the programme.

The Sherborne House programme, at the time of this study, was beginning to engage with emerging findings about 'what works', and as an intensive, multifaceted programme for high risk offenders, the programme incorporated many of what have been put forward as being the characteristics of effective programmes (see for example McGuire, 1995.) However the mix of programme elements was not based on a coherent theory of offending. Nor was the programme based on research evidence as to effective methods of intervention. In contrast, the next chapter describes an evaluation of a fully-blown evidence based programme. This, the 'Reasoning and Rehabilitation' programme is based on a theory of offending
which was derived from structured review of research evidence, and on a theory of intervention similarly derived from structured review of effective methods.

Note.

1. This assertion about the changing case-mix of probation centres is based on informal interviews with probation centre staff.
Chapter 6

Reasoning and Rehabilitation: an evidence-based programme for offenders

This chapter describes an evaluation of an explicitly evidence-based programme for offenders: the Reasoning and Rehabilitation programme (see Fabiano and Ross, 1987). This intervention has been implemented widely in the probation service across England and Wales (see Hedderman and Sugg, 1997, McGuire, 1993). It was introduced from mid-1992 as the key component of a day programme for adult offenders at the Camberwell Probation Centre in South London.

Earlier chapters developed the idea that work with offenders, where possible, should be justified by evidence that it is likely to be effective in reducing offending. However the four interventions described in previous chapters were not explicitly rooted in ideas linked to research findings. Rather they were based on looser ideas about why people offend, and about what might prove effective in reducing this. In contrast, the Reasoning and Rehabilitation programme can be characterised as a model of the evidence-based approach. It is a well-defined programme, based on a theory of offending, which in turn derives from systematic reviews of research evidence (see Ross and Fabiano, 1985 Ross et al, 1986.)

The Reasoning and Rehabilitation programme has been implemented and evaluated on a large number of sites across North America (see e.g. Robinson, 1995, Ross, Fabiano and Ewles, 1988) and the United Kingdom (see e.g. McGuire, 1995, Raynor and Vanstone, 1997). This programme has been commended as a good example of an evidence-based programme in a report published by the Probation Inspectorate (Underdown, 1998.) There therefore seem good grounds to hope that this may be the best example so far of an effective programme, which will achieve significant reductions in offending.
This chapter reviews the research evidence and the theory on which the Reasoning and Rehabilitation programme is based. This chapter goes on to look at how the programme has been implemented, in the UK and in North America, and considers why the programme may have proven popular with probation services in England and Wales. This chapter then reviews the main evaluations carried out to date of the Reasoning and Rehabilitation programme's effectiveness in reducing offending.

The remainder of the chapter consists of an evaluation of the programme at Camberwell Probation Centre. As in previous chapters, this focuses on the programme's impact on reoffending, as indicated by reconviction. However this research also aims to assess the immediate impact of the programme on offenders' attitudes and thinking styles, using a set of intermediate outcome measures. These are of intrinsic interest, utilising psychometric testing as a way of assessing whether offenders' thinking styles are changed by the Reasoning and Rehabilitation Programme. These measures are intended to function as "dynamic predictors" of the type recommended by Gerandreau et al (1995). The research also uses feedback from offenders who completed the programme, obtained via structured interview. Analysis focuses on factors which predict whether an offender will complete the programme; changes in intermediate outcome measures; impact on reconviction; and how well intermediate outcome measures, and changes in these measures predict reconviction.

The chapter concludes with a brief discussion of the main lessons to be learned about what works, using research evidence in programme design, and using intermediate outcome measures to evaluate whether a programme has been successful.

The origin and design of the Reasoning and Rehabilitation Programme

Theory and evidence base
The Reasoning and Rehabilitation programme originated in Canada, in the work of a group of psychologists and criminologists working around the University of Ottawa. The theory and research evidence on which the Reasoning and Rehabilitation programme is based are set out by Ross and Fabiano (1985). The authors started out by surveying the (mainly psychological) literature about offending, in terms of what it had to say about offenders' thinking styles and what the authors term 'interpersonal cognition'. This latter term refers to the domain of understanding others and their actions, including:

'the ability to make inferences about others, to take the perspective of others, to understand the perceptions others have of oneself, and to understand social phenomena...' (Page 34).

The authors stress that there is no evidence that what they term "social intelligence" correlates with what are more commonly considered to be cognitive abilities of the type commonly assessed using IQ tests (pace other reviewers of this field, for example Farrington, 1992, who cites evidence that deficits in cognitive abilities of the type measured by 'IQ' tests, and in particular abstract reasoning skills, correlate with offending.)

Fabiano and Ross' review of literature about the psychological characteristics of offenders reaches the conclusion that a significant proportion of persistent offenders share certain common deficits, or developmental delays in the acquisition of what they term the 'cognitive skills for social competence.' These are described in detail in Fabiano and Ross (1986), and in Ross (1995), and can be summarised as:

- **impulsivity**: inadequacies in self-control and reflecting on the consequences of behaviour;

- **externality**: believing events in their lives are controlled by fate, or other external factors;

- **concrete thinking and conceptual rigidity**: lack of abstract reasoning skills,
and concomitant rigidity in thinking style, and lack of ability to associate cause and effect.

- **interpersonal problem solving skills**: lack of ability to recognise and solve problems which arise in relations with others; egocentricity and lack of ability to take on the perspective of others, lack of interpersonal understanding, judging by appearances, and missing the nuances of social interaction.

Ross and Fabiano reasoned that an individual with these deficits, in the absence of protective factors, would have problems with interpersonal relationships, be less likely to learn from errors, and ultimately be more likely to offend. Or in other words, a substantial proportion of people commit offences because of deficiencies in their 'social intelligence'. This is very much in line with other psychological explanations of why people offend (see for example Farrington, 1991.)

Fabiano and Ross (1985) then review the 'what works' literature, looking at evaluative studies of offender-interventions which took place between 1973 and 1978, in order to identify those which showed the largest impact on offending. Similar reviews contributing to the design of the reasoning and rehabilitation programme are described in Gerandreau and Ross, (1979), Ross, (1990), Ross, Fabiano & Ewles, (1988). Ross *et al* (1988) state that 'many well-controlled studies were identified which had found reductions in recidivism ranging from 30% to 74%.' Ross and Fabiano (1985) report that:

'A common component of many effective rehabilitation programmes for juvenile and adult offenders is an intervention strategy which could be expected to enhance the offenders cognitive development.'

(Page 71.)

In a structured review these authors report that of 16 programmes which contained what they describe as 'cognitive components', 15 were effective in reducing offending. The intervention methods used in these programmes are described as training in cognitive skills, such as interpersonal problem-solving skills; decision making skills; negotiating skills; and alternative thinking-skills. They also report
that the programmes which were successful in reducing offending, employed
modelling and role-playing training techniques, and that some treated the offenders
as students learning new skills, rather than as the passive subjects of intervention.
Fabiano and Ross set out their main conclusions as follows:

' - Many offenders exhibit inadequacies in their development of a number of cognitive skills.
These inadequacies may limit their ability to function effectively in a prosocial manner...

- Many correctional programmes which provide a cognitive training component have been
effective in reducing the recidivism of juvenile, adolescent and adult offenders...

- Cognitive training may be critical to the success of these programmes.' (Page 117-118)

The next, and crucial stage in Fabiano and Ross' thesis, is that the thinking skills
required to correct the deficits outlined above can be taught, via a programme of
training and education; and that this can lead to a substantial reduction in offending.
The next phase of this strand of research was to test this theory prospectively, by
designing an intervention on these principles, and testing its effectiveness in
reducing offending. The intervention designed by Fabiano and Ross is known in the
UK as 'The Reasoning and Rehabilitation Programme', and in the Correctional
Service of Canada as the Cognitive Skills Training Programme.

The Reasoning and Rehabilitation programme

For a detailed description of the Reasoning and Rehabilitation Programme, see
Ross, Fabiano and Ross (1985.) The programme is set out in highly prescriptive
terms in a manual. The personnel responsible for delivering the programme act as
teachers/trainers, guiding the offenders, who are cast as students or trainees, through
a set of exercises. The manual provides detailed lesson-plans for each session. The
actual material taught to offenders consists of a set of exercises in how to
distinguish between fact and opinion; ways of making decisions; and ways of
dealing with social situations. The training techniques used include direct
presentation of programme content, whereby the trainer acts as a 'teacher'. The teaching is combined with individual and group exercises (including role playing and video feedback) to allow offenders to practice the cognitive skills. The aim is that by using a variety of techniques and learning exercises the programme should engage participants' learning styles.

The programme is given in groups of up to 10 participants at one time, with 8 being the ideal group size. The programme takes place in 36-sessions, each session lasting 2 hours. Sessions are delivered in a pre-determined sequence, the intention being that those learned later in the programme should build on those learned earlier. The intention is that with sessions delivered on successive days, this should constitute intensive programming offering sufficient exposure for offenders to master the prescribed skills.

The intention is that referral and selection procedures should ensure the programme is only delivered to offenders likely to benefit (i.e. repeat offenders who appear to exhibit the cognitive deficits referred to above). The staff responsible for delivering the programme must all have completed a 2-week training programme, which covers the theory underlying the programme, the programme's aims, and how it should be taught.

**Piloting the programme**

The Reasoning and Rehabilitation Programme was piloted on a group of probationers, in the Pickering Project, run in the community by staff of the Correctional Service of Canada. This is described by Ross *et al* (1988) in a widely cited paper. In this study, the authors show that probationers in Ontario, who underwent the Reasoning and Rehabilitation programme were far less likely to be reconvicted than offenders on regular probation or attending a life skills course, run by the same staff. Only 19% of those who attended the Reasoning and Rehabilitation Project were reconvicted, compared to 70% of those on regular
probation and 48% of probationers attending a life skills training course. Similar positive findings are reported for the proportion of offenders on each treatment condition sentenced to custody: none of those on Reasoning and Rehabilitation, compared to 30% of regular probationers, and 11% of those on life skills training. These were described as “initial but dramatic results” (Ross et al 1991, p34). But whilst findings seem impressive, it is not clear whether the reconviction figures for Reasoning and Rehabilitation apply to all who started the programme, or only to those who completed the programme. If the latter is the case, no information is presented as to the reconvictions of people who dropped out. Also there are only small numbers of offenders in each group, there were only 62 offenders in the entire study, with 22 offenders in the Reasoning and Rehabilitation group. None the less, these are impressive findings for a pilot study.

As the programme came to be implemented more widely across North America, and the UK, other small scale studies also reported positive findings. These include De Maret, (1991), Fabiano, Porporino and Robinson (1991), Garrido and Sanchis (1991), Johnson and Hunter (1992), Ross and Ross (1989), Smith, Cox, and Mealy (1991), Valiant and Antonowicz (1990), McGuire et al (1995). However these studies either had insufficient numbers of offenders for results to be statistically significant, or relied on intermediate outcome measures, such as psychometric test results, to assess effectiveness. Nevertheless, partly as a result of the positive findings of this pilot work, the programme came to be implemented widely.

**Take-up of Reasoning and Rehabilitation by the probation service**

Reasoning and Rehabilitation has been implemented widely in the probation service in England and Wales. Hedderman and Sugg (1997) surveying the use of cognitive behavioural programmes found that 24 probation areas were running 'self control' programmes, predominantly the Reasoning and Rehabilitation Programme. McGuire (1993), surveying probation services in 1992/3 found that of 22 areas surveyed, 13
had run Reasoning and Rehabilitation programmes, and a further four were planning to do so, many of these areas running programmes at more than one location.

McGuire (1993) cites a number of reasons for this rapid and wide take-up of the programme, including:

- the existence of a sound evidence-base;
- ease of implementation, due to accessibility of programme materials, and minimal training requirements;
- successful evaluation findings for implemented programmes;
- readiness and need on the part of the probation service to demonstrate its effectiveness in reducing crime (see Chapter 1).

An important reason for the wide take-up was the high profile in probation circles of the Mid-Glamorgan probation service STOP (Straight Thinking on Probation) programme. This small South-Wales probation service trained most of its officers in delivering the Reasoning and Rehabilitation course, and was the subject of an evaluation which reported positive early findings (Raynor and Vanstone, 1992, 1993a, 1993b, 1994.)

Another factor facilitating take-up is the ease with which the programme fits within a conditional probation order, or probation centre programme.

Finally, take-up has been authoritatively reinforced by the Home Office Research and Statistics Directorate recommending the use of cognitive behavioural methods, including the Reasoning and Rehabilitation programme, (Hedderman and Sugg, 1997); and also by the Probation Inspectorate recommending the Reasoning and Rehabilitation as a good evidence-based programme (Underdown, 1998).
Evidence for the effectiveness of the Reasoning and Rehabilitation Programme

Despite the wide implementation of the Reasoning and Rehabilitation programme, evidence for its effectiveness in reducing offending, on anything like a par with that produced by the pilot study, has yet to emerge. A major study of the programme as implemented in the Correctional Service of Canadian (mainly in prisons (Research and Statistics Branch, Correctional Service of Canada, 1991) showed less pronounced reductions in readmission to prison, with 20% (8 out of 40 offenders) of the treatment group being readmitted for new convictions, compared to 30 percent (7 out of 23 offenders) of the comparison group. This finding was not statistically significant.

In the largest UK-based study of the reasoning and rehabilitation programme published to date, Raynor and Vanstone (1996) showed that 70 percent of those commencing the STOP programme (probation with a condition of attending a tailored Reasoning and Rehabilitation course) were reconvicted within 2 years, compared to 65 percent of those sentenced to regular probation, and 53 percent of those sentenced to community service. These findings look less negative when the relative levels of risk of reconviction of offenders sentenced to each type of disposal are taken into account. Only 2 percent more of offenders sentenced to STOP offend than predicted by their age and previous offending, using a reconviction predictor (Copas et al 1994). Raynor and Vanstone report more positive findings for offenders who complete the programme. There are also more pronounced reductions in offending, for STOP participants over the first year of follow-up, and STOP participants were also found to have committed less serious offences, on reconviction.

Table 6.1 shows a comparison of the findings of the largest evaluations of Reasoning and Rehabilitation presented to date. The first thing apparent in Table 6.1 is that offenders in South Wales seem to be reconvicted considerably more than
offenders in Canada. This probably illustrates a danger in making international comparisons. Without information about risk of reconviction little can be learned from comparisons. In an international context, the construct 'risk of reconviction' will as well as relating to offender histories and characteristics, be likely to encompass differences in detection rates between the towns of Mid-Glamorgan, where the STOP programme was carried out, and the whole of Canada.) More interesting is that the trends over a one-year follow-up period are very much the same (2-year information is not available for the Canadian programme.) Offenders who complete the programme are less likely to be reconvicted than offenders in the comparison group; but offenders who do not complete the programme actually are more likely to be reconvicted. In the Canadian study, overall offenders starting the Reasoning Rehabilitation programme do somewhat better than offenders who do not (though even with the large number in the study, this does not attain statistical significance.) In the STOP programme this appears not to be the case for the figures presented in Table 6.1, though the STOP participants are at higher risk of reconviction than offenders in the comparison group, which accounts for most of this difference.

<table>
<thead>
<tr>
<th></th>
<th>Correct. Serv. Canada, reconvicted in 9 mths*</th>
<th>Correct. Serv. Canada, recon. in 12 mths**</th>
<th>STOP Programme Recon. in 12 mths***</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>percent recon.</td>
<td>n</td>
</tr>
<tr>
<td>R&amp;R drop-outs</td>
<td>104</td>
<td>23</td>
<td>302</td>
</tr>
<tr>
<td>R&amp;R completers</td>
<td>446</td>
<td>13</td>
<td>1444</td>
</tr>
<tr>
<td>R&amp;R total</td>
<td>550</td>
<td>15</td>
<td>1749</td>
</tr>
<tr>
<td>control/ prob. comp.</td>
<td>207</td>
<td>14</td>
<td>379</td>
</tr>
</tbody>
</table>

*figures are taken from Table 4 and Table 9 of Ross, 1995, randomly allocated control group.
Figures in italics are calculated from those presented by Ross

**Figures are taken from Table B1, Robinson, 1995, part randomly allocated control group.

*** Figures taken from Tables 3 and 4. Raynor and Vanstone, 1996, proportion reconvicted omits false positives from pseudo-reconvictions.

The main finding which clearly emerges from Table 6.1, is that the current verdict of large-scale and well controlled research, is that the Reasoning and Rehabilitation
programme may produce modest reductions in reconviction for those who complete the programme.

A number of evaluations of the Reasoning and Rehabilitation Programme have, in addition to reconviction, looked at how offenders' attitudes have been changed by the programme (this is discussed below in the context of intermediate outcome measures). This has involved using psychometric tests (see for example Raynor, 1997, McGuire et al, 1995, Robinson et al 1991.) These report positive findings. But in terms of impact on reconviction, the best evidence available is that this programme does not seem to represent a major advance on earlier programmes less rooted in theory and research evidence.

The Camberwell Probation Centre programme

Camberwell Probation Centre (CPC), is located in Camberwell in South London. This centre was set up in 1974 as on of the four original day training centres (see Vanstone, 1996), becoming a probation day centre following the 1982 Criminal Justice Act. The place of probation day centres in probation practice is reviewed in the previous chapter. Following implementation of the 1991 Criminal Justice Act, a new Senior Probation Officer took over the management of the centre, and significant changes were made to the centre's programme. These changes were influenced by awareness of evidence about 'what works', and of the Mid-Glamorgan STOP programme. It was decided that the Reasoning and Rehabilitation programme would form the core of the new CPC programme. The Reasoning and Rehabilitation Programme was implemented from June 1992. Other elements of the programme included a black empowerment group, for black offenders, a men and offending group, for white male offenders, and craft workshops, of various types for all offenders.

The Camberwell Probation Centre was staffed by a team of seven probation officers and a senior probation officer, all of whom were trained (by Robert Ross, the
designer of the programme) in running the Reasoning and Rehabilitation Programme. There was some turnover amongst probation-officer staff, following the change of orientation of the centre. Lawrence (1992) reports that in June 1992, two thirds of staff at the centre had been there for a year or less.

Programme elements

The Reasoning and Rehabilitation Programme formed the core of the Camberwell Probation Centre programme, sessions from this taking up about half of each day. However there were other significant elements to the programme, the main elements being:

- **craft**: programme members each undertook a craft project, a creative piece of work which may have involved making a piece of furniture, or other constructive activity.

- **black empowerment**: about half the offenders who attended CPC were black, and all attended the black empowerment course element. This was part of the Inner London Probation Service Black Groups Initiative, which has been evaluated and described elsewhere (see Jeffers, 1995, Jenkins and Lawrence, 1993, Lawrence, 1992.)

- **men and offending**: instead of black empowerment, white offenders attended this programme element, which focused on the relationship between gender and offending, based on the fact most offenders are men (see for example Bented, 1994 for a discussion of the thinking underlying similar probation programmes.)

Research questions

The research reported here was commissioned at the time of initial implementation
of the Reasoning and Rehabilitation programme, as part of the overall CPC programme, in order to evaluate whether the programme, and in particular Reasoning and Rehabilitation, was effective in meeting its aim of reducing offending. Some key questions were as follows:

**targeting:** is the programme targeting offenders likely to benefit from the course, in terms of their being persistent offenders with the appropriate cognitive deficits?

**completion:** what distinguishes people who complete the course from those who drop out? Completion is an important outcome indicator in itself. Offenders who complete the course can be said to have had the 'full treatment' (Robinson, 1995). Equally completion of a part of a sentence passed by a criminal court is intrinsically important.

**changes in thinking:** the R&R Program aims to change their offending by changing their thinking styles. Do these changes happen?

**changes in self-predictions:** the programme aims to reduce the likelihood of reoffending, do offenders think this has happened?

**problem reduction:** though the programme does not address offenders' individual problems directly, (though helping offenders address their problems and making help available is a central feature of probation, see Chapter 1), it could be significant to see how offenders with different levels of problems respond to the course. It is also of interest how their perceived level of problems may change between the start of the course and completion.

**offending:** (as indicated by reconvictions) what impact is there on offending of attending the course? (Clearly the most important question as this is the raison d'être of the centre.) How does the subsequent offending of those on the programme compare with that of a similar group of offenders given different sentences? Also how do changes in
thinking, self-prediction and problem relate to offending?

**Evaluation method**

Two research designs were used in the current evaluation. The programme's immediate impact on offenders was assessed using a pre/post-test design, to measure changes in attitudes thinking styles, and offenders self-assessment as to their likelihood of reoffending. Coupled with this was a quasi-experimental pseudo-control group design, which compared the reconvictions of offenders who attended the programme with those of a similar group, assessed as suitable for the programme, but who did not attend. In addition an exit interview was used to assess a sub-set of offenders' opinions of and satisfaction with the programme.

**Intermediate outcome measures as indicators of programme effectiveness**

The ultimate aim of the Reasoning and Rehabilitation programme, in common with most offender-interventions, is to reduce offending. The best available measures of this come from information about reconvictions, collected after the programme. However reconviction data has serious drawbacks as a measure of programme effectiveness. These are discussed in depth in Chapter 2, but briefly may be summarised as:

- problems with making sense of what constitutes a good or bad reconviction outcome, which entails having an appropriate comparator (this is returned to below.)

- logistic problems arising from the delays that a follow-up period entails, and practical difficulties in obtaining information about reconvictions.

- well-analysed reconviction data may indicate whether a programme has been
successful or otherwise, but it is unlikely to be able to suggest why this may have happened.

One way of addressing these problems is by using what Mair (1991) terms 'intermediate outcome measures', which relate to whether a programme has achieved its intermediate aims. For a programme which sets out to reduce offending by, as in Chapter 3, helping offenders to learn to drive, a salient intermediate outcome measure would be the number of offenders getting driving licences as a result of the programme. For a programme such as Reasoning and Rehabilitation, which sets out to reduce offending by changing offenders thinking styles, relevant intermediate outcome measures would include measures of offenders' thinking styles, and attitudes to offending.

Intermediate outcome measures also include what Gerandreau et al (1995) term 'dynamic prediction factors'. Many factors which are known to predict offending such as 'age of first conviction' or 'number of previous convictions' can not be changed by intervention. They will be the same after a programme as before, and so are no use for indicating whether the programme has been successful. However other factors which predict offending, such as anti-social attitudes or substance-misuse, can change as a result of intervention, and so may indicate a programme is likely to reduce offending. Intermediate outcome measures also include what Losel (1995, after Andrews et al, 1990) terms 'reductions in criminogenic need'. In the first example given above, if an individuals offending directly resulted from his or her being unable to drive legitimately, this would constitute a criminogenic need. Reductions in criminogenic need (i.e. being able to drive legitimately) should be associated with reductions in offending.

The promise of intermediate outcome measures is that as dynamic prediction factors, they may provide rapid indications as to whether a programme is likely to be successful. Also where intermediate outcome measures are linked in a theoretical way to final intended outcomes, they may provide evidence as to why a programme is effective. In the case of the Reasoning and Rehabilitation programme, which aims
to change offenders’ attitudes and thinking styles, intermediate outcome indicators must include measures of attitudes to crime, and of the way offenders think, linked to the theory which underlies the programme. These indicators include a battery of psychometric tests and questionnaires which aim to measure attitudes to crime, impulsiveness, self-control behaviours, and locus of control. Also used was a problem check-list, scales whereby offenders indicated their likelihood of reoffending. In addition a standardised 'consumer feedback' interview was used to elicit offenders' views of the programme.

A monitoring package was designed and put in place at the start of the programme. This was intended to obtain information about offenders, which would enable the above questions to be answered. The choice of monitoring instruments and interim outcome measures used was informed by those used to monitor the Reasoning and Rehabilitation Programme in other settings, encompassing instruments used to evaluate the programme in other settings, including the STOP programme (see Raynor and Vanstone, 1996), the Cognitive Skills Training Programme (see Robinson et al 1991), and the Probation Evaluation Project (see McGuire et al, 1995). These included psychometric tests, a reconviction predictor scale, self assessments, problem inventories and qualitative feedback.

**Measures used**

An assessment package was administered at the start of the Camberwell Probation Centre programme. This was repeated at the end of the programme, for all offenders who completed, to assess change. Unfortunately the assessment package could only be administered to offenders who complete the course, because those who fail to complete generally do not return to the probation centre, where it was administered. Once in place, the package was administered mainly by Camberwell Probation Centre's Senior Probation Officer. The assessment package consisted of the following measures:
The CrimePics Questionnaire

One of the indicators used in the current research to assess attitudes to offending was the CrimePics questionnaire. Frude et al (1990) describe the development of this psychometric scale, which is intended to measure probationers' attitudes to offending. This was commissioned to be part of the STOP evaluation of Reasoning and Rehabilitation in probation, to be used to assess how the programme changed such attitudes (see Raynor, 1997 for a full account.)

The CrimePics Questionnaire consists of a number of scales which aim to assess different aspects of criminality, and reasons for offending. It also includes a problem checklist whereby offenders state the extent to which they have problems with various factors which may be linked to offending. There are two versions of CrimePics in existence: CrimePics I, the version used in the current research, and CrimePics II (Frude et al, 1994) a revised version.

The CrimePics questionnaire has been widely adopted by the probation service across England and Wales, being used to assess criminality, and as an outcome indicator to show the effectiveness of programmes in influencing likelihood of offending. Ellis and Underdown (1998b) state that 'CrimePics has to some extent, been adopted as the 'industry standard' for measuring attitudinal change on general offending programmes'. (para 9.7.4). The research reported in this chapter provides only the second assessment (following that of Raynor, 1997) of the predictive validity of CrimePics, in terms of its power to predict further reconvictions, and hence its utility as an intermediate outcome indicator.

- Psychometric tests

The psychometric tests, used to assess offenders' thinking styles were as follows:

- IVE Impulsiveness Inventory (Eysenck and Eysenck, 1991) this is a
standardised (on a UK sample of adults) published test, which forms part of the Eysenck personality scales. The test consists of 54 items (along the lines of 'I like to leave things to the last minute'), with which people indicate they either agree or disagree. The inventory assesses impulsivity, the extent to which people act without thinking, something which the Reasoning and Rehabilitation seeks to reduce. This test also measures 'venturesomeness', which the authors define as a healthier form of risk-taking and adventurousness; and empathy, defined as the extent to which individuals empathise with others. These latter two scales were analysed as part of the current research, but results are not presented here, because these personality factors relate less to the aims of Reasoning and Rehabilitation than does impulsiveness. Also there were no significant findings, so in the interests of space, these findings are not included.

- **Rosenbaum Schedule of Self-Control Behaviours** (Rosenbaum, 1980) this schedule aims to assess self-management behaviours, which the Reasoning and Rehabilitation programme aims to increase. It consists of 36 items, similar to: 'when I am depressed, I try to keep myself busy with things that I like', with which people indicate whether they agree or disagree, on a six-point scale. The scale was developed and validated on groups, mainly of undergraduates, in the US and in Israel.

- **Levinson Locus of Control Scales** (Levinson, 1973) this scale aims to assess the extent to which individuals feel in control of their lives, something which Reasoning and Rehabilitation seeks to increase. It consists of 24 items, along the lines of 'to a great extent my life is controlled by accidental happenings', which form three scales which assess the degree to which individuals feel internally controlled, controlled by chance events, or controlled by powerful others. The scales were developed on groups of people in the US, and validated on a sample of people admitted to a US mental hospital.

- **Self assessment information**
A straightforward, but what proved to be a very useful measure, was asking offenders to indicate the chances expressed in odds out of ten, of reoffending and of being reconvicted in the next year. For example an offender might typically indicate a five in ten chance of reoffending, but only a three in ten chance of being reconvicted. It has been found that offenders own predictions are highly predictive of their actual offending (see Bottoms et al, 1995, Burnet, 1994.) This is also a highly face-valid indicator of what the Reasoning and Rehabilitation programme is trying to achieve.

- Information reconviction and risk of reconviction

Information was obtained from the National Identification Bureau about the convictions of offenders before and after the Reasoning and Rehabilitation course. This information was also obtained for a comparison group of offenders. This information was also used to calculate offenders' risk of reconviction, using the Offender Group Reconviction Score (Probation Circular 43/97).

- Programme content / programme integrity

It is also necessary to know what was actually done in order to know what is being evaluated. The issue here is that of programme integrity: is the course what it says it is and what it is meant to be? Ideally this would include looking directly at work done in the programme, as was done in the STOP evaluation (Raynor and Vanstone, 1996). Here all Reasoning and Rehabilitation sessions were recorded on video, and a random sample viewed and rated according to a checklist as to how closely the programme was being adhered to. Unfortunately resources did not permit this in the current research. However, all staff delivering the course had received standardised training in how Reasoning and Rehabilitation should be delivered, and the programme was delivered according to the manual. In addition to this, the researcher took part in a number of the sessions, alongside the offenders.
Offenders' opinions about and satisfaction with the programme

Exit interviews were carried out with a group of twenty offenders who completed the programme, to assess what they thought of the various programme elements and the way they were delivered. These took the form of a highly structured interview, whereby offenders' responses to a standard set of questions about the programme were recorded on a set of Likert scales, and comments were recorded in writing.

Information collected

Data was available for offenders sentenced between 5 July 1991, who attended the first CPC Reasoning and Reconviction Course and 16 December 1993, by which time the 25th course was being run. Data was also available for 99 offenders referred to CPC over the same period who did not start the programme, but who were assessed as suitable for the course in Social Inquiry Reports, and (following enactment of the 1991 Criminal Justice Act) in Pre-Sentence Reports.

There is some variation in the availability of information due to minor inconsistencies in data collection and availability, set out in Table 6.2. Because of logistical problems (i.e. problems of organising the assessment days so that offenders, assessment instruments, and an experienced person to perform the assessment were present simultaneously) it was not always possible to administer the battery of assessment instruments entirely consistently.

Psychometric test data is available for the first six groups only, for a total of 58 offenders who started the programme, (28 percent of all those who started.) The use of these tests was discontinued after this time, for reasons given below. Reconviction data is available for 104 offenders who attended CPC, 99 of whom were from groups 1 to 14, and for 5 offenders from group 25.
Reconviction data was not available for 13 offenders from groups 1 to 14, nor for 4 offenders from group 25. Reconviction data was not sought for groups 15 to 24. Reconviction data is also available for the 98 offenders who make up the comparison group. Offender Group reconviction Score data is also available for all offenders for whom reconviction data is available, data from the National Identification Bureau having been used to calculate the ORGS score.

Table 6.2. Outcome data availability

<table>
<thead>
<tr>
<th>Programme</th>
<th>Starters in</th>
<th>Starters included in recon. follow-up</th>
<th>Completers included in reconv. follow-up</th>
<th>Completers in reconv. follow-up</th>
<th>Comparison group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Programme</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>starters</td>
<td>included in recon. follow-up</td>
<td>completers</td>
<td>completers in reconv. follow-up</td>
<td>group</td>
</tr>
<tr>
<td>IVE impulsiveness</td>
<td>58</td>
<td>53</td>
<td>25</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>Rosenbaum</td>
<td>57</td>
<td>53</td>
<td>25</td>
<td>12</td>
<td>-</td>
</tr>
<tr>
<td>Levinson</td>
<td>58</td>
<td>54</td>
<td>26</td>
<td>14</td>
<td>-</td>
</tr>
<tr>
<td>Crimepines</td>
<td>203</td>
<td>102</td>
<td>101</td>
<td>36</td>
<td>-</td>
</tr>
<tr>
<td>Problem check-list</td>
<td>205</td>
<td>103</td>
<td>104</td>
<td>40</td>
<td>-</td>
</tr>
<tr>
<td>Self-assessment</td>
<td>204</td>
<td>103</td>
<td>103</td>
<td>38</td>
<td>-</td>
</tr>
<tr>
<td>Reconviction</td>
<td>105</td>
<td>105</td>
<td>40</td>
<td>105</td>
<td>98</td>
</tr>
<tr>
<td>ORGS</td>
<td>99</td>
<td>99</td>
<td>40</td>
<td>99</td>
<td>98</td>
</tr>
</tbody>
</table>

The factors set out above mean that base figures for certain tables in the results section vary slightly. Table 6.2, which shows the number of offenders for whom each type of information is available, allows the reasons for this to be traced.

Results and discussion

This section sets out the main results, as they address the main research questions.

Targeting / Offender Profile

Table 6.3 shows the demographic characteristics of programme members. The average age, and the race profile closely match that of the Inner London Probation Service caseload. The reason there were so few women programme members is that
soon after the inception of the programme, a new probation centre was opened solely for women offenders, so the Camberwell Probation Centre course became solely for men after this time.

Table 6.3. Demographic characteristics

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Percent (n=104)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed at offence</td>
<td>10.6</td>
</tr>
<tr>
<td>White</td>
<td>57.1</td>
</tr>
<tr>
<td>Black*</td>
<td>38.8</td>
</tr>
<tr>
<td>Female</td>
<td>3.8</td>
</tr>
<tr>
<td>Age at offence</td>
<td>28.8</td>
</tr>
</tbody>
</table>

*Information about ethnicity was available for a subset of 98 offenders.

Table 6.4 shows clearly that this is a highly persistent group of offenders, at very high risk of reconviction. The OGRS prediction that 71 percent of offenders who attended CPC will be reconvicted within 2 years, compares to an average ORGS score of 64 percent for conditional probation orders (4A and 4B conditions, Lloyd et al 1994.) This high score reflects the fact this is a heavily convicted group, at high risk of reconviction.

Table 6.4. Previous convictions and risk of reconviction

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Mean Number or Proportion (n=99)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior sentencing occasions</td>
<td>15.4</td>
</tr>
<tr>
<td>Prior probation orders</td>
<td>1.7</td>
</tr>
<tr>
<td>Prior custodial sentences</td>
<td>5.9</td>
</tr>
<tr>
<td>OGRS predicted 2 year reconviction rate</td>
<td>71%</td>
</tr>
</tbody>
</table>

Table 6.5 shows the psychometric test scores of the Camberwell Probation Centre programme members. The advantage of using published validated psychometric tests is that it is possible to compare scores on those tests, obtained for an experimental group, with 'norm' scores: those for other populations. The CPC population should be more impulsive, less internally controlled and less self-
controlled than the general population, in order for the R&R course to be effective. Table 6.5 shows this clearly seems to be the case. Scores on these questionnaires indicate that this group of offenders is significantly less internally controlled, less self-controlled, and more impulsive than groups from the general population. This supports Ross and Fabiano's (1986) hypothesis about why these offenders offend. Furthermore, coupled with the high level of previous convictions, this seems exactly to be the group most likely to benefit from the Reasoning and Rehabilitation course.

It should be noted that the Levinson and Rosenbaum scales were validated on mainly US undergraduate, or mental hospital groups, so the relevance of the validation groups to groups of offenders in South London is questionable. (Note: the sample of CPC attendees is smaller in this table than in previous tables because this battery of psychometric tests was used for the first year only of the programme, as described above.)

<table>
<thead>
<tr>
<th></th>
<th>Levinson Locus of Control</th>
<th>Chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td>Others</td>
<td></td>
</tr>
<tr>
<td>CPC (n=58)</td>
<td>24.5</td>
<td>29.6</td>
</tr>
<tr>
<td>'Normal' sample (n=96)</td>
<td>35.5</td>
<td>16.7</td>
</tr>
<tr>
<td>Hospital sample (n=165)</td>
<td>35.4</td>
<td>23.8</td>
</tr>
</tbody>
</table>

IVE Impulsiveness

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CPC (n=58)</td>
<td>11.2</td>
</tr>
<tr>
<td>General pop. (n=559)</td>
<td>6.6</td>
</tr>
</tbody>
</table>

Rosenbaum Schedule of Self-Control Behaviours

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CPC (n=57)</td>
<td>14.3</td>
</tr>
<tr>
<td>US students (n=35)</td>
<td>25.9</td>
</tr>
<tr>
<td>Israeli student samples</td>
<td>23 to 27.</td>
</tr>
<tr>
<td>Israeli non-student, male</td>
<td>31.3</td>
</tr>
</tbody>
</table>

Table 6.6 shows data from the Crimepics I questionnaire. In the absence of norm data (i.e. of information about the sort of scores to be expected from a relevant
population), there is limited scope for interpretation. However the scores in this table are at a similar level to the Crime scale scores of offenders in STOP programme (see Raynor, 1997), and the pattern of scores is very similar. For example, the ranking of scores amongst the crime index components follows a similar pattern, with 'cost of crime' having the highest score, and 'moral attitudes' the lowest. This may indicate that the scale possesses a degree of reliability.

Table 6.6. Crime Index Components: scores at the start of the CPC programme.

<table>
<thead>
<tr>
<th>Crime index components</th>
<th>all offenders attending CPC (n=205)</th>
<th>CPC attmeanders in recon. follow-up (n=103)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of crime (C)</td>
<td>5.9</td>
<td>5.7</td>
</tr>
<tr>
<td>Responsibility (R)</td>
<td>5.6</td>
<td>5.5</td>
</tr>
<tr>
<td>Impulsiveness (I)</td>
<td>5.0</td>
<td>5.1</td>
</tr>
<tr>
<td>Moral Attitudes (M)</td>
<td>4.8</td>
<td>4.5</td>
</tr>
<tr>
<td>Awareness of effect on victims (E)</td>
<td>4.3</td>
<td>5.2</td>
</tr>
</tbody>
</table>

| Other Crimepicons components            |                                     |                                             |
|-----------------------------------------|                                     |                                             |
| Identity as criminal                    | 4.1                                 | 4.2                                         |
| Self Prediction                         | 5.6                                 | 5.6                                         |

| Problem profile (ranked)                |                                     |                                             |
|-----------------------------------------|                                     |                                             |
| Problem with money                      | 3.2                                 | 3.0                                         |
| Problem with employment                 | 3.1                                 | 3.2                                         |
| Tendency to get bored                   | 2.5                                 | 2.6                                         |
| Lots of worries                         | 2.4                                 | 2.3                                         |
| Lack of confidence                      | 2.1                                 | 2.1                                         |
| Losing temper                           | 2.0                                 | 2.0                                         |
| Problem with housing                    | 2.0                                 | 2.1                                         |
| Problem with relationships              | 1.9                                 | 2.0                                         |
| Depressed                               | 1.9                                 | 1.8                                         |
| Family problems                         | 1.8                                 | 1.8                                         |
| Problem with self-image                 | 1.8                                 | 1.8                                         |
| Problem with drink/drugs                | 1.7                                 | 1.7                                         |
| Need for extra excitement               | 1.6                                 | 1.7                                         |
| Problem with health/fitness             | 1.5                                 | 1.5                                         |
| Problem with gambling                   | 1.2                                 | 1.2                                         |

The pattern of scores in the problem checklist is also similar to that found by Raynor (1997). For example in both studies the highest scoring problems were problem with money, problem with employment, and with getting bored. This is also in common with the findings of Aubrey and Hough, 1997, who reported in a
study of the need profiles of almost 700 offenders that 'problems with employment or finance were the most commonly identified problems...' (page viii.) The fact that findings from the problem checklist data are in common with other studies is evidence that it possesses at least some degree of reliability and validity.

Completing the Camberwell Probation Centre programme

The aim of the programme, in the interests of maximising benefit to offenders and to the community, is to get as many offenders as possible to complete the course, within the rules (of attendance etc.) of the centre. If there were significant differences between completers and non-completers, this might have implications for assessment of offenders, so those most likely to complete, and benefit thereby, could be selected.

Over the period of the study 51 percent of those who attended CPC completed the course. Table 6.7. shows that completion rates improved slightly over time. However offenders included in the reconviction follow-up were drawn largely from earlier groups, to allow time for a follow-up period. So this means that the proportion of these offenders who completed the course was somewhat lower than the average, at 41 percent. This is a low completion rate in the context of community sentences, comparing to completion rates of 60 percent for later groups. These are more in the mainstream of completion rates for community programmes for high risk offenders, comparing for example to a completion rate of 63 percent for a slightly higher risk group of younger offenders at Sherborne House (see Chapter 5.) The current research did not encompass reasons for this improvement in completion rates. It may reflect the programme 'bedding in', as staff become more experienced in running the programme, and perhaps as more appropriate referrals are obtained.
Table 6.8. Predicting which offenders will complete the CPC course

<table>
<thead>
<tr>
<th>CRIME total</th>
<th>did not complete (n)</th>
<th>completed (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>26.6</td>
<td>102</td>
<td>26.6</td>
</tr>
<tr>
<td>Problem Total</td>
<td>2.1</td>
<td>2.0</td>
</tr>
<tr>
<td>Chance of reoff.</td>
<td>4.0</td>
<td>4.0</td>
</tr>
<tr>
<td>Chance of recon.</td>
<td>3.2</td>
<td>3.0</td>
</tr>
<tr>
<td>OGRS score (pct)</td>
<td>69</td>
<td>72</td>
</tr>
<tr>
<td>Rosenbaum Total</td>
<td>10.4</td>
<td>19.4</td>
</tr>
<tr>
<td>Impulsiveness</td>
<td>12.2</td>
<td>9.9</td>
</tr>
<tr>
<td>Levinson: Chance</td>
<td>31.1</td>
<td>33.1</td>
</tr>
<tr>
<td>Levinson: Others</td>
<td>28.8</td>
<td>30.7</td>
</tr>
<tr>
<td>Levinson: Internal</td>
<td>23.8</td>
<td>25.4</td>
</tr>
</tbody>
</table>

Table 6.8 shows indicator scores for all offenders assessed. The pattern of scores is the same for the sub-set of offenders included in the reconviction follow-up study, so this is not shown here. There are only very minor differences on the various indicators between offenders who completed the CPC course and those who dropped out. The only difference to attain statistical significance is for the Eysenck IVE Impulsiveness scale. The more impulsive programme participants are less likely to complete, as might be predicted.

This raises an interesting paradox. The intention behind the programme is that it should target offenders who possess certain cognitive deficits, which they need the programme to make good. However these same deficits might mean that offenders are more likely to drop out, and hence not benefit. However this is entirely academic in the current scenario, where the most significant difference, that those who drop out score more highly on the impulsiveness questionnaire, nowhere approaches a level where it might help inform targeting and selection for the
programme.

Offender feedback

Structured interviews were carried out with a group of offenders leaving the programme as part of a wider exercise to develop standardised ways of assessing offender's satisfaction with probation programmes. Twenty-one interviews were carried out in late 1994, with offenders completing the course in Groups 21 to 25. Feedback from offenders was generally positive about the wider probation centre programme, about staff, and about Reasoning and Rehabilitation.

Of the 21 offenders, 18 indicated that overall they considered the Reasoning and Rehabilitation Course to be either very good or good. Only two offenders indicated that they found the course to be poor. Similarly 18 of the 21 indicated that overall the course had had positive or very positive effect on them. 15 of the course participants indicated that they felt positive or very positive about the staff. The remaining 6 course participants indicated that they felt neutral about the staff of the centre.

With regard to different modules of the Reasoning and Rehabilitation Programme, 12 offenders pinpointed the creative thinking components of the course as being the most helpful. This module included exercises in how to evaluate issues in a structured way, weighing the plusses, minuses and interesting implications of a given issue or situation.

Answers to the question: 'what was the main thing that the Reasoning and Rehabilitation Programme has taught you?' were generally positive, and along lines which would suggest that the programme was achieving its intended aims, some typical comments were as follows:

'To think more positively';
'How not to wreck 9 months in one day';
'To be more straight';
'To sit back and analyse things for myself';

However responses to the question: 'what did you think of the Reasoning and Rehabilitation course overall?', were generally more mixed, with some typical responses were:

'It gives you time to stop and think where you've been going wrong'.
'No one's going to like it, but it's needed';
'It was like going back to school with a teacher and a blackboard';
'Some of it was insulting, because it made you feel stupid'.

When asked what effect the Reasoning and Rehabilitation had had on them, and how it had changed them, course members' responses were again very much in line with their developing the desired thinking skills, typical comments being:

'How to conduct yourself without getting into trouble';
'More chilled out';
'Learn more by listening than by acting';
'Taught me how to get through things without losing my temper';
'My thoughts and actions are clearer and slower';
'It helped to keep me clean because my time was occupied';
'I've got better friends and have dumped the bad ones';
'I box around things differently now';

Though there was some negative comment:

'It'll take more than 10 weeks to change me'
'In some ways the same as prison in that I don't want to come back'

- When offenders were asked to rate the main components of the CPC program in order of preference, the craft component was ranked highest, followed by the men and offending, and black empowerment courses. Reasoning and Rehabilitation was
ranked third.

When asked how much of a factor offending was in their lives, on average the 21 interviewees indicated that this had greatly decreased as a result of the programme, with their average rating moving from 65 down to 37 on scale of 1 to 100. Alonside the comments about the impact of the programme, this seems to be positive evidence that the Reasoning and Rehabilitation Programme was working as intended.

Changes: comparing before / after measures

The key premise of the Reasoning and Rehabilitation programme is that by changing the way offenders think, their offending will be reduced. Results reported thus far are consistent with the first part of the theory underlying the Reasoning and Rehabilitation programme, in as much as the repeat offenders in the programme do appear to possess the appropriate thinking styles. This section considers whether programme members' thinking styles did change, after the programme compared to beforehand. Table 6.9 shows the extent of these changes as indicated by the battery of psychometric tests and self-assessment schedules. Because the tests and schedules were administered at the start and on completion of the course, it is only possible to monitor change amongst programme members who completed the course.

There were significant changes on a number of the indicators. The largest and most statistically significant changes were in CrimePics crime and problem checklist scores. Increases in Crime-pics scores indicate reduced criminal attitudes. Reduction in problem score shows reduction in problems. The average problem reduction from 2.0 to 1.8 may seem slight, but this is on a scale of 1 to 4, with a minimum score of 1. So in fact this represents a 20 percent reduction in problem score, from 1.0 to 0.8, a major reduction which is highly statistically significant.

201
Table 6.9. Changes:

<table>
<thead>
<tr>
<th>Indicator</th>
<th>no. of pairs</th>
<th>mean before</th>
<th>score after</th>
<th>significance (matched pairs t-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRIME total</td>
<td>89</td>
<td>1.8</td>
<td>1.9</td>
<td>p&lt;0.0001</td>
</tr>
<tr>
<td>(CRIME total, reconviction follow-up)</td>
<td>36</td>
<td>1.7</td>
<td>1.8</td>
<td>p&lt;0.05</td>
</tr>
<tr>
<td>Problem Total</td>
<td>96</td>
<td>2.0</td>
<td>1.8</td>
<td>p&lt;0.0001</td>
</tr>
<tr>
<td>(Problem total, reconviction follow-up)</td>
<td>41</td>
<td>1.9</td>
<td>1.7</td>
<td>p&lt;0.001</td>
</tr>
<tr>
<td>Crime-pics identity as criminal</td>
<td>89</td>
<td>4.0</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>Crime-pics self prediction</td>
<td>89</td>
<td>5.7</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>Self assessed odds of reoffending</td>
<td>92</td>
<td>4.1</td>
<td>3.0</td>
<td>p&lt;.001</td>
</tr>
<tr>
<td>Self assessed odds of reconviction</td>
<td>90</td>
<td>3.1</td>
<td>2.5</td>
<td>p&lt;.05</td>
</tr>
<tr>
<td>Rosenbaum Scale Total</td>
<td>15</td>
<td>21.5</td>
<td>23.3</td>
<td></td>
</tr>
<tr>
<td>Impulsiveness</td>
<td>16</td>
<td>9.8</td>
<td>10.1</td>
<td></td>
</tr>
<tr>
<td>Levinson: Chance</td>
<td>18</td>
<td>32.1</td>
<td>32.8</td>
<td></td>
</tr>
<tr>
<td>Levinson: Others</td>
<td>18</td>
<td>29.1</td>
<td>32.3</td>
<td></td>
</tr>
<tr>
<td>Levinson: Internal</td>
<td>18</td>
<td>25.9</td>
<td>25.3</td>
<td></td>
</tr>
</tbody>
</table>

There were also significant shifts in offenders’ assessment as to their likelihood, expressed in odds out of ten, of reoffending and being reconvicted, though it is interesting to note that offenders predict that they are more likely to offend than to be reconvicted. It is also interesting that the larger average change was in predicted likelihood of offending, rather than reconviction.

The fact no significant changes were observed in psychometric test scores led to the decision to abandon their use after the first year of the programme. This was on two main grounds. Firstly it is clear that either the programme was not working (either because the theory on which it was based was wrong, or because it was badly implemented) or the tests were not working. It was hoped that the latter was the case. Second, the tests were time-consuming and problematic to administer, whilst yielding no obvious benefits for the running of the programme.

Table 6.10 shows the level of changes in problem and crime profile for offenders included in the reconviction follow-up. The overall changes in this group are the same as for the wider group shown in Table 6.9. This analysis is presented separately for the follow-up group only, because later analyses look at the
relationship between these changes and subsequent reconviction for this group.

Within the components of Crime-pics, the main positive changes were in increased awareness of the effects of crime on victims. (As before, increased Crime-pics scores reflect less criminal and more pro-social attitudes.)

Table 6.10. Mean changes in problem and crime profile

Crime index components (ranked by level of pro-social change, n=39)

<table>
<thead>
<tr>
<th></th>
<th>before</th>
<th>after</th>
<th>direction of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect - awareness of effects on victims</td>
<td>4.9</td>
<td>5.6</td>
<td>more pro-social</td>
</tr>
<tr>
<td>Impulsiveness</td>
<td>5.2</td>
<td>5.7</td>
<td>more pro-social</td>
</tr>
<tr>
<td>Responsibility - for actions</td>
<td>5.5</td>
<td>5.8</td>
<td>more pro-social</td>
</tr>
<tr>
<td>Cost - awareness of costs of crime</td>
<td>5.7</td>
<td>5.7</td>
<td>no change</td>
</tr>
<tr>
<td>Moral Attitudes - to crime</td>
<td>4.6</td>
<td>4.5</td>
<td>more pro-criminal</td>
</tr>
</tbody>
</table>

Problem check-list, (ranked by size of problem reduction, n=42)

<table>
<thead>
<tr>
<th></th>
<th>before</th>
<th>after</th>
<th>direction of change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of confidence</td>
<td>1.9</td>
<td>1.3</td>
<td>reduction</td>
</tr>
<tr>
<td>Problem with employment</td>
<td>3.1</td>
<td>2.6</td>
<td>&quot;</td>
</tr>
<tr>
<td>Problem with self-image</td>
<td>1.8</td>
<td>1.3</td>
<td>&quot;</td>
</tr>
<tr>
<td>Lots of worries</td>
<td>2.1</td>
<td>1.7</td>
<td>&quot;</td>
</tr>
<tr>
<td>Need for extra excitement</td>
<td>1.6</td>
<td>1.3</td>
<td>&quot;</td>
</tr>
<tr>
<td>Tendency to get boredom</td>
<td>2.4</td>
<td>2.1</td>
<td>&quot;</td>
</tr>
<tr>
<td>Problem with money</td>
<td>2.9</td>
<td>2.7</td>
<td>&quot;</td>
</tr>
<tr>
<td>Problem with drink/drugs</td>
<td>1.5</td>
<td>1.3</td>
<td>&quot;</td>
</tr>
<tr>
<td>Problem with housing</td>
<td>2.0</td>
<td>1.8</td>
<td>&quot;</td>
</tr>
<tr>
<td>Family problems</td>
<td>1.6</td>
<td>1.4</td>
<td>&quot;</td>
</tr>
<tr>
<td>Problem with relationships</td>
<td>1.8</td>
<td>1.7</td>
<td>&quot;</td>
</tr>
<tr>
<td>Depressed</td>
<td>1.7</td>
<td>1.6</td>
<td>&quot;</td>
</tr>
<tr>
<td>Problem with gambling</td>
<td>1.1</td>
<td>1.1</td>
<td>no change</td>
</tr>
<tr>
<td>Losing temper</td>
<td>1.7</td>
<td>1.7</td>
<td>&quot;</td>
</tr>
<tr>
<td>Problem with health/fitness</td>
<td>1.3</td>
<td>1.4</td>
<td>increase</td>
</tr>
<tr>
<td>Identity as criminal</td>
<td>4.0</td>
<td>3.8</td>
<td>more pro-social</td>
</tr>
<tr>
<td>Self Prediction</td>
<td>5.7</td>
<td>6.0</td>
<td>more pro-criminal</td>
</tr>
</tbody>
</table>

Within the problem checklist, the largest changes that offenders reported were with reduced problems relating to lack of confidence, with employment and with having lots of worries. There were no or negative changes on a number of items (these included gambling, which was a problem for only handful of offenders; losing temper, and health.) It is not clear the extent to which the problem areas on which
programme participants reported the greatest degree of positive change are the ones most likely to be criminogenic.

Table 6.11. Number of offenders whose scores increased or decreased (included in reconviction follow-up).

<table>
<thead>
<tr>
<th></th>
<th>increased</th>
<th>no change</th>
<th>decreased</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crime index: criminal attitudes</td>
<td>9</td>
<td>1</td>
<td>26</td>
</tr>
<tr>
<td>Problem profile: problems</td>
<td>9</td>
<td>3</td>
<td>29</td>
</tr>
<tr>
<td>Self prediction: chance of reoffending</td>
<td>9</td>
<td>15</td>
<td>14</td>
</tr>
</tbody>
</table>

Though the trend was towards reduced criminal attitudes and reduced problems, not all offenders’ scores improved in this respect, with some offenders actually showing increased criminal attitudes and increased problems (see Table 6.11.) The relation between the increase and decrease groups, and subsequent reconviction is explored below.

Reconviction

The most important set of results relate to reconviction, as impact on reconviction is the real test of whether the Reasoning and Rehabilitation programme achieved its aims. If the programme has no impact on reconviction, then findings relating to intermediate outcomes are of limited interest.

A comparison group is used to help make sense of reconviction findings. The comparison group was made up of offenders for whom a probation order including a condition of attending the Camberwell Probation Centre programme was proposed to the court as a sentencing option, but who for a variety of reasons received different sentences. The need is to demonstrate that the comparison group is sufficiently similar to the CPC group for the further offending of this group to be taken as a good indicator of how the CPC group would have offended if given an alternative sentence.
Are offenders in the comparison group similar to those who attended the programme?

Table 6.12, below, shows that offenders in the comparison group had a very similar offending profile to CPC programme participants, as far as the reference offence is concerned.

Table 6.12. Comparing CPC and comparison groups, reference offence

<table>
<thead>
<tr>
<th>Offence</th>
<th>Percent referred but did not attend (n=98)</th>
<th>Percent attended CPC (n=104)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violence</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Burglary</td>
<td>37</td>
<td>39</td>
</tr>
<tr>
<td>Robbery</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>TDA/A to be C/MV interference</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Theft/handling</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>Fraud/forgery</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Criminal damage/arson</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Motoring offences</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>Drug offences</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Other offences</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 6.13. Comparison group: alternative sentences

<table>
<thead>
<tr>
<th>Main sentence</th>
<th>percent of offenders (n=98)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine/discharge/bind-over</td>
<td>4</td>
</tr>
<tr>
<td>Probation/supervision</td>
<td>25</td>
</tr>
<tr>
<td>CS (up to 100 hrs)</td>
<td>4</td>
</tr>
<tr>
<td>CS (over 180 hrs)</td>
<td>4</td>
</tr>
<tr>
<td>Combination order</td>
<td>2</td>
</tr>
<tr>
<td>Suspended sentence</td>
<td>3</td>
</tr>
<tr>
<td>Custody (up to 6 mths)</td>
<td>27</td>
</tr>
<tr>
<td>Custody (over 6 mths)</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 6.13 shows that the sentences received by offenders in the comparison group were distributed at a similar level of seriousness to probation with a condition of attending CPC. This is not surprising given the intention that the CPC should be used by the courts as an alternative to a custodial sentence.
Table 6.14. Offending history and risk of further offending: for all offenders in the reconviction follow-up study*

<table>
<thead>
<tr>
<th></th>
<th>comparison group (n=98)</th>
<th>attenders (n=104)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean no. previous sentences</td>
<td>14.2</td>
<td>15.4</td>
</tr>
<tr>
<td>median no. previous sentences</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>mean no. previous custodial sentences</td>
<td>5.1</td>
<td>5.9</td>
</tr>
<tr>
<td>median no. previous custodial sentences</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>OGRS score (pct)</td>
<td>66</td>
<td>71</td>
</tr>
</tbody>
</table>

*Information in this table was obtained from the National Identification Bureau, see next section.

Table 6.14 shows that offenders who attended Camberwell Probation Centre had on average slightly more convictions, leading to a higher OGRS score predicting a higher risk of reconviction. However these differences were not statistically significant, and with a mean number of 14 and 15 previous convictions respectively, and of 5 and 6 previous custodial sentences, both were heavily convicted groups, at high risk of further offending. The comparison group seems sufficiently similar to the CPC group for the further offending of this group to be taken as an indicator of how the CPC group would have offended if given alternative sentences.

Analysis of reconviction data

A full print-out of each offender's convictions was obtained from the National Identification Bureau. Missing clients were randomly distributed. Data analysed included:

- the offence and sentence which occasioned referral to CPC
- the offence and sentence (if any occurred) immediately following that which occasioned referral to CPC.
- the Offender Group Reconviction Score (Home Office 1996) calculated manually from the full print-out of offending obtained from the NIB.

Information was sought for first reconviction after the initial sentence to either CPC
or to an alternative sentence, as set out in Table 6.13 above. Looking at first
reconviction means that only offences are analysed which can be considered as
being an outcome applying to either the CPC or to the alternative sentences. It is
possible to know the proportion of offenders reconvicted within one, two and three
years of attending CPC, but not the number or seriousness of offences committed,
nor of sentences, because of earlier less serious offences masking later more serious
ones. Thus an offender may be reconvicted after six months and sentenced by way
of a fine, but if a month later the offender is reconvicted for a more serious
offence, and perhaps given a custodial sentence, this is not taken into account in the
analysis. Despite this limitation, the data serves to replicate data available for the
evaluation of the STOP programme. However, unlike the reconviction outcome
measures used in previous chapters, it means it is not possible to show the total
amount of offending within a given period after sentence. Only the cumulative
proportion who offended and nature of offence and sentence on first reconviction
are available for analysis.

Information about reconvictions was not available from the National Identification
Bureau for 13 percent of the offenders for whom it was sought. This is a similar
proportion to that found in Chapter 2, a follow-up study of the Demonstration Unit.
Here the equivalent figure for offenders who could not be identified was 14
percent. As with the Demonstration Unit, the missing information seems to be
randomly distributed, and unlikely to influence the findings of this study.

Analysis includes one-year reconviction rates, but concentrates more on two year
rates. Whilst criticised these remain the most common measure of the effectiveness
of sentences (Lloyd et al 1995), and two years is also the period over which the
Offender Group Reconviction Score applies (Home Office, 1997).

**Reconviction outcomes**

Table 6.15 shows there was no difference in the proportion reconvicted within 2
years of offenders who attended CPC, compared to those who were given other sentences. However offenders sentenced to CPC were less likely to be sentenced to custody at first reconviction: 29 percent of the CPC group compared to 35 percent of those given other sentences. Thus attending CPC could be said to have been associated with a 17 percent reduction in custody at reconviction. This implies that the CPC group when reconvicted were committing less serious offences. This is to some extent illustrated by the fact that the CPC group were less likely to be reconvicted for an offence of violence, sex or burglary. However none of these findings are statistically significant.

Table 6.15. Percent reconvicted within 2 years of sentence for all offenders referred to CPC and followed-up.

<table>
<thead>
<tr>
<th></th>
<th>attended (n=105)</th>
<th>referred, but did not attend (n=98)</th>
</tr>
</thead>
<tbody>
<tr>
<td>for any offence</td>
<td>68</td>
<td>68</td>
</tr>
<tr>
<td>cust. sent. at recon.</td>
<td>29</td>
<td>35</td>
</tr>
<tr>
<td>violence/sex/burglary</td>
<td>20</td>
<td>23</td>
</tr>
</tbody>
</table>

Table 6.16. Percent reconvicted within 2 years for all who attended CPC

<table>
<thead>
<tr>
<th></th>
<th>completed (n=43)</th>
<th>did not complete (n=62)</th>
</tr>
</thead>
<tbody>
<tr>
<td>for any offence</td>
<td>60</td>
<td>73</td>
</tr>
<tr>
<td>custodial sentence</td>
<td>16</td>
<td>37 (p&lt;0.02)</td>
</tr>
<tr>
<td>violence/sex/burglary</td>
<td>19</td>
<td>21</td>
</tr>
</tbody>
</table>

Table 6.16 shows that offenders who complete the CPC course are less likely to be reconvicted, and significantly less likely to be sentenced to custody on reconviction. Indeed the proportion sentenced to custody of offenders who completed the course was less than half that of those who did not complete, or of the comparison group. However this does not necessarily reflect a treatment effect. Offenders who completed the CPC course may have been a self selecting group who were committed to staying out of trouble regardless of any intervention. Equally, a
treatment from which only a proportion of people can benefit is of limited utility if those who will benefit can only be identified post hoc. These issues were discussed at greater length in Chapter 2, and will be returned to in a later section. However in terms of these results, the most important are those for all who started the CPC course.

At first sight these raw findings do not seem to support the idea that the Reasoning and Rehabilitation Programme is reducing reconviction. However even the most negative finding, in Table 6.15 above shows all offenders who started the course over two years were reconvicted no more than a comparison group, most of whom spent a significant portion of that time in custody (68 percent in each case.)

The most positive interpretation of the raw findings would compare the proportion of offenders who completed the programme and were sentenced to custody, with this figure for the entire comparison group: 16 percent compared to 35 percent, a more than halving of offending at this level of seriousness. However it is not valid to compare programme completers with a mixed comparison group. Even so, only 29 percent of those who started the programme were reconvicted and sentenced to custody within 2 years, still a 17 percent reduction in offending at this level of seriousness.

Looking in more detail at reconviction outcomes

To make more sense of the raw figures, it is necessary to look in more detail at the reconviction information, and at the alternative disposals to which the comparison group were sentenced. Some reconvictions in the tables above are in fact pseudo-reconvictions: convictions for offences committed before the start of the CPC or alternative sentence, which cannot therefore be deemed outcomes of those sentences.
Correcting for pseudo-reconvictions.

Pseudo reconvictions are controlled for in this study by finding out which reconvictions are likely to be in this category, and removing the offenders with these reconvictions from further analysis. The criteria for defining a reconviction as a likely pseudo-reconviction are based on the time taken by London courts to process offenders, at the time offenders were attending the CPC. Any sentence passed in a magistrate's court within 4 weeks of the initial sentence, and any sentence passed in a crown court within 12 weeks of the initial sentence, at that time must be for an offence which predates the offence which occasioned the pre-sentence report proposing a probation order plus CPC. This can only be an estimate, based on court dates gleaned from the National Identification Bureau.

Table 6.17 shows that overall the reconvictions for 9 percent of offenders were very likely to have been pseudo-reconvictions, this proportion being slightly lower for CPC programme participants than for offenders sentenced to other disposals. As no more information was available about these offenders' reconvictions, offenders with pseudo-reconvictions were excluded from further analyses. This leaves a dataset of 185 offenders.

Table 6.17. Pseudo-reconvictions by sentence

<table>
<thead>
<tr>
<th>Initial sentence category</th>
<th>percent of offenders reconvicted, for an offence committed prior to initial sent.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine/disch (n=4)</td>
<td>50</td>
</tr>
<tr>
<td>Probation (n=24)</td>
<td>13</td>
</tr>
<tr>
<td>CS (n=8)</td>
<td>25</td>
</tr>
<tr>
<td>Comb order (n=2)</td>
<td>50</td>
</tr>
<tr>
<td>Susp Cust (n=3)</td>
<td>0</td>
</tr>
<tr>
<td>Cust &lt;6 mths (n=26)</td>
<td>12</td>
</tr>
<tr>
<td>Cust &gt;6 mths (n=31)</td>
<td>6</td>
</tr>
<tr>
<td>CPC non comp (n=62)</td>
<td>5</td>
</tr>
<tr>
<td>CPC completer (n=43)</td>
<td>5</td>
</tr>
<tr>
<td>total (n=203)</td>
<td>9</td>
</tr>
</tbody>
</table>
Reconviction outcomes by sentence-type

Table 6.18 shows the reconvictions broken down by detailed sentence category, and to aid comparisons, combined into the main categories. The most significant sentence groups are standard probation, custody following release, and CPC programme starters.

<table>
<thead>
<tr>
<th>sentence</th>
<th>reconvicted within 1 year</th>
<th>reconvicted within 2 years</th>
<th>sentenced to custody on first recoviction within 2 yrs</th>
<th>OGRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probation (n=21)</td>
<td>62</td>
<td>86</td>
<td>43</td>
<td>71</td>
</tr>
<tr>
<td>Com. Serv / Comb. (n=7)</td>
<td>43</td>
<td>71</td>
<td>57</td>
<td>64</td>
</tr>
<tr>
<td>custody: from sentence (n=52)</td>
<td>33</td>
<td>56</td>
<td>23</td>
<td>63</td>
</tr>
<tr>
<td>cust: after release (n=46, 1yr; n=30, 2 yrs) 54</td>
<td>33</td>
<td>56</td>
<td>23</td>
<td>63</td>
</tr>
<tr>
<td>CPC total (n=100, OGRS n=95)</td>
<td>45</td>
<td>67</td>
<td>28</td>
<td>70</td>
</tr>
<tr>
<td>CPC non comp (n=59, OGRS n=56)</td>
<td>48</td>
<td>71</td>
<td>36</td>
<td>72</td>
</tr>
<tr>
<td>CPC completer (n=41, OGRS n=39)</td>
<td>42</td>
<td>61</td>
<td>17</td>
<td>69</td>
</tr>
<tr>
<td>total (n=185, OGRS n=180)</td>
<td>43</td>
<td>66</td>
<td>29</td>
<td>68</td>
</tr>
</tbody>
</table>

Comparing the main sentence types in Table 6.18: regular probation, custody after release, and CPC programme starters, it is clear the CPC group shows the most favourable outcomes: 45 percent of CPC programme starters being convicted within a year, compared to 62 percent of offenders on regular probation, and 54 percent of offenders sentenced to custody, in the year following their release: 27 percent and 17 percent reductions in reconviction respectively. But it also has to be noted that looking at the custody group's reconvictions from time of sentence, a smaller proportion is convicted than are of the CPC programme starters or completers, showing a slight advantage for custody in terms of the public safety criterion, if not of the offender treatment criterion.

Table 6.19 takes the analysis a step further, by taking into account the relative risk of reconviction of the different sentence groups. This means looking at two-year
reconviction rates only, the period over which the OGRS score applies. For the custody group strictly speaking this should mean looking at reconvictions from time of release only, again the period over which the OGRS score applies.

Table 6.19. Comparing actual with predicted reconviction rates for the main types of sentence, correcting for pseudo-reconvictions: percent reconvicted within 2 years

<table>
<thead>
<tr>
<th>sentence</th>
<th>reconvicted within 2 years</th>
<th>OGRS</th>
<th>percent more who offended than predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probation (n=21)</td>
<td>86</td>
<td>71</td>
<td>17</td>
</tr>
<tr>
<td>Com. Serv / Comb. (n=7)</td>
<td>71</td>
<td>64</td>
<td>10</td>
</tr>
<tr>
<td>custody: from sentence (n=52)</td>
<td>56</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>cust: after release (n=30)</td>
<td>73</td>
<td>63</td>
<td>14</td>
</tr>
<tr>
<td>CPC total (n=100, OGRS n=95)</td>
<td>67</td>
<td>70</td>
<td>-5</td>
</tr>
<tr>
<td>CPC non-completer (n=59, OGRS n=56)</td>
<td>71</td>
<td>72</td>
<td>-0</td>
</tr>
<tr>
<td>CPC completer (n=41, OGRS n=39)</td>
<td>61</td>
<td>69</td>
<td>-13</td>
</tr>
<tr>
<td>total (n=183, OGRS n=180)</td>
<td>66</td>
<td>68</td>
<td>-3</td>
</tr>
</tbody>
</table>

In terms of predicted offending, those at highest risk of reconviction were offenders sentenced to probation, and to probation with a condition of attending CPC, the prediction being that 71 and 70 percent respectively would be reconvicted within 2 years (see Table 6.23). Offenders sentenced to custody were at lowest risk of reconviction, with 63 percent predicted to be reconvicted within 2 years. This may well reflect the way the probation service targeted offenders for the CPC, the course being aimed at persistent offenders who have committed serious offences. The offenders for whom CPC was proposed, but who were sentenced to custody, are likely to have committed more serious offences but have had fewer previous convictions, and hence be at lower risk of reconviction according to the OGRS. For offenders given community sentences the converse applies. These offenders are likely to have committed a less serious offence, but to have been deemed suitable for the CPC because of their previous offending, which places them at high risk of reconviction.

What is striking is that offenders sentenced to community disposals other than CPC
(community service, probation without conditions, or a combination order) offended at a much higher rate than predicted by their OGRS score: 30 percent more for probation and 13 percent more for community service and combination orders.

The lower than predicted reconviction rate for offenders sentenced to custody, when looking at reconvictions accrued from time of sentence, reflects the fact that the OGRS score is designed to predict reconviction from time of release, so this group has a much shorter follow-up period than allowed for by the OGRS. This does serve to illustrate how the over two years there is an incarceration effect, which means the community is protected from crime. However once released, this group continue to offend at a high rate.

**Intermediate outcome measures and reconviction**

This is the most important set of results, in methodological terms. A key theme of the thesis has been that of how to determine whether work with offenders is effective. There are drawbacks to relying on reconviction outcomes, which have been considered at length in Chapter 2. However the use of intermediate outcome measures, or dynamic predictors of reoffending, including instruments such as CrimePics, holds out the promise of quicker and more readily available feedback. The key questions are whether dynamic predictors predict reconviction (dealt with below), and the extent to which changes in dynamic predictors relate to actual offending (next section.)

**Predicting reconviction**

Table 6.20 looks at how well each assessment instrument discriminates between offenders who were reconvicted within 2 years, and those who were not. The indicator which discriminates most significantly is the Offender Group Reconviction
Score, calculated from an offenders age, gender and information about previous reconvictions. However the OGRS score is not a dynamic predictor, as none of the information from which it is calculated is amenable to change. The next most significant predictor is that of offenders self-predictions, on a scale of one to ten of how likely they are to reoffend. In fact this is a better predictor than OGRS, in terms of how well it discriminates, the only reason it ranks as less significant than OGRS is that it is available for a larger group of offenders (i.e for all offenders in the comparison group) being calculated from information which did not have to be obtained from the offenders themselves. In fact the self prediction for offenders who were reconvicted is 70 percent higher than that for offenders who did not, whereas the OGRS prediction is only 18 percent higher. This is reflected by the higher correlation with self prediction (see Table 6.20.)

<table>
<thead>
<tr>
<th></th>
<th>not recon.</th>
<th>reconvicted</th>
<th>significance (2-tailed t-test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OGRS</td>
<td>61 (n=62)</td>
<td>72 (n=118)</td>
<td>p=0.001</td>
</tr>
<tr>
<td>self-assessed chance of offend</td>
<td>2.7 (n=32)</td>
<td>4.6 (n=66)</td>
<td>p=0.003</td>
</tr>
<tr>
<td>Rosenbaum</td>
<td>18.9 (n=21)</td>
<td>9.1 (n=30)</td>
<td>p=0.07 (i.e. nearing sig.)</td>
</tr>
<tr>
<td>CRIME total</td>
<td>26.8 (n=32)</td>
<td>25.5 (n=65)</td>
<td>p=0.07</td>
</tr>
<tr>
<td>Levinson Others</td>
<td>31.7 (n=21)</td>
<td>28.6 (n=31)</td>
<td>p=0.09</td>
</tr>
<tr>
<td>Self-assess reconvict</td>
<td>2.3 (n=32)</td>
<td>3.0 (n=66)</td>
<td></td>
</tr>
<tr>
<td>Levinson Chance</td>
<td>31.5 (n=21)</td>
<td>32.7 (n=31)</td>
<td></td>
</tr>
<tr>
<td>Levinson Internal</td>
<td>23.3 (n=21)</td>
<td>25.9 (n=31)</td>
<td></td>
</tr>
<tr>
<td>IVE Impulsiveness</td>
<td>10.9 (n=20)</td>
<td>11.5 (n=31)</td>
<td></td>
</tr>
<tr>
<td>Problem total</td>
<td>2.1 (n=31)</td>
<td>2.0 (n=67)</td>
<td></td>
</tr>
</tbody>
</table>

The Rosenbaum self-control schedule appears to discriminate well between those who were and were not reconvicted, but this is an effect of the wide range of possible scores (a range of 288 is theoretically possible.) Nevertheless, this indicator does approach significance in predicting reconviction, with less self-controlled offenders being more likely to be reconvicted. This is reflected by the lower level of correlation with reconviction shown in Table 6.21.

214
Table 6.21. Comparing indicators: predicting reconviction: correlation (Pearson’s R) of indicator with reconviction

<table>
<thead>
<tr>
<th>Indicator</th>
<th>1 year of CPC</th>
<th>2 yrs of CPC</th>
<th>Custodial sentence within 2 yrs of CPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-assessment: offending (n=98)</td>
<td>0.19</td>
<td>0.30*</td>
<td>0.24</td>
</tr>
<tr>
<td>OGRS (n=180)</td>
<td>0.24**</td>
<td>0.26**</td>
<td>0.32**</td>
</tr>
<tr>
<td>Rosenbaum (n=51)</td>
<td>-0.15</td>
<td>-0.26</td>
<td>0.00</td>
</tr>
<tr>
<td>Levinson: others (n=52)</td>
<td>-0.03</td>
<td>-0.23</td>
<td>-0.03</td>
</tr>
<tr>
<td>Crime-pics total (n=98)</td>
<td>-0.29*</td>
<td>-0.18</td>
<td>-0.00</td>
</tr>
<tr>
<td>Levinson: internal (n=52)</td>
<td>0.22</td>
<td>0.15</td>
<td>-0.05</td>
</tr>
<tr>
<td>Self-assessment: reconv. (n=51)</td>
<td>0.13</td>
<td>0.14</td>
<td>0.25</td>
</tr>
<tr>
<td>Completed CPC (n=185)</td>
<td>-0.06</td>
<td>-0.11</td>
<td>-0.20</td>
</tr>
<tr>
<td>Levinson: chance (n=52)</td>
<td>0.07</td>
<td>0.09</td>
<td>-0.02</td>
</tr>
<tr>
<td>IVE: Impulsiveness (n=51)</td>
<td>-0.04</td>
<td>0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>Attended CPC (n=180)</td>
<td>0.05</td>
<td>0.04</td>
<td>-0.03</td>
</tr>
<tr>
<td>Problem total (n=97)</td>
<td>-0.08</td>
<td>-0.02</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

* Correlation is significant at the p<0.05 level.
** Correlation is significant at the p<0.01 level.

Crimepics scores are not associated significantly with reconviction over two years in a two-tailed matched groups t-test. In a one tailed test (the justification for using a one tailed test being that there is a clear hypothesis that lower scores will be associated with a higher chance of reconviction) there is a significant difference in the Crime-pics scores of offenders who are reconvicted (p=0.035). Also CrimePics does correlate significantly with reconviction over the first year of follow-up.

No other scales correlate significantly with reconviction. There is no association between impulsiveness, or locus of control and offending. This means either that locus of control, and impulsiveness do not have a bearing on likelihood of reconviction, or that the tests do not work with this group, or that they were wrongly administered. A combination of the first and second reasons seems most likely.

There is a very slight negative correlation between problem check-list total and reconviction (see Tables 6.20 and 6.21). In other words, offenders who report the most or highest level of problems are actually slightly less likely to be reconvicted than the average. Table 6.22 shows the components of the total problem score,
showing that seven types of problem correlate positively with offending, though no correlations are statistically significant, and all are at a low level. The problems which correlate most strongly with reconviction are problems with relationships, housing and with needing excitement. But eight types of problem correlate negatively with offending. Offenders who report having lots of worries, problems with money, or problems with losing their temper are actually less likely to be reconvicted. What this seems to show is that reported problems do not relate in any systematic way to reconviction, but appear to be a different kind of construct.

<table>
<thead>
<tr>
<th>problem checklist item</th>
<th>correlation (Pearson's R) with offending within 2 years (n=98)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem with relationships</td>
<td>0.19</td>
</tr>
<tr>
<td>Need for extra excitement</td>
<td>0.10</td>
</tr>
<tr>
<td>Problem with housing</td>
<td>0.09</td>
</tr>
<tr>
<td>Problem with self-image</td>
<td>0.08</td>
</tr>
<tr>
<td>Problem with gambling</td>
<td>0.04</td>
</tr>
<tr>
<td>Problem with employment</td>
<td>0.02</td>
</tr>
<tr>
<td>Family problems</td>
<td>0.00</td>
</tr>
<tr>
<td>Tendency to get bored</td>
<td>-0.02</td>
</tr>
<tr>
<td>Problem with drink/drugs</td>
<td>-0.04</td>
</tr>
<tr>
<td>Lack of confidence</td>
<td>-0.03</td>
</tr>
<tr>
<td>Problem with health/</td>
<td>-0.04</td>
</tr>
<tr>
<td>Depressed</td>
<td>-0.06</td>
</tr>
<tr>
<td>Problem with money</td>
<td>-0.14</td>
</tr>
<tr>
<td>Losing temper</td>
<td>-0.15</td>
</tr>
<tr>
<td>Lots of worries</td>
<td>-0.23</td>
</tr>
</tbody>
</table>

There is a fairly low level of overlap between problems which are predictive of offending (shown in Table 6.22 above) and problems which change over the programme (shown in Table 6.10.) The five problems areas that changed most after the course were with lack of confidence, employment, self-image, lots of worries and need for excitement. So it should perhaps not be surprising that an overall reduction in problems is not related to reconviction.
Table 6.23. Reconviction outcomes and Crime-pics scores

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>reconvicted within 1 year (percent)</th>
<th>reconvicted within 2 years (percent)</th>
<th>sentenced to custody on 1st recon. within 2 yrs (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>below ave crime index (n=48)</td>
<td>60**</td>
<td>73</td>
<td>29</td>
</tr>
<tr>
<td>above ave crime index (n=49)</td>
<td>31</td>
<td>61</td>
<td>29</td>
</tr>
<tr>
<td>lower problem score (n=49)</td>
<td>53</td>
<td>73</td>
<td>29</td>
</tr>
<tr>
<td>higher problem score (n=49)</td>
<td>38</td>
<td>63</td>
<td>29</td>
</tr>
</tbody>
</table>

**The difference in proportion of offenders with above or below average Crime-pics scores reconvicted within one year is significant at p=0.003 on a 2-tailed t-test.

Another way of looking at indicator scores is by looking at the reconviction patterns of offenders with above or below average scores (see Table 6.23, which replicates an analysis of CrimePics data carried out by Raynor, 1997). This shows a similar pattern of findings to that shown in earlier tables: offenders with more pro-criminal attitudes as assessed by Crimepics are more likely to be reconvicted, as are offenders showing below average problem-checklist scores. This effect is most pronounced in the first year after sentence.

Table 6.24. Combining attitudes, problems and reconvictions: percent reconvicted

<table>
<thead>
<tr>
<th></th>
<th>reconvicted within 1 year</th>
<th>reconvicted within 2 years</th>
<th>sentenced to cust. within 2 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>low criminal attitudes, low problems (n=27)</td>
<td>37</td>
<td>63</td>
<td>30</td>
</tr>
<tr>
<td>high criminal attitudes, low problems (n=20)</td>
<td>75</td>
<td>85</td>
<td>30</td>
</tr>
<tr>
<td>low criminal attitudes, high problems (n=21)</td>
<td>24</td>
<td>62</td>
<td>29</td>
</tr>
<tr>
<td>high criminal attitudes, high problems (n=28)</td>
<td>50</td>
<td>65</td>
<td>29</td>
</tr>
</tbody>
</table>

Table 6.24, also replicates analysis of CrimePics data carried out by Raynor (1997), and shows that problems and attitude to crime interact as might be predicted: the offenders most likely to be reconvicted are those who report above average criminal attitudes and below average problems. Similarly offenders least likely to be reconvicted are those who report below average criminal attitudes and above average problems.
The findings reported in Table 6.24 contradict those of Raynor, (1997), who found that summed problem scores (using the same checklist) were predictive of offending, and that offenders most likely to be reconvicted were those who reported high criminal attitude scores, and high problem scores.

Changes in outcome indicators and reconviction

The reason for using intermediate outcome indicators is that they should signal whether the objectives of a programme are being achieved. In this case that means that changes in intermediate outcome measures should indicate whether offending is reduced.

This research gathered information about change in attitudes and thinking by taking measures of these factors by assessing offenders before and after the CPC course. A problem is that 'after' measures are available only for those offenders who completed the course. This, combined with the fact that some of the outcome measures were used for the first year of the course only, means that the number of offenders in this section of the study are relatively small.

Table 6.25. Correlations between intermediate outcome indicators and offending within the first two years, before and after the CPC course.

<table>
<thead>
<tr>
<th></th>
<th>before</th>
<th>after</th>
<th>high score=more..</th>
<th>change in correlation with offending, comparing before and after measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rosenbaum (n=12)</td>
<td>.086</td>
<td>.248</td>
<td>self controlled</td>
<td>increases</td>
</tr>
<tr>
<td>IVE Impulsiveness (n=12)</td>
<td>.255</td>
<td>.336</td>
<td>impulsive</td>
<td>increases</td>
</tr>
<tr>
<td>Levinson Chance (n=14)</td>
<td>-.305</td>
<td>.599</td>
<td>controlled by chance</td>
<td>increases</td>
</tr>
<tr>
<td>Levinson Others (n=14)</td>
<td>-.208</td>
<td>.148</td>
<td>controlled by others</td>
<td>increases</td>
</tr>
<tr>
<td>Levinson Internal (n=14)</td>
<td>-.054</td>
<td>.449</td>
<td>controlled by self</td>
<td>increases</td>
</tr>
<tr>
<td>Crime-pics index (n=35)</td>
<td>-.385</td>
<td>-.031</td>
<td>pro-social / anti-crime problems</td>
<td>decreases</td>
</tr>
<tr>
<td>Problem scale (n=40)</td>
<td>.252</td>
<td>.142</td>
<td>likely to reoffend</td>
<td>decreases</td>
</tr>
<tr>
<td>Self assessed (n=38)</td>
<td>.328</td>
<td>.069</td>
<td></td>
<td>decreases</td>
</tr>
<tr>
<td>chance of offending</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self assessed (n=38)</td>
<td>.393</td>
<td>-.009</td>
<td>likely to be reconvicted</td>
<td>decreases</td>
</tr>
<tr>
<td>chance of reconviction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OGRS (n=40)</td>
<td>.316</td>
<td>.316</td>
<td>likely to be reconvicted</td>
<td>no change</td>
</tr>
</tbody>
</table>

218
Table 6.25 looks at how well each of the outcome indicators predicts reconviction at the outset and at the conclusion of the programme. Though none of the correlations (between indicators and reconviction within 2 years) are statistically significant, it is interesting to note that the psychometric tests became slightly better predictors of offending after the CPC course. In other words, the indicators that look at what the programme members think became less accurate as predictors of their behaviour, but the indicators of how they thought, became more accurate.

Table 6.26 looks at the changes in indicator scores in more detail.

<table>
<thead>
<tr>
<th>indicator</th>
<th>reconviction within 2 years of CPC, offenders who were...</th>
<th>significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>not recon. (n=16)</td>
<td>reconvicted (n=25)</td>
</tr>
<tr>
<td>self assessed chance of offending...</td>
<td>2.2</td>
<td>4.6</td>
</tr>
<tr>
<td>before CPC</td>
<td>3.1</td>
<td>3.6</td>
</tr>
<tr>
<td>after CPC</td>
<td>0.8</td>
<td>-1.0</td>
</tr>
<tr>
<td>change in mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosenbaum self-control scale</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- before CPC</td>
<td>22.6</td>
<td>10.8</td>
</tr>
<tr>
<td>- after CPC</td>
<td>15.0</td>
<td>27.8</td>
</tr>
<tr>
<td>- change in mean</td>
<td>-3.0</td>
<td>13.3</td>
</tr>
<tr>
<td>Crimepocs CRIME total...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- before CPC</td>
<td>27.4</td>
<td>24.8</td>
</tr>
<tr>
<td>- after CPC</td>
<td>27.5</td>
<td>27.1</td>
</tr>
<tr>
<td>- change in mean</td>
<td>0.4</td>
<td>2.7</td>
</tr>
<tr>
<td>Problem total...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- before CPC</td>
<td>1.75</td>
<td>1.98</td>
</tr>
<tr>
<td>- after CPC</td>
<td>1.60</td>
<td>1.72</td>
</tr>
<tr>
<td>- change in mean</td>
<td>0.13</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Table 6.26 shows the average before and after scores, and the average change in the indicator score for the outcome indicators which showed some association with reconviction, or which showed some significant shift over the programme. These findings necessarily are for programme completers only, the table divides them into those who later were reconvicted, and those who were not reconvicted within a two-year follow-up period. Looking first at self-assessed chance of reconviction and at offenders who were not later reconvicted, before the programme these offenders on average assessed their chances of reoffending as being 2.2 out of 10. But after
the programme, on average they revised their assessment upwards, to 3.1 out of 10, an apparent failure according to this indicator. This means that after the programme, they said they were more likely to be reconvicted on average, but they were not. Paradoxically, looking at offenders who were later reconvicted, before the programme they assessed their likelihood of reoffending on average as being 4.6 out of 10. But after the programme they revised this estimate downward to 3.6 out of 10, an apparent success according to this indicator. After the programme they said they were less likely to reoffend, but they did.

Similar findings apply to the other intermediate outcome indictors in Table 6.26. Looking just at the column for offenders who were not reconvicted, these people appear not to have benefited from the CPC course, in that they show:

- an upward (i.e. more likely to offend) shift in their self assessed likelihood of reconviction;
- almost no change in their Crime-pics CRIME index score;
- a downward (i.e. less self-controlled) shift in their Rosenbaum Schedule of Self-Control behaviours;
- a small downward shift in their problems.

In contrast offenders who were reconvicted were the ones who reported the greatest improvements on all these indices.

The findings in this table are counter intuitive, however referring back to Table 6.11 makes it somewhat easier to understand: whilst the average shifts in indicator scores for all offenders who completed tend to be in a more self-controlled, more pro-social direction, a significant minority of offenders show no change, or report deterioration on these indicators. In fact 14 offenders shifted their predictions downwards, as to the likelihood of their reoffending. But of these 14 people, 12 (86 percent) were reconvicted within 2 years. 15 offenders reported the same chances out of ten of reoffending, at the end of the course as they did at the beginning. Only 8 (53 percent) of these offenders were reconvicted within 2 years. There were
9 offenders who revised their self-predicted likelihood of offending upward after the programme, of whom 5 (56 percent) were reconvicted. It is the offenders most likely to report positive change who are most likely to be reconvicted.

Table 6.27 presents data on changes in outcome measures in a different way again, showing a more complex picture. The columns of this table divide offenders into different groups for each of the intermediate outcome indicators. The first column is for those offenders who showed either negative or no change according to that indicator. The second column is for the generally larger group of offenders who showed prosocial change. These trends of Table 6.26 are repeated, though less strongly for the 1 year and custody reconviction outcome measures. For example looking at self-prediction in chances out of ten, it is only for the 2-year reconviction that offenders who report prosocial change offend significantly more.

<table>
<thead>
<tr>
<th>Shift in self-prediction</th>
<th>no change (n=15) or more likely to reoffend (n=9)</th>
<th>less likely to reoffend (n=13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>reconvicted in first year</td>
<td>46</td>
<td>43</td>
</tr>
<tr>
<td>reconvicted within two years</td>
<td>54</td>
<td>86 p&lt; 0.05 (2-tail t-test)</td>
</tr>
<tr>
<td>cust. sent. at reconviction in 2 yrs</td>
<td>17</td>
<td>21</td>
</tr>
<tr>
<td>Shift in Crime-pics crime score</td>
<td>negative or no change (n=10)</td>
<td>positive (pro-social change (n=25)</td>
</tr>
<tr>
<td>reconvicted within 1 year</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>reconvicted within 2 years</td>
<td>50</td>
<td>60</td>
</tr>
<tr>
<td>sentenced to custody within 2 years</td>
<td>30</td>
<td>12</td>
</tr>
<tr>
<td>Shift in problem score</td>
<td>probs got worse /no change (n=12)</td>
<td>problems improved (n=27)</td>
</tr>
<tr>
<td>reconvicted within 1 year</td>
<td>33</td>
<td>44</td>
</tr>
<tr>
<td>reconvicted within 2 years</td>
<td>50</td>
<td>67</td>
</tr>
<tr>
<td>sentenced to custody within 2 years</td>
<td>0</td>
<td>26 p&lt;0.01 (2-tailed t-test)</td>
</tr>
</tbody>
</table>

The situation is more complex again for CrimePics scores. There is no difference,
In terms of one-year reconviction outcomes between 'prosocial shifters' and 'negative/non-shifters'. For 2-year reconvictions, the 'prosocial shifters' are somewhat more likely to be reconvicted. But for 2-year reconvictions sentenced by way of custody, the pro-social shifters do markedly better. Those whose score shifted positively were less than half as likely to be sentenced to custody. However because of the small numbers none of the findings relating to CrimePics are statistically significant, so these could be chance findings.

The findings for changes in offenders assessments of their problems, whilst remaining counter-intuitive are at least consistent with the earlier finding that more problems appears to equate with less offending. Offenders who reported their problems in general improving were significantly more likely according to each of the reconviction outcome measures: the 1, 2 and 2-year custody rates (the latter being statistically significant at the p<0.02 level in a 2-tailed t-test.

Summary

This section summarises the main findings presented above, relating them to the initial research questions:

targeting:

- the programme successfully targeted repeat offenders who were at high risk of reoffending, many of whom might otherwise have been sentenced to custody;

- according to a set of psychometric tests, these offenders appeared to have thinking styles of the sort the Reasoning and Rehabilitation Programme was designed to correct. These included impulsiveness, lack of self-control, criminal and antisocial attitudes, and a tendency to attribute events in their lives to external factors such as chance events. In addition the programme participants reported
having a wide range of problems in their lives.

**completion:**

- at the time offenders in the follow-up research attended the Camberwell Probation Centre, completion rates were low, at 41 percent, but as the course bedded in over time completion rates rose to 60 percent.

- there were only very minor differences on the set of intermediate outcome indicators between offenders who completed the course and those who dropped out, the only statistically significant difference being that offenders who dropped out were more impulsive.

- offenders completing the programme reported a high level of satisfaction, many making comments that would suggest the programme was achieving its intended impact.

**changes in thinking, in self-predictions and in problems:**

- offenders who completed the course showed statistically significant reductions in criminal attitudes, reported problems, and self assessed likelihood of reoffending. But they showed only very minor shifts on self-control, locus of control and impulsiveness.

- the problems reported by offenders as being most reduced were with lack of confidence, employment, self image, having lots of worries, and needing extra excitement.

- not all offenders shifted in a more prosocial direction. On some indicators, some offenders showed no change and some showed more antisocial scores after the programme. Similarly some offenders reported increased problems after the programme.
**Impact on Offending:**

Overall the programme seems to be effective in reducing reconviction, but how effective the programme is deemed to be in reducing offending depends on the outcome measure chosen and on how comparisons are made. For example:

- 45 percent of offenders who started, and 42 percent those who completed the programme were reconvicted within a year, compared to 54 percent of a lower-risk group released to custody, and 62 percent of a group sentenced to regular probation.

- Only 17 percent of offenders who completed the programme were reconvicted and sentenced to custody within 2 years, compared to 33 percent of offenders sentenced to custody within 2 years of release. But alternatively:

- 67 percent of the CPC group were reconvicted within 2 years, exactly the same proportion as were reconvicted of the comparison group (when the period spent in custody by comparison group members is included in the follow-up period.)

- from time of sentence, only 56 percent of offenders sentenced to custody instead of the CPC were reconvicted within 2 years (again because of the incarceration effect), compared to 67 percent of the CPC group.

**Relation of Changes in Thinking, Self-Prediction and Problems to Offending:**

- a number of the indicators predicted reconviction, the best being offenders' own predictions. The CrimePics questionnaire, of attitudes to offending, only neared significance as a predictor, as did the self-control schedule. Offenders who reported having more problems were actually slightly less likely to be reconvicted.

- no individual type of problem correlated significantly with reconviction,
though that which correlated most strongly was 'problem with relationships'.

- offenders who after the programme said they were less likely to be reconvicted, compared to beforehand, were significantly more likely to be reconvicted.

- on no outcome measures were prosocial shifts associated with reduced reconviction.

- of the intermediate outcome indicators, those which measured attitude, or what offenders think, were better predictors of reconviction before the programme than after. Measures of personality, or how offenders think were slightly better predictors after the programme.

Conclusions

This section focuses on what the results of the current study have to say about the Reasoning and Rehabilitation Programme, and its theoretical grounding. This study raises important issues about the relationship of theory and evidence; and how this in turn relates to practical work with offenders, but these are discussed in the next chapter.

From these findings it is reasonable to conclude that the Reasoning and Rehabilitation Programme has a modest impact in reducing reconviction (though that depends on how comparisons are made vis a vis custodial sentences.) These findings, whilst less positive than those of Ross (1988) are considerably more positive than those of the STOP programme evaluation (Raynor and Vanstone, 1996), and are slightly more positive than those of Robinson (1995). What is less clear is how reductions in reconviction were achieved, and what this conveys about the theory which underlies the programme.
A main strand which runs through these findings is the difficulty of separating conceptual from methodological issues. This study illustrates how problematic it is to separate questions as to "what works' from those of 'how do we know what works?' The next section offers some comments with regard to methodology.

Methodology: outcome measures and their interpretation

It would be easier to distinguish theoretical issues, in making sense of how and how well programmes reduce offending, were it possible to agree protocols across evaluative studies of offender-programmes. These could resolve, or at least set out a consistent position, on issues of methodologies used to make sense of reconviction rates. Such protocols could resolve how comparisons are to be made between groups sentenced in the community, or to custody; and whether the criterion measure should be one of community safety, or of offender rehabilitation. They could also suggest consistency over length of follow-up period.

Consistency is also needed with regard to how programme completion relates to programme effectiveness. It is meaningless to compare a group of offenders who complete a programme without infraction, to a group of offenders sentenced to other conditions where their progress is not known. Yet this is often done when comparison group are used. Whether an offender completes a programme is something which can only be known post-hoc, so information about success rates for completers is of little use in isolation from the course's completion rate. It follows that the best indicator of the effectiveness of a programme is one which combines success rates for completers with those for drop-outs, in other words the success rate for programme starters.

It is specious to argue that drop-outs from the Reasoning and Rehabilitation Programme should not be included in assessments of the programme's effectiveness, because they did not have the 'full dose' (as does Robinson, 1995). This is to separate the 'completability' of a programme from its other aspects, when
'compleatability' is one of the most important features of an offender-programme. Many interventions which are not conceptualised in terms of 'dose', for example the Sherborne House programme, show similar lower reconvictions on the part of offenders who complete (see for example Maitland and Keegan, 1988). This could be represented as a treatment effect, given evidence that the course does have impact on offenders. Or it could be a result of offenders who have decided not to reoffend being more likely to complete the programme, and not offend in future regardless of programme content. This is discussed below in more detail.

One clear conclusion from this study is that intermediate outcome measures, which are rooted in attitudinal change, can be no substitute for behavioural measures such as reconviction as indicators of programme success. Perhaps this should not be surprising given the well known tenuousness, in psychological research, of links between attitudes and behaviour (see for example Jaspers, 1978.) Specifically, it is hard to see what value the use of scales such as CrimePics adds to a programme evaluation.

What is surprising is how little change occurs on the intermediate outcome measures. Loesel (1995) observes that the closer to treatment an outcome measure is, the larger the effect size that is observed. Whilst the current study only is able to analyse intermediate outcomes for programme completers, this group showed much larger differences between predicted and actual offending, than were observed for any intermediate outcome measures.

A useful approach may be to regard intermediate outcome measures as having intrinsic value, rather than view them as predictors of whether reconviction is likely to be reduced. Mair (1991) advocates an analogous approach of using secondary outcome measures to help take account of the multiple aims of sentences. For example a relevant outcome measure for community service could be number of hours work carried out for the benefit of the community. In the same way offender feedback which is positive about a programme would be seen as valuable evidence as to the quality of the programme. The quality of a programme is something which
is important in its self, and does not depend for significance on its relation with reconviction. Other intermediate outcome measures of intrinsic value might include the number of offenders achieving places in training, or education; achieving literacy; getting jobs; or maintaining relationships with their partners and children. These things are worthwhile regardless of how they relate to reconviction (see for example Williams and Webb, 1992 for more examples of this approach, which is explored in more detail in the next chapter.)

If intermediate outcome measures are used which do not have intrinsic value, they must have proven validity in being strong predictors of reoffending. None of the indicators used in the current study satisfy this condition.

Offenders' problems

Particularly interesting is the relationship of problems reported by offenders, and how these change and in turn relate to reconviction. Self-reported problems are one of the domains in which offenders report most change, in spite of the fact this is not a matter specifically addressed by Reasoning and Rehabilitation, or for the most part by other components of the Camberwell Probation Centre programme. It seems likely that the changes are in their subjective perceptions. An offender who reports having less of a 'problem with employment' at the end of the programme is more likely to regard unemployment as less of a problem, than to have found a job.

Raynor and Vanstone (1996) suggest that social problems faced by offenders may be so intractable as to mask any progress on the part of the individual offenders, and lead to reoffending. Raynor and Vanstone conclude of the STOP programme: 'the future success of intervention of this kind will be governed by the degree to which further offending is complemented by attempts to assist them with the problems that they encounter in their everyday lives in the real world' (page 282).

However the relationship between problems and reconviction in the current study
contradicts findings from the STOP programme, in which offenders who reported more problems were more likely to be reconvicted. In the current research the reverse was found to be the case. This also contradicts findings which show an association between social problems such as unemployment, and offending (see for example Field, 1997.) However these findings apply to general populations. The group of offenders targeted for the Camberwell Probation Centre is a highly selected group, with a long history of sometimes serious offending. This group is quite different from less committed offenders, and from the general population, so perhaps it should not be surprising if different relationships apply between problems and offending. It is also interesting to note that the research reported in Chapter 3, on the first ILPS Demonstration Unit reports no clear association between social problems and offending (see also Harraway et al, 1985.)

Are these findings consistent with the theories on which Reasoning and Rehabilitation is based? Attitudes, personality factors and theories of offending

The Reasoning and Rehabilitation Programme is based on a theory of criminal behaviour which is in the mainstream of psychological theories of offending. Farrington, 1991, gives an account of such psychological theories of criminal behaviour, which explain offending as resulting from the interaction of the person (in this case who has antisocial tendencies), and the environment, (which provides criminal opportunities.) The reasoning and rehabilitation programme rests on the theory that the antisocial tendency is the result of cognitive deficits which, crucially, can be made good via training in thinking skills. But are the findings from the current study consistent with Ross and Fabiano's (1985) theory? The following predictions were derived from this theory:

- many repeat offenders have cognitive deficits, relating to their social intelligence. This is entirely consistent with the findings of the current study.

- these deficits cause their offending. The current research did not look at
causation.

- **to be effective, programmes must target these offenders.** The Camberwell Probation Centre targeted appropriate offenders.

- **their cognitive deficits can be corrected through training.** The current research does not show any evidence that this happens. There is some evidence of attitudinal change, but less evidence of any change in underlying personality factors, or thinking styles.

- **correction of deficits will result in less offending.** There was evidence of reduction in offending, but no evidence this was linked to changes in attitudes or thinking styles.

**Conclusion**

The Reasoning and Rehabilitation Programme did appear to bring about modest reductions in offending, but the current study produced no evidence that this was achieved by correcting cognitive deficits. This must either be because the measures did not work, or because offenders really did not change. (If offenders did not change, this could be either because the theory underlying the programme is wrong, because the programme was not well rooted in the theory, or because the programme was badly implemented.) The measures used in this study seem more to tap personality and attitudes, than to be indicators of cognitive style. It may be that more behaviourally or ability-focused measures of social interaction would provide better measures of whether thinking **skills** had indeed been taught, which would influence offenders' behaviour.

Of the scales used in the current study, the CrimePics and Rosenbaum scales and the self-assessment of risk seem more likely to tap attitudinal change, than change in underlying personality factors. The Eysenck IVE, and the Levinson scale seem

230
more likely to tap underlying personality factors, being linked to theories of personality. So it may be significant that the largest movements are seen on the attitudinal measures, but that these are not associated with changes in reconviction. On the subject of personality factors in offending, Andrews (1995) comments:

For many writers, from Lombroso and Freud, through the Yale school, and up to Gottfredson and Hirschi... the essence of criminality has to do with a lack of self-control. To my knowledge, the field, after years of research and conceptualisation has not been able to reach a high consensus conceptualisation of what is meant by self-control, not does it have available a set of assessment instruments that actually demonstrate the dominance of self-control deficits in the analysis of criminal conduct. My prediction is that 'what people think' (antisocial attitudes) will prove more important than 'how people think'. Andrews, 1995, p51.

The Reasoning and Rehabilitation programme seeks to change 'how people think', but seems to be more influential in changing attitudes, (what people think).

However pace Andrews above, those who report attitudinal change are more likely to go on and reoffend. This weak link between attitude and behaviour should be no surprise, given the wealth of psychological research which shows that what people say, and what they do, are rarely more than weakly related. This may especially be the case when incentives may exist, as could apply in the current research, for offenders to say what they think the assessor wants to hear.

Were treatment to have taken place in accordance with Ross and Fabiano's theory, one would expect to see shifts in the intermediate outcome measures which were not found. In fact it is the offenders who change least who are also least likely to be reconvicted. This is evidence against the intended treatment effect causing reductions in offending. It suggests that some of the reduction in offending amongst completers is because of a selection effect. Perhaps the main reason CPC completers do so well is that the drop out rate is so high, leaving behind only those offenders who would not reoffend, regardless of treatment. This could be tested by comparing CPC completers with completers of other community sentences which have similar high drop-out rates, or by greatly reducing the drop out rate. For it to be said to be a treatment effect, there would need to be stronger evidence that
treatment has occurred. This is an argument for looking at all offenders who start a
programme. It is also evidence which contradicts Ross and Fabiano's theory of
thinking styles being linked to offending. Either the measures of thinking style do
not work (and given they show that appropriate offenders are targetted for the
programme, it seems they work to some extent), or positive changes in attitudes
and thinking are not linked to offending.

Ross and Fabiano should be commended for testing their theory in an intervention-
based study, in a field where many theories of offending go untested in the real
world. As well as looking at what reconviction outcomes were achieved, this study
sought to determine how they were achieved, in terms of changes in offenders
linked to a theory-based programme. As Glaser (1975) has observed, in the
evaluation field questions which start with 'how' whilst more problematic to answer,
have the promise of yielding more interesting results. However the inconclusive
results of the current study, with regard to the psychological mechanisms by which
results were achieved, should not be seen as surprising given the complexity of the
world in which offenders live. It is possible to control only a small part of this
world in intervention, and to encompass only part of it in evaluation.

This has been the final, and most ambitious in terms of its scope, of the set of
research case-studies. The next, and concluding, chapter draws together the findings
of each of the case-studies, so comparisons can be made between programmes
worked with similar offenders and similar aims. The next chapter also considers
how this research can contribute to an evidence-based for probation interventions
with offenders.
Chapter 7

Building the evidence-base: summary and conclusions

The research reported here was carried out with the aim of adding to an evidence-base which underpins the probation service's work with offenders, by contributing to knowledge about interventions which are likely to be effective in reducing offending. This work also had a strong methodological focus, aiming to help determine which evaluation methods and outcome indicators can best determine whether offender-interventions are effective in reducing offending.

The research has necessarily taken as its focus what Pitts (1992) claims is an emerging 'something works doctrine', and which Bottoms, (1995) terms 'the new orthodoxy'. According to this orthodoxy, (see for example Andrews, 1995) the 'principles of effective programming' are well understood, and community-based interventions with offenders, which focus on and challenge their offending, are likely to be effective. Pitts (1992) claims these assumption are based more on political considerations than on research evidence. This concluding chapter brings together the series of evaluative research studies, reported in previous chapters. Each of these was carried out with the aim of demonstrating effective practice. The current chapter looks at what together they tell us about what works in offender-interventions, and looks at what this means for the 'something works' orthodoxy. Each of the types of programme included in the research has been influential in the probation field, and all are broadly representative of many mainstream offender-interventions.

The research presented here is significant in scale, and has involved analysing in some detail the offending careers of over 700 offenders, (including all offenders who attended the programmes, and matched groups of offenders included in
comparison groups.)

This chapter draws together the main findings from the earlier research-based chapters, allowing findings to be compared. It evaluates how this new evidence fits with other research findings as to 'what works'; and sets out the most effective methodologies for monitoring work with offenders. Finally, this chapter looks at how evidence-based work with offenders can advance.

Why the research was carried out: pressure to demonstrate the effectiveness of probation

This research was carried out to determine whether key programmes for offenders, run by Inner London Probation Service, were successful in reducing offending. The research was done at a time when the probation service was regaining confidence that its interventions could be effective in combatting offending (see for example McIvor, 1990.) The work was given additional impetus by wider pressures on the probation service to demonstrate effectiveness in reducing offending. These included:

- the need in common with other public sector organisations to demonstrate cost-effectiveness and accountability (see Audit Commission, 1988);.

- an increased focus on monitoring publicly funded programmes in terms of the outcomes secured for service-users (see Williams and Webb, 1991, Department of Health, 1995);

- increased confidence that the service could be effective, and that 'something works' did not compromise the service's values (see Raynor and Vanstone, 1994).

- insecurity about possible reorganisation of a service unable to demonstrate
effectiveness (Sutton, 1994; Thomas, 1995) in a volatile environment with regard to central governmental criminal justice policy (see Nellis, 1995.)

The research: programmes included

The first stage of the research focused on the development of suitable outcome measures. Reconviction is the most obvious measure of reoffending, but is not as straightforward as it at first seems. The use of reconviction as a way of assessing the outcome of probation is reviewed by Wilkinson (1994a, 1994b, 1994c), focusing on the use of prediction techniques to make sense of reconviction rates. Having resolved how to monitor outcomes, a set of programme evaluations was carried out, which is summarised in the next section. Emerging findings were disseminated via internal and external publication, conference papers and workshops (references are given under appropriate headings below.)

The programmes

The current research was carried out on a set of five programmes run by Inner London Probation Service. Each of these programmes is in the mainstream of probation interventions (see NPRIE, 1994 for a catalogue of probation projects.) In fact each of the five projects is well known in probation circles, and has had some degree of wider influence on probation practice. Three of the five programmes were categorised by Ellis and Underdown (1998) as being 'well-evaluated programmes which provided some evidence of success' (page 108). This was in a survey of 210 probation programmes, carried out for HM Inspectorate of Probation, with the aim of finding examples of good practice in evaluation. Two of the programme-evaluations described in this thesis were not so categorised. However one, the Sherborne House evaluation, had not been completed at the time of Ellis and Underdown's (1998) survey. The other research evaluation, of the Demonstration Unit programme for offenders who take cars, showed no evidence that this
programme reduced offending.

The Demonstration Unit TDA and Burglary programmes were run by the first Inner London Probation Service Demonstration Unit, a pioneering unit which supervised offenders between 1982 and 1984. Though no longer functioning, this unit represented a significant innovation within probation practice, in many ways being a forerunner for what are now mainstream ways of working with offenders. The Demonstration Unit set out to increase sentencers' confidence in probation as a sentencing option for higher risk offenders, aiming to show that offenders who might otherwise have been sent to prison could be held safely in the community. This was to be done by concentrating supervision on offending, rather than on offenders' personal and social problems (at the time this was innovative.) The Demonstration Unit ran two entirely separate offender-programmes:

DUB: Demonstration Unit: Programme for Burglary Offenders: The programme for burglary offenders used a combination of one-to-one and group work, aimed at showing that burglary was not an effective way of making money, and that committing burglary was not purely a matter of impulse. Some social skills training techniques were used in doing this. (see Harraway et al, 1985, Wilkinson, 1995b, 1996, reported in Ellis and Underdown, 1998.)

DUT: Demonstration Unit: Programme for TDA Offenders: the second programme component of the Demonstration Unit was for generally younger offenders who had been convicted of Taking and Driving Away (TDA) a car. As with offenders convicted of burglary, the focus of supervision was on the specific type of offending, based this time on the notion that if these offenders could drive legitimately, they would have no need to continue to offend. So all attended a driver retraining course, and obtained provisional driving licences, and took driving lessons (see Harraway et al, 1985, 1986, Harraway, 1986, Wilkinson, 1995b, 1996).

The remaining three projects are long-established probation interventions. All have been running since the early 1980s, and remain operational, though their component
programmes have evolved. The research focused on offenders passing through these programmes mainly in the years 1990 to 1992.

**IMP:**  *Ilderton Motor Project:* this programme is for young offenders who take cars. The emphasis in this programme is on closely matching to the programme offenders who are most likely to benefit, that is those offenders whose offending in bound up with taking cars, and who are particularly interested in cars. The programme allows offenders to pursue their interest in cars in a structured and disciplined environment. They work on restoring cars for a minimum number of three two-hour sessions per week, with the incentive of racing them in competitions. (For reports on this evaluation see Wilkinson and Morgan, 1995, Wilkinson, 1997, reported in Ellis and Underdown, 1998.)

**SH:**  *Sherborne House Probation Centre:* this probation centre offers a programme for offenders aged between 17 and 21, who are at very high risk of reoffending, and of custody. This aims to involve offenders by offering a wide range of recreational and creative activities, offers information in key areas, connects offenders with resources to pursue their goals (e.g. for education and employment), and directly confronts their offending in a series of group-work exercises. (For a report on an earlier evaluation see Maitland and Keegan, 1988).

**CPC:**  *Camberwell Probation Centre:* this centre has existed from the early 1970s. However in 1991 its programme was reconfigured around the Reasoning and Rehabilitation Programme. This is a theory-driven, evidence-based programme, which has been widely implemented in North America (see for example Ross, 1995), and in England and Wales (see McGuire, 1993.) This programme is based on the idea that high-risk repeat offenders commonly have faulty thinking skills, being impulsive, not good at relating to others, and not given to thinking through the consequences of their actions. The programme consists of training exercises intended to rectify these deficits. (For reports on the evaluation see Wilkinson, 1993, Wilkinson, Ed. Stanley, 1995. reported in Ellis and Underdown, 1998.)
<table>
<thead>
<tr>
<th>Programme</th>
<th>Aims</th>
<th>Theory / Ideas</th>
<th>Methods</th>
<th>'Dose'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demo Unit: Burglary</td>
<td>- divert high risk; offenders from custody; - hold in comm.'ty without offending; - reduce burglary.</td>
<td>- work with offenders should focus on offending; - help offenders to understand negatives of offending; - show burglary is a bad way to make money; - show offending is not a matter of impulse, but can be avoided.</td>
<td>- groupwork; - social skills training; - individual sessions</td>
<td>2 to 3 cont-acts per mth for 6 mths; 1 contact per mths for the next 6 mths i.e. 'low dose'</td>
</tr>
<tr>
<td>Demo Unit: TDA</td>
<td>- divert from custody; - hold in community without offending; - reduce car-crime;</td>
<td>- work with offenders should focus on offending; - legitimatise offending behaviour and offending will cease.</td>
<td>- groupwork; - information and training in responsible driving; - driving lessons.</td>
<td>12 to 16 2 hour sessions over six months then contact once per mth for 6 months i.e. low dose</td>
</tr>
<tr>
<td>Ilderton Motor Project</td>
<td>- reduce car-crime amongst high-risk young offenders.</td>
<td>- offenders who are obsessed with cars can indulge this in a constructive way, learning responsibility.</td>
<td>- training in restoring old cars, in responsible environment; incentive of taking part in banger racing;</td>
<td>2 hrs, 3 eve's a week ave... mths</td>
</tr>
<tr>
<td>Sherborne House</td>
<td>- reduce offending - divert from custody - work with high-risk young offenders</td>
<td>- give offenders opportunities, enhance self-esteem, challenge rationale for offending - engage with offenders in any way possible</td>
<td>- groupwork exercises to challenge offending - chance for education and training - craft project - drama workshops - adventure activities - men and offending,</td>
<td>4.5 days week for 10 weeks</td>
</tr>
<tr>
<td>Camberwell Probation Centre</td>
<td>- reduce offending - divert from custody - work with high risk offenders.</td>
<td>- offenders offend because of wrong thinking styles which can be corrected by training.</td>
<td>- training in thinking skills; - set of group-work exercises</td>
<td>35 2-hour R&amp;R session in a 12-week prog.</td>
</tr>
</tbody>
</table>
Table 7.1 offers a summary of the programmes, setting out when the research was carried out, programme aims, the ideas or theories on which they were based, the main intervention methods which they used; and the amount of contact with offenders was involved in delivering the intervention.

**Evaluation methods**

As well as the actual results of the research being important, a significant strand running through this work was how results were to be obtained, which involved developing and refining ways to gauge the effectiveness of offender-programmes.

**Offending outcomes**

All five programme-evaluations were designed as offending follow-up studies, each using a comparison group to help make sense of offending outcomes. In most cases each programme's comparison group was made up of offenders referred, assessed and found suitable for the programme, but then sentenced to other community or custodial penalties. The Ilderton Motor Project's comparison group was a matched group selected specifically for the purpose of comparing patterns of offending.

Four of the studies used information about reconvictions as their main outcome indicator. This was obtained from the Home Office National Identification Bureau. Reconviction may not be an ideal measure of offending given only a small minority of offences committed result in a conviction, (see Barklay, 1993). Reconviction may also derive to an extent from the performance of criminal justice agencies rather than the behaviour of offenders. Reconviction also includes clouding factors such as pseudo-reconvictions (see Lloyd et al, 1994.) Nevertheless, reconviction correlates strongly with self-reported offending (see for example Bottoms, 1995, Tarling, 1993), and measures can be taken to eliminate pseudo-reconvictions from
follow-up data. Despite its drawbacks, reconviction remains a robust, useful and above all obtainable measure of offending.

The Ilderton Motor Project study used information about arrests, obtained from the Metropolitan Police. In this case the main outcome indicators were derived from the occasion on which offences were committed, where there was no later court finding of not-guilty.

In all the studies, number, type and seriousness of reconviction were collated, as relevant programme outcomes, as well as the proportion of offenders reconvicted within given time periods.

Making sense of information about reconvictions

Comparison groups were used to help make sense of reconviction outcomes, using a quasi-experimental design. Allocation to conditions was non-random. This falls short of the ideal of random allocation, in that it is always open to question whether groups are truly equivalent (see e.g. McDonald et al, 1992.) Because allocation depends usually on the decision of the sentencing court, the need is to demonstrate that at time of sentence, the intervention and comparison groups are effectively drawn from the same population of offenders, in terms of criminogenic factors. This is so that the offending of the comparison group can be taken as indicating how the intervention group would have behaved had the intervention not existed. Reconviction predictors were used to assess whether intervention and comparison groups were at similar risk of offending. These were the National Risk of Reconviction Predictor, (see Copas et al, 1994; and its successor the Offender Group Reconviction Predictor, see Copas and Marshall, 1998.)

Quasi-experimental techniques can be uninformative as to how programme results are achieved (see e.g. Pawson and Tilley, 1994.) Nevertheless, comparison groups are relatively straightforward to set up, for criminal justice programmes which
already exist. They were the best design available, when this research was planned, for indicating within a reasonable time-frame whether programmes were effective in reducing reconviction.

A problem with comparing intervention and comparison groups' offending arises from the fact a substantial proportion of offenders in each comparison group were sentenced to custody. This means that follow-up periods from time of sentence include times when a substantial proportion of the comparison group were not available to be reconvicted. Comparing reconvictions from time of sentence in this way indicates the degree to which the community has been protected from offending, but it is less useful in showing the impact of treatment on offenders. All five component studies look at offending from time of sentence, the Sherborne House and Camberwell Probation Centre studies also look at offending from time of release for comparison group members sentenced to custody.

Another issue is that of programme completion. Some researchers have argued (e.g. Robinson 1995), that offenders who leave a programme early have not received the full intervention, and so should be excluded from reconviction follow-up studies. Maitland and Keegan, (1988) argue this is particularly the case for offenders who drop out in the early weeks of an intervention. But it is only possible to distinguish completers post hoc. If only half the offenders starting a programme are likely to complete, this is an important fact about the intervention. It may also be that offenders who complete a programme self-select, by not offending during the programme and by being more committed to ceasing offending, regardless of programme content. Whilst the current studies did distinguish separate outcomes for programme completers and drop-outs, the most important results are those for all who start a programme, and it is only these which are brought together in this summary.

Other outcomes

For most of these programmes reducing offending was only one aim. Other
important aims included targeting offenders for the community programmes who might otherwise have received custodial sentences and who were at high risk of reoffending. In addition each programme had its own aims relating to the particular nature of the work done with offenders.

Reducing offending is an ultimate goal of offender-programmes, but there may be steps *en route* to this aim, which can be termed 'intermediate outcome measures' (see Fitz-Gibbon and Morris, 1987). This also relates to the concept of the 'dynamic reconviction predictor' (Gerandreau *et al.*, 1995b), a factor linked to offending, which unlike gender, or offending history, can be changed by intervention. The evaluation of the Camberwell Probation Centre programme included a significant number of intermediate outcome measures / dynamic reconviction predictors. This programme set out to change offenders' thinking styles, so they would become more self-controlled, less impulsive, and more pro-social in their attitudes. A range of attitudinal and psychometric measures were selected, linked to the theory underlying the programme, to assess whether or not these changes had occurred. Most significantly, one of the scales used, 'CrimePics' (Frude *et al.*, 1990), has become regarded as an 'industry standard' (Underdown, 1998) in probation, for assessing whether offender-programmes have changed attitudes to offending.

**Results and discussion**

This section summarises the results of each of the five programmes, allowing results to be compared. Distinguishing more and less effective methods of working necessitates making comparisons. Comparing programmes may also allow common aspects of effective programmes to be identified (what Antonowicz and Ross, 1994, term the 'essential components shared by effective programmes'). However there are dangers in making such comparisons. These programmes target different group of offenders, so like may not be being compared with like. Each of the
research projects summarised here bears its own distinct messages for work with offenders. It would be wrong for this section to be seen purely as a league table for identifying the 'best' programmes.

This section can provide only a brief summary of some of the main findings. Full results are given in each of the earlier research-based chapters. In the interests of space, summary tables only are shown here. Table 7.2 shows the origin of the base figures for all subsequent tables.

<table>
<thead>
<tr>
<th>Number of offenders...</th>
<th>1 year</th>
<th>2 years</th>
<th>1 year</th>
<th>2 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attended programme</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUB</td>
<td>51</td>
<td>51</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>DUT</td>
<td>69</td>
<td>69</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>IMP</td>
<td>34</td>
<td>26</td>
<td>40</td>
<td>35</td>
</tr>
<tr>
<td>SH</td>
<td>122</td>
<td>113</td>
<td>104</td>
<td>90</td>
</tr>
<tr>
<td>CPC*</td>
<td>100</td>
<td>100</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>Total</td>
<td>376</td>
<td>359</td>
<td>333</td>
<td>313</td>
</tr>
</tbody>
</table>

*Analysis from point of sentence only

A total of 709 offenders were followed-up for at least a year after sentence, with 646 offenders being followed up for at least 2 years.

Targeting aims and outcomes

All the programmes targeted offenders at high risk of reoffending. Table 7.3 shows the programmes were successful in doing this, all programmes targeted offenders at higher risk than the average for conditional probation orders, which is the prediction that 60 percent of offenders will be reconvicted within 2 years of start of sentence (see Lloyd et al, 1995.) It also seems likely that all programmes were successful in diverting offenders from custodial sentences, considering the nature of offences committed, the high level of previous convictions (see Table 7.3), and the
high proportion of offenders unsuccessfully referred to the programmes who were sentenced to custody. This seems particularly to have been the case for the Sherborne House programme.

<table>
<thead>
<tr>
<th>Previous convictions</th>
<th>Age</th>
<th>Risk of reconviction</th>
</tr>
</thead>
<tbody>
<tr>
<td>intv.</td>
<td>comp.</td>
<td>intv.</td>
</tr>
<tr>
<td>Demo Burglary</td>
<td>6.5</td>
<td>6.8</td>
</tr>
<tr>
<td>Demo TDA</td>
<td>3.0</td>
<td>4.9</td>
</tr>
<tr>
<td>Ildefont</td>
<td>6.6</td>
<td>5.4</td>
</tr>
<tr>
<td>Sherborne</td>
<td>7.1</td>
<td>5.4</td>
</tr>
<tr>
<td>CPC</td>
<td>15.4</td>
<td>14.2</td>
</tr>
</tbody>
</table>

All comparison groups can be deemed sufficiently similar at time of sentence, in terms of factors linked to offending, for their subsequent offending to be taken as indicating how the intervention group would have behaved, had the intervention not taken place. In fact most of the intervention groups were at higher risk of reconviction than the comparison groups, markedly (but not statistically significantly) so in the case of the two probation centre programmes: Sherborne House and Camberwell Probation Centre.

Reconviction outcomes

Tables 7.4, 7.5 and 7.6 summarise reconviction outcomes, looking at proportion convicted (or in the case of Ilderton arrested for an offence and not subsequently found not guilty), over 1 and 2 year follow-up periods; at how this compares with what would be predicted on the basis of offenders' age, gender and previous convictions; and at custodial sentences over a 2 year follow-up.

These findings, in terms of the effectiveness of each programme, are broadly similar regardless of the precise reconviction outcome used, be it follow-up period, comparator, or seriousness of reconviction. This is to be expected on the basis of
Table 7.4. Reconviction outcomes

<table>
<thead>
<tr>
<th></th>
<th>percent of offenders reconvicted within a one or two year follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>attended programme</td>
</tr>
<tr>
<td></td>
<td>1 year 2 years</td>
</tr>
<tr>
<td>DUB</td>
<td>29 57</td>
</tr>
<tr>
<td>DUT</td>
<td>48 68</td>
</tr>
<tr>
<td>IMP\textsuperscript{*}</td>
<td>53 65</td>
</tr>
<tr>
<td>SH</td>
<td>63 73</td>
</tr>
<tr>
<td>CPC\textsuperscript{**}</td>
<td>45 67</td>
</tr>
</tbody>
</table>

\textsuperscript{*} Figures for Ilderton Motor Project are for 'offending occasions' rather than for reconvictions. Offending occasions are essentially arrests, which were not later followed by a not-guilty finding.

\textsuperscript{**} Figures for Camberwell Probation Centre omit known pseudo-reconvictions.

1. Significant difference between programme and comparison groups in the direction of less offending on the part of the programme group equivalent to a 48 percent reduction, \( p<0.01 \) in a 2-tailed \( t \)-test.

2. Significant difference between programme and comparison groups in the direction of less offending on the part of the programme group equivalent to a 34 percent reduction, \( p<0.05 \) in a 2-tailed \( t \)-test.

3. Significant difference between programme and comparison groups in the direction of less offending on the part of the programme group equivalent to a 29 percent reduction, \( p<0.05 \) in a 2-tailed \( t \)-test.

Table 7.5. Percent reconvicted within 2 years: comparing predicted and actual rates.

<table>
<thead>
<tr>
<th></th>
<th>predicted</th>
<th>actual</th>
<th>what pct of predicted</th>
<th>percent offending less than predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>DUB (n=51)</td>
<td>68</td>
<td>57</td>
<td>84</td>
<td>16</td>
</tr>
<tr>
<td>DUT (n=69)</td>
<td>64</td>
<td>68</td>
<td>106</td>
<td>6 more</td>
</tr>
<tr>
<td>IMP\textsuperscript{*} (n=26)</td>
<td>75</td>
<td>65</td>
<td>87</td>
<td>13</td>
</tr>
<tr>
<td>SH (n=113)</td>
<td>79</td>
<td>73</td>
<td>92</td>
<td>8</td>
</tr>
<tr>
<td>CPC (n=100)</td>
<td>70</td>
<td>67</td>
<td>96</td>
<td>4</td>
</tr>
</tbody>
</table>

\textsuperscript{*} The figure for Ilderton Motor Project is for offending occasion, rather than reconviction, so the predicted figure is included to aid comparisons across projects only.

Table 7.6. Percent reconvicted and sentenced to custody within 2 years

<table>
<thead>
<tr>
<th>Programme</th>
<th>Comparison (custodials from sentence)</th>
<th>Comparison (custodials from release)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DUB</td>
<td>39</td>
<td>-</td>
</tr>
<tr>
<td>DUT</td>
<td>30</td>
<td>-</td>
</tr>
<tr>
<td>IMP</td>
<td>15</td>
<td>46\textsuperscript{**} -</td>
</tr>
<tr>
<td></td>
<td>( p&lt;0.01 ) (matched groups ( t )-test, 2 tailed.)</td>
<td></td>
</tr>
<tr>
<td>SH</td>
<td>55</td>
<td>33</td>
</tr>
<tr>
<td>CPC</td>
<td>28</td>
<td>31</td>
</tr>
</tbody>
</table>

245
large-scale research into criminal careers, which shows that for the most part, offending seems to be unidimensional in terms of frequency and seriousness (see Farringdon, 1992.)

Table 7.7 ranks the programmes in terms of which are most and least effective, according to the size of the difference between the programme group and the comparator for each outcome indicator. Notwithstanding warnings given above about the dangers of comparing dissimilar programmes, there is some value in this. All were probation programmes working largely with young male offenders, with several previous convictions, mainly for acquisative crime. In particular the Ilderton Motor Project, and Demonstration Unit TDA projects were focused on similar offenders with similar objectives, but with markedly different outcomes.

<table>
<thead>
<tr>
<th>ranking by effectiveness</th>
<th>by comp. group proportion recon.</th>
<th>by prediction at 2 years</th>
<th>by custody at 2 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (most effective)</td>
<td>IMP</td>
<td>DUB</td>
<td>IMP</td>
</tr>
<tr>
<td>2</td>
<td>DUB</td>
<td>IMP</td>
<td>DUB</td>
</tr>
<tr>
<td>3</td>
<td>CPC</td>
<td>SH</td>
<td>CPC</td>
</tr>
<tr>
<td>4</td>
<td>SH</td>
<td>CPC</td>
<td>DUT</td>
</tr>
<tr>
<td>5</td>
<td>DUT</td>
<td>SH</td>
<td></td>
</tr>
</tbody>
</table>

What seems clear is that Ilderton Motor Project, the Demonstration Unit Burglary and the Camberwell Probation Centre programmes were effective. The status of Sherborne House is less clear. The group who attended Sherborne were at the very highest risk of reconviction, markedly (though not statistically significantly) more so than the comparison group. The Sherborne Group go on to offend more than the comparison group, even when time spent by the comparison group sentenced to custody is eliminated from the analysis. However when the Sherborne Group are compared to a group of similar age and gender, with a similar offending history (that is by using the OGRS predictor, Copas and Marshall, 1998), the Sherborne Group in fact offend less than predicted.
Only the Demonstration Unit TDA project seems to have been an unequivocal failure in reducing reconviction. Yet paradoxically this project was the only one to have been the subject of a substantive research-based publication, prior to the current research initiative. This paper (Harraway et al, 1985) reported the programme to have been a success, in terms of increasing the use of the probation order for motor-vehicle related offenders; in terms of offenders' interest in and engagement with the programme; and in terms of offenders completing their probation orders without breach action being taken.

**Offence specific outcomes**

The two most effective programmes, (and the least effective programme) worked with offenders convicted of a specific type of crime, and focused intervention on that type of offending. An obvious question is whether offending outcomes reflected the offence-specificity of the intervention.

Table 7.8 shows that Ilderton offenders were more committed to motor vehicle offending than were the Demonstration Unit offenders. The Ilderton group had committed on average 4.9 motor-vehicle related offences each, in their prior offending career, 74 percent of their overall offending being motor vehicle related. However the Demonstration Unit offenders showed less evidence of specialisation. On average they had been convicted of one earlier motor-vehicle related offence, from a total of 3 previous convictions.

Following sentence the proportion of motor vehicle related offences committed by the Ilderton group falls significantly compared to the comparison group. The Demonstration Unit TDA offenders show far less evidence of change. This seems to show that Ilderton's was a successful offence specific intervention. However the Demonstration Unit TDA programme seemed not to be selecting specialist car-crime offenders. Perhaps it should not be surprising that an intervention focused not on offending related to cars, but on driving, should prove ineffective with what was
essentially a mixed group of young offenders. In contrast the Ilderton offenders show significant evidence of specialisation, and evidence that reduction occurs in motor-vehicle related offending which has a significant impact on the total amount of crime committed by this group.

Table 7.8. Offence specificity

<table>
<thead>
<tr>
<th></th>
<th>Demo Unit TDA</th>
<th></th>
<th>Ilderton Motor Project</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DUT</td>
<td>Comp.</td>
<td>IMP</td>
</tr>
<tr>
<td>Average no. previous...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>convictions</td>
<td>3.0</td>
<td>4.9</td>
<td>6.6</td>
</tr>
<tr>
<td>m/v related</td>
<td>1.0</td>
<td>1.6</td>
<td>4.9</td>
</tr>
<tr>
<td>percent m/v related</td>
<td>33</td>
<td>33</td>
<td>74</td>
</tr>
<tr>
<td>in the year after</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>convictions</td>
<td>0.5</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>m/v related</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>percent m/v related</td>
<td>24</td>
<td>18</td>
<td>29</td>
</tr>
</tbody>
</table>

* Statistically significant at the p<0.05 level in a two-tailed (matched groups t-test).

"m-v related" refers to any motor vehicle related offence; percent m-v related refers to the proportion of all offences committed which involved motor vehicles.

An important finding from the Demonstration Unit TDA project research, is that offenders who specialise in car-crime are likely to be no more than a subset of the group of offenders convicted of TDA at any one time. This is consistent with research into criminal careers, which shows that with the exception of sex-offenders and offenders who commit fraud over a long period, specialisation in one type of crime is uncommon, (see for example Farrington, 1992; Stander et al 1989, Kempf, 1986, Klein, 1984). There is no evidence that a majority of the offenders who are convicted of car crime are specialists in this type of offending.

Other aims and outcomes

The current research did not set out to evaluate how well the programmes achieved.
other aims, of securing community integration, and providing practical help. However it was reported (by Harraway et al, 1985) that offenders attending the Demonstration Unit programmes showed some improvement in what the researchers termed 'social variables', referring to stability of housing and employment, and positive leisure activities. But these observations were made without reference to a comparison group. Interestingly Harraway et al (1985) reported that improvement in these social variables was not linked to reduction in reconviction.

Maitland and Keegan (1988) reported, on an individual case-study basis, how the Sherborne House programme changes how offenders perceive their social world, the main impact being that after the programme offenders see the world as more complicated, but are more positive about their place in it. However offenders who showed positive change in this regard were no less likely to be reconvicted.

The study of Camberwell Probation Centre reported here looked at offenders' self-reported problems in a range of social and personal areas, including with drugs, relationships, employment, and with self-esteem. Most offenders who completed the programme showed overall improvement in problems, the biggest average gains being with 'lack of confidence' and with 'problem with employment'. However a limitation of this study was that there was no objective validation of the self-reported information. In common with Harraway (1985), there was no clear relationship between problem reduction, where this occurred, and reconviction.

**Intermediate Outcomes and Reconviction**

Table 7.9 reprises a key finding of the Camberwell Probation Centre study, with regard to intermediate outcome measures / dynamic reconviction predictors. This table shows that offenders who showed pro-social changes were actually slightly more likely to be reconvicted within 2 years (though they were actually less likely to be sentenced to custody on reconviction.) Because of the small numbers it is not possible to generalise from either of these findings. But given similar findings for
the other intermediate outcome measures used in this study, it seems clear that for
this group of offenders at least, there is no straightforward correlation between
positive change on measures of impulsiveness, self-control, attitudes to offending, or problems; and reconviction. Given the common finding in social psychology, of
the tenuousness of links between attitudes and behaviour (see for example Jaspars,
1978), perhaps this lack of correlation should not be surprising.

Table 7.9. Camberwell Probation Centre: attitudinal change and reconviction

<table>
<thead>
<tr>
<th>Shift in attitude to offending (i.e. change in CrimePics score)</th>
<th>negative or no change (n=10)</th>
<th>positive (pro-social change (n=25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>reconvicted within 1 year (percent)</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>reconvicted within 2 years (percent)</td>
<td>50</td>
<td>60</td>
</tr>
</tbody>
</table>

**Reviewing findings in the light of evidence about 'what works'**

The more closely the outcome data from each study is scrutinised, the more
pronounced are the reductions observed in offending outcomes. Impact on offending
becomes clearer once corrections have been made for inadequacies in the raw
reconviction data, and for pre-existing differences between the intervention and
comparison groups. The following ways of treating the reconviction data from the
four effective programmes all tend to reveal more favourable reconviction outcomes
for the intervention group:

- correcting for pseudo-reconvictions (i.e. convictions for offences committed
  before the start of intervention);
- getting outcome data which is linked more closely to offending (i.e. by
  looking at data relating to arrest);
- taking into account risk of reconviction;
- looking at more serious reconvictions, (i.e. those sentenced by way of custody);
- looking at reconvictions for offences committed within a year of the start of intervention.

Unfortunately a significant degree of analysis is required to do most of the above, to correct for differences between intervention and comparison groups, resulting from the fact allocation to conditions is non-random. This tends to make findings less immediately impressive, and is one reason why for many reviewers (see for example Losel, 1995, McDonald et al, 1992) quasi-experimental methods will ultimately never yield as satisfactory findings as true random-allocation experiments.

This section attempts to assess why particular programme outcomes occurred. This must be done tentatively because of the dangers of rationalising facts to fit theoretical models *post hoc*. Findings can be explained away in a manner not allowed by prospective hypothesis-testing.

An idea advanced by proponents of the 'what works' orthodoxy, is that there now exist a set of principles rooted in theory (see e.g. Andrews, 1996) and in evidence (see e.g. McGuire, 1995), and that programmes which incorporate these principles are more likely to be effective in reducing offending (see McGuire *et al*, 1995). Table 7.10 draws on Table 1.1 from Chapter 1 of the current work, to set out a checklist of what a number of reviewers of the field have reported to be the characteristics of effective programmes (though as Losel, 1995, observes, there is only limited agreement between reviewers as to these characteristics.) This makes it possible to assess the extent to which the five programmes evaluated in the current research incorporate the characteristics of effective programmes. The issue this table seeks to address is that of how differences between programmes relate to their effectiveness, and in turn how this matches what could be predicted from the 'principles of what works'.
Table 7.10. How do the programmes match the 'principles of What Works'? (y: programme incorporates this principle; Y: programme strongly incorporates this principle.)

<table>
<thead>
<tr>
<th>'What works' principle*</th>
<th>Projects**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DU:T</td>
</tr>
<tr>
<td><strong>number of elements</strong></td>
<td>(i.e. no. of y's)</td>
</tr>
<tr>
<td><strong>ranking in terms of effectiveness</strong></td>
<td>5</td>
</tr>
<tr>
<td><strong>effective</strong>*?**</td>
<td>no</td>
</tr>
<tr>
<td><strong>selection of offenders</strong></td>
<td>-</td>
</tr>
<tr>
<td>- high risk offenders (for high intensity serv.s)</td>
<td>-</td>
</tr>
<tr>
<td><strong>methods / content</strong></td>
<td>-</td>
</tr>
<tr>
<td>- behavioural / cognitive behavioural / social cog.</td>
<td>-</td>
</tr>
<tr>
<td>- founded on a theoretical / conceptual model</td>
<td>y</td>
</tr>
<tr>
<td>- training / skills orientated</td>
<td>y</td>
</tr>
<tr>
<td>- multimodal, multifaceted, structured</td>
<td>y</td>
</tr>
<tr>
<td>- matched to offenders' learning styles</td>
<td>y</td>
</tr>
<tr>
<td>- focus on criminogenic needs</td>
<td>-</td>
</tr>
<tr>
<td>- improve non-criminal opportunities and facilitate community reintegration</td>
<td>y</td>
</tr>
<tr>
<td>- responsive to circumstantial and personal needs</td>
<td>y</td>
</tr>
<tr>
<td><strong>dose / amount</strong></td>
<td>-</td>
</tr>
<tr>
<td>- intensive, short-term, time-limited</td>
<td>-</td>
</tr>
<tr>
<td>- 2+ contacts/week for 26+ weeks;</td>
<td>-</td>
</tr>
<tr>
<td>&gt;40% of offenders' time</td>
<td>-</td>
</tr>
<tr>
<td><strong>implementation / enforcement</strong></td>
<td>-</td>
</tr>
<tr>
<td>- well monitored, supervised and implemented</td>
<td>y</td>
</tr>
<tr>
<td>- clear treatment plan delivered to each client</td>
<td>-</td>
</tr>
<tr>
<td>- planned, with clear goals</td>
<td>y</td>
</tr>
<tr>
<td>- apply restrictions and control only when justified by risk of reoffending, and where services commensurate with need are also applied;</td>
<td>-</td>
</tr>
<tr>
<td>- enforcement of programme contingencies / behavioural strategies firm but fair</td>
<td>y</td>
</tr>
<tr>
<td>- trained therapists relate to offenders in sensitive and constructive ways.</td>
<td>-</td>
</tr>
<tr>
<td>- disruption of the criminal network by programme structure and activities.</td>
<td>-</td>
</tr>
<tr>
<td><strong>context</strong></td>
<td>-</td>
</tr>
<tr>
<td>- community-based</td>
<td>y</td>
</tr>
<tr>
<td>- allow professional override</td>
<td>y</td>
</tr>
<tr>
<td>- do not indirectly increase the risk of offenders going into custody.</td>
<td>y</td>
</tr>
<tr>
<td>- high level of advocacy and brokerage</td>
<td>y</td>
</tr>
<tr>
<td>- researcher involved in delivering service.</td>
<td>y</td>
</tr>
</tbody>
</table>

* The summary in this table of the findings of research reviews derives from the following meta-analyses and narrative syntheses of the 'what works' literature: Andrews et al, 1990; Antonowicz and Ross, 1994; Bonta et al 1994; Gerandreau et al, 1994; Lipsey, 1990; Loxel, 1995; McGuire, 1994, 1995; McVor, 1990; Ross and Fabiano, 1987; Sheldon, 1994. See Chapter 1 for an explanation of the derivation of these principles from a sometimes contradictory set of reviews.

** Project key: DUT: Demonstration Unit, TDA offenders; DUB: Demonstration Unit, Burglary offenders; IMP: Ildefon Motor Project; SH: Sherborne House CPC: Camberwell Probation Centre

*** By effective, what is meant that this project shows strong evidence of reducing recidivism, compared to prediction, or to comparison group.
These programmes share a considerable number of features in common. This is to be expected given that all took place within the context of a conditional probation order. This means that all programmes were run in the community, by the probation service and mainly by social-work trained probation officers. All programmes worked with relatively high risk offenders, convicted mainly of acquisitive offences.

Some key questions, reviewing the success of programmes against the extent to which they incorporate 'what works' features, are as follows:

- **why did the Demonstration Unit TDA programme do so badly?** The outcome is puzzling in that this programme was clearly structured, rooted in a theory of why young people offend, well-run, and well-matched to offenders' learning styles.

- **why did the Ilderton Motor Project do so well?** This outcome might not be expected, given that this project has no explicit theoretical underpinning, offers minimal social work input, and includes relatively few of what reviewers of the what works literature have identified as the essential components of rehabilitation programmes (see for example Antonowicz and Ross, 1994.)

- **why did the Demonstration Unit programme for burglars do so well?** As a very low intensity programme working with high risk offenders, and founded with unclear aims, there was no reason to suppose this programme would turn out to be so apparently effective in reducing reconviction.

- **why did the Camberwell Probation Centre programme not do better?** Though this programme did seem to reduce offending, it might have been expected to do this more so, given it was based round the Reasoning and Rehabilitation Programme, for which strong claims of effectiveness have been made (see for example Ross, 1990.) This is an explicitly evidence-based and theory-based programme, which combines many features of effective programmes.
- why did Sherborne House not do better? This too is an unexpected finding. As an intensive, time-limited, highly structured, multi-modal programme, targeted on high risk offenders, and well matched to their learning styles, it might have been expected to be highly effective in reducing offending.

It may well be that the Demonstration Unit TDA programme did not reduce reconviction because the theory on which it was based was wrong. This was that all offenders convicted of car crime are 'car freaks', who if they could drive legitimately would have no need to offend. However car crime made up a relatively small proportion of the offences committed by the offenders targeted by this programme, so making these people legitimate drivers could be expected to only have a very minimal impact on their offending. A second possible reason is that this programme failed to help many offenders to become legitimate drivers. Given that under 15 percent were found to have obtained driving licences, this raises the question of what the remainder were likely to do.

In contrast the Ilderton Motor Project targeted offenders at very high risk of reconviction, and three quarters of whose previous offending was car related. This programme offered genuine opportunities to work on and to drive cars, being matched to learning styles and addressing criminogenic need. The programme occupied a significant amount of offenders' time, likely to disrupt criminal networks and offering constructive activities at times of day when otherwise these offenders might have been involved in car crime.

It is harder to explain the success of the Demonstration Unit burglary programme in terms of the principles of effective programming. This programme utilised some social skills training methods, and addressed offending in what seems to have been an appropriate way. But otherwise this intervention used fewer of the principles of effective programming than the less successful probation centre programmes.

The Camberwell Probation Centre programme, whilst modestly successful seems somehow less than the sum of its parts. It is possible the centre's relatively low
success rate, compared say to Ilderton Motor Project, or to the Demonstration Unit burglary project, is because over the period studied, the programme had a very high drop-out rate of 60 percent. This later stabilised at a more acceptable 40 percent, so perhaps a further study might show still more evidence of effectiveness. However offenders who completed the programme and who did not reoffend seemed to be the ones who were not changed by the programme, according to the intermediate outcome measures of attitude to crime, of impulsiveness and of self-control.

Nevertheless, these results for Camberwell Probation Centre are actually somewhat more positive than those reported by Robinson (1995) or by Raynor and Vanstone (1996), for the Reasoning and Rehabilitation Programme run in other settings.

It is also difficult to explain the relative lack of impact on offending of the Sherborne House programme in terms of the principles of effective programming, because Sherborne House appeared to combine many of these principles (though the programme did focus to a large extent on improving self esteem, and on other non-criminogenic factors.) However an explanation, rooted in the mechanics of the criminal justice system, can be advanced as follows: offenders who attended Sherborne House had committed serious offences and were at the highest risk of reconviction. By sentencing them to Sherborne House the court was 'giving them a chance' when otherwise custody might have been expected. The comparison group was at lower risk of reconviction, having been proposed a sentence including Sherborne House (albeit unsuccessfully, with in most cases a custodial sentence being passed instead), more because of the seriousness of their current offence, than for their risk of reconviction. So it should be no surprise the comparison group reoffend less. The Sherborne group actually offended at a markedly lower rate than predicted by their age and previous convictions. But when reconvicted, having already been given one chance by the court, the Sherborne group were more likely to be sentenced to custody than might be warranted by the seriousness of offence alone. In other words the Sherborne House programme might have worked, but a quasi-experiment with reconviction as the sole outcome indicator might simply fail to show this.
It seems the most likely reason the Demo Unit programme for TDA offenders failed to reduce reconviction, is that the programme failed to identify and address the appropriate criminogenic needs. Offenders on this programme were not specialised TDA offenders, and so focusing on legitimising their driving was never likely to have much impact on their offending. In contrast the Ilderton programme members clearly were specialised in TDA, and very interested in cars. Diverting this interest in a constructive direction, in a disciplined and structured environment, which also meant they spent a substantial proportion of 'at risk' time involved in purposeful activity did focus on criminogenic need, and did have an impact on offending.

The Demonstration Unit burglary programme was instigated without a clear idea as to how to work with offenders. This meant working with the offenders, and exploring their reasons for offending, challenging their reasoning, and using social-skills training methods found to be effective in other settings (see Priestly and McGuire, 1985) to teach skills for avoiding further offending. It may be that this approach of evolving a programme, and developing methods of working which seem to match offenders' needs may have ensured the programme was focused on true criminogenic need.

The Sherborne House programme placed a lot of emphasis on self-esteem, and providing opportunities for legitimate activities. Whilst these things are important, lack of self-esteem and lack of opportunity are not themselves strongly predictive of offending.

**Conclusions**

This research was done at a time when the probation service's interest in reducing offending had been reawakened by an combination of an increasing focus on the outcome of social work interventions; a gathering body of positive research
findings; and increasing pressure to justify the service's role. Three main strands of conclusion accordingly can be drawn. The first strand of conclusions relates to the research findings summarised above, and the additional light they may help to shed on 'what works', in terms of reducing offending. Secondly is a strand of methodological issues, which can be placed under the banner: 'how do we know what works, and how do we know if it's working'. The main issues here are how best to research and monitor work with offenders, to identify effective methods for reducing offending, and assess how well these methods are implemented and working in practice. The final strand of conclusions relates to the role of an evidence-base, and how research evidence can translate into policy and practice in work with offenders.

What works?

The positive findings of the current research are consistent with those of recent reviews of research in this area (see Chapter 1, and for example Losel, 1995), in that overall they confirm a modest positive effect of intervention. But as was shown above, the pattern of results could not have been predicted by applying what various reviewers of the literature claim are the principles of effective intervention. Reviewing this literature Losel (1993) concludes that:

'there are encouraging effect sizes for the most adequate types of programme (e.g. specific cognitive-behavioural, multimodal, need and responsivity orientated approaches.) Even so many differential effects drawn out by meta-analysis are only low-to-moderate and not yet consistent... We are still far from a conclusive answer with respect to the question of what works, with whom and under what conditions.' (Page 431.)

There is more agreement amongst reviewers as to which intervention methods do not work. This was discussed in Chapter 1 (see also McGuire, 1995); briefly ineffective interventions include those based on punishment and deterrence, or
intensive supervision which does not attempt to change offenders; unfocussed and psychodynamic counselling, not directed to behaviour modification, and casework; and disposals which fail to challenge offending, or the factors which lead to offending.

It can be particularly difficult to interpret negative or equivocal findings, because there are many reasons why an intervention may not work. Each programme included in the current research starts with an idea or theory about offending. Each programme, operating in the context of a community sentence, then applies an intervention intended to act on causes of offending, thereby reducing further offending and reconviction. Negative or equivocal findings as to reconviction outcomes can indicate either the theory is wrong; the programme-context is not conducive; the method of intervention is inappropriate, or that the intervention is badly applied.

Information about processes or intermediate outcomes may help to pinpoint the weaker links in the chain. For example if desired intermediate outcomes are achieved, but prove to be unrelated to offending, this might indicate that the theory on which a programme is based is wrong. This situation applies to the Camberwell Probation Centre study. Here though results overall were modestly positive, offenders who reported positive change in intermediate outcome indicators were actually more likely to be reconvicted. This could indicate that the theory on which the intervention was based was wrong, and something else was causing lower reconvictions. Or it could mean that the method for measuring intermediate outcomes was wrong, or that the programme was badly implemented. Rosenbaum (1986) in this regard refers to theory failure, methodological failure, or programme failure.

Positive findings are much more straightforward to interpret, especially when dynamic predictors of offending indicate the programme is achieving desired intermediate outcomes, linked to the theory on which the programme was based, and in turn linked to reductions in offending. From the mixed findings of the
current set of studies it is possible to offer some tentative conclusions about 'what works'.

- **focus on true criminogenic need;**

Criminogenic need refers to the key underlying reasons for offending. For example the Ilderton offenders seem really to have had a dominating interest in cars, which was a main reason for their offending. Similarly for the Demonstration Unit burglary offenders, burglary and offending generally may well have been seen by programme participants as a useful and valid way of getting money.

Perhaps more general programmes, like that of Sherborne House, lose a certain clarity of aim. These programmes have a more 'blunderbuss' approach, in terms of ideas about why young people offend, including external factors such as social deprivation and lack of opportunities, as well as internal factors such as low self-esteem, impulsiveness, and perceived benefits from offending. These may all be real needs, and reflect the complexity of offenders' lives, but this approach could also mean that key criminogenic needs may be addressed less clearly.

Findings for the Demonstration Unit programme for TDA offenders illustrate that if programmes are intended for offenders who specialise in one type of crime (excluding sex offenders), they should be defined as specialising in that type of crime by more than just their most recent offence.

- **disrupt criminal networks;**

It is possible for a offender to be on probation without coming into contact with other offenders, or only doing so in well-supervised settings (as is not the case for custodial sentences.) This means that probation programmes have the potential not to reinforce criminal identities and networks. From the current research, it is not
possible to assess the degree to which offenders were brought together in less supervised ways, though it seems probable the two probation centre programmes, Sherborne House and Camberwell, were more likely to have done this than the more successful Ilderton and Demonstration Unit burglary programmes.

the need for caution.

The two evaluation case-studies with the lowest client numbers (the Demonstration Unit programme for burglars, and Ilderton Motor Project) produced the most positive findings. The staff to programme participant ratio is unlikely to have been a factor given that the programmes with most clients were based at well-resourced probation centres with relatively large staff teams. Lipsey (1995) observed interventions with smaller numbers of clients tend to produce more positive findings. Copas and Mair (1998) attribute such findings to publication bias: small studies are only published when they produce interesting findings. However the fact that the smallest in terms of client-numbers of the current set of evaluation case-studies produce the most positive findings, is not because of publication bias. All effectiveness studies carried out within Inner London Probation Service, which used reconviction as their primary outcome measure, were included. No study was excluded because of negative findings. Furthermore the findings of these two most positive studies were statistically significant. This means that these findings are highly unlikely to have come about by chance, and are likely to be replicable with other similar offenders.

Conclusions about methods: identifying 'what works' and monitoring effectiveness

One of the main conclusions to be drawn from this project is that better primary research is needed.
random allocation;

Increasingly large and sophisticated syntheses will never be as persuasive as high-quality primary studies, preferably involving random allocation to intervention conditions (see Losel, 1995, McDonald et al, 1995). Random control studies have been carried out in probation settings (see for example Stanley and Goldberg, 1984), but have a tendency to break down especially when carried out on a large-scale, so researchers are unable to ensure protocols are adhered to, as found by Robinson (1995). As Geddes (1996) observes of psychiatry:

Large, pragmatic (i.e., reflecting everyday clinical practice), randomised, controlled trials are increasingly being recognised as providing the most accurate estimates of the effectiveness of clinical interventions.

- attention to processes, and to questions starting with 'how?'

As Pawson and Tilley (1994) have stated, evaluations which neglect to look at the processes involved in what is being evaluated risk being uninformative. Glaser (1975) makes a similar point when he argues that the most interesting questions in evaluation research are those which start with 'how?' (Meaning 'how were results acheived?' Rather than 'how much / how many'.) A drawback of the current set of research case-studies is outcomes are emphasised over processes, which makes the more equivocal findings difficult to interpret. Attention also needs to be given to process evaluation, and case studies. (Though this is not to de-emphasise the need for outcome evaluation, given the paucity of evidence in this area: see Ellis and Underdown, 1998.)

- theoretically grounded research (into theoretically grounded interventions)
More research is needed that is theoretically grounded, into interventions which are similarly based in theories of human behaviour, crime and deviency. Research which is uninformed by theory is less likely to contribute to a wider body of findings, adding to knowledge and relevant beyond the single study. For example the finding from the Camberwell Probation Centre study, that changes in criminal attitudes are not necessarily correlated with reduced likelihood of reconviction can be explained in terms of theories of social psychology. This finding is immediately relevant to programmes which rely on non-behaviourally correlated attitudinal measures for evaluating their effectiveness. But more importantly it could be followed up by further work to look at how attitudes to offending relate to offenders' behaviour.

- **prospective intervention-based studies;**

Intervention is used here in the sense of changes to an offender programme, rather than to offenders. Such studies can look at how changes to programme elements can affect outcomes. For example with the Camberwell Probation Centre study it would have been interesting to see how intermediate outcomes, completion and reconviction rates changed as new components of the programme (in particular reasoning and rehabilitation) were introduced and bedded in. This type of study would need to control for changes in case-mix and risk profile of offenders on the programme. This is the essence of the action research approach, which requires good integration of research and service delivery (the Demonstration Unit programmes in their work to provide change sentencing patterns for TDA offenders and burglars provide some illustration of this approach.)

- **standard performance indicators;**

The probation service should be capable of formulating a standard set of information for evaluating offender-programmes, about offenders, programmes and
outcomes. Developments in information technology which facilitate sharing information about clients across the probation service could support this. These include the CRAMS case-management system, and data-sharing mechanisms including intranets and the internet (see Home Office / Probation Service, 1993.) Standardisation of criteria of success across studies, or benchmarking of outcomes would also be useful. But it would need a cultural leap forward on the part of the probation service, for practitioners to understand better the value of rigorous monitoring, and participate in this. There might also be a need to overcome possible suspicions that monitoring is about staff appraisal.

- **the importance of objective measures of programme success;**

As shown by the Demonstration Unit TDA programme, whilst a project may generate enthusiasm and enjoyment amongst those who attend, and be experienced as a successful project by those running it, this does not mean that the project will necessarily reduce offending. Equally, as shown by the Demonstration Unit burglary programme, just because those running a project are less than confident about what they are doing, this does not necessarily mean that offending will not be reduced.

- **the need for suitable comparators;**

Findings based only on those who completed a project can be confusing. The Demonstration Unit team concluded their project had been a success, which in terms of its objectives it had been even with the TDA offenders. It expanded the use of probation amongst this group and held them in the community without markedly worse consequences than would have otherwise been the case. But only when a comparison group is used and convictions looked at over some years is it apparent that work with burglars succeeded in temporarily but significantly reducing offending, whilst that with TDA offenders did not.
the importance of reconviction as an outcome measure;

It is only possible to tell whether programmes reduce offending by looking at offending behaviour. There is simply no substitute for behavioural measures. The most easily available measure is reconviction. Intermediate outcome measures which seek to predict reconviction are not adequately developed at present (though this is a promising area for research - see Gerandreau et al, 1994, Losel, 1995, and particularly see Clark et al, 1993). The best indicator of offending behaviour comes from information about reconvictions, or better still when appropriately used (see chapter 3) arrest data. Closer working with police to get better data about arrest may be useful, in the interest of getting better feedback about offenders' behaviour. Arrest-data excludes much of the 'noise' introduced by court processes, and is closer to the actual behaviour probation seeks to influence. When evaluating reconviction outcomes for individuals, the nature, seriousness and frequency of offending can be taken into account. However as the research case studies have shown, and as might be expected on the basis of research into criminal careers (see Farington, 1992), on a programme level, information about frequency, nature and seriousness of reconviction conveys little more information about programme success than does a simple bipolar measure.

Because information simply about number of arrests may reflect operational policy in the police (for example to target known offenders), this data would need to be used in an informed way, perhaps by only including arrests at the scene of a crime, closely linked to a crime, or not followed by a not-guilty finding in court (methods for doing this were described in Chapter 3, and in Wilkinson, 1997.) There is a further argument that information about arrests may be a poor measure of offending because it could reflect bias in policing. But if information from the police were to highlight discrimination against some probation clients, this is not a reason for the probation service not to use this information. If discrimination in policing is revealed, this is a matter for the police service to do something about.
Mair (1992) argues in favour of using a range of indicators to evaluate achievement of the multiple aims of offender-programmes. But because offenders are defined by the fact they have been convicted and programmes exist to reduce offending, reconviction must stand preeminent as the key outcome indicator. For example in the case of the Demonstration Unit, it is only when reconvictions are looked at that the success of the programme for burglary offenders becomes apparent, otherwise this programme was not regarded as being as successful as the parallel programme for TDA offenders, which in fact was associated with increased convictions.

There is a need for more widely available high quality information about reconviction, to allow those responsible for the supervision of offenders to monitor the success or otherwise of their work. The Home Office Offenders Index may help with this. Information about arrests, as information technology makes this more readily accessible within criminal justice partnerships, may provide a more sensitive outcome measure than reconviction.

- making sense of information about reconvictions

Short term reconviction rates seem to yield the best measures of programme success, being closer in time to the programme in question.

Prediction techniques for making sense of reconviction data have been around for a long time (see for example Simon, 1971). However the OGRS (Copas and Marshall 1998) represents an advance, in offering an 'industry standard'. With increasing use of computers (NPSIS, 1994), and increased access to information about individuals' reconvictions (via the Home Office Offenders' Index), which could automatically include OGRS risk score, it could become entirely straightforward to look at the average risk levels of offenders on a given programme, or a given probation officer's caseload, and to compare this with actual reconvictions, in a manner initially proposed by the Audit Commission (1988).
- the role of intermediate outcome measures

More work is needed on intermediate outcome measures. Results reported above show that the use of attitudinal or psychometric scales as outcome measures must for the moment remain of doubtful value. The drawback of using instruments such as CrimePics as outcome measures, is that scores on such scales have value only in as much as they predict future offending. The extent to which they do this seems highly questionable (see Chapter 6, and see also Raynor, 1997.) A better approach is to use immediate outcome measures which measure something of intrinsic value, preferably but not necessarily linked to criminogenic need. Depending on the particular programme, such measures could include the number of offenders getting NVQs, or getting training or work placements, sustaining periods drug free, or on legal substitute prescribing, or taking up referrals to community resources. All these points measure real benefits, in a way scores on psychometric tests simply do not.

Intermediate outcome measures which have intrinsic value may also help to predict further offending. For example in the current set of studies, alarm bells should have been set ringing by the fact only a handful of the Demonstration Unit TDA offenders who satisfactorily completed the scheme were known to have passed a driving test, the original aim of the scheme.

- the importance of programme-completion

How completion rates are used as performance indicators is crucial. The danger of looking only at success rates for completers, certainly of community programmes, is that this reduces the pressure on those running programmes to drive up success rates for all participants. Completion also relates to targeting. A programme which determines that say older, lower risk offenders are more likely to complete might target this group, in order to obtain more programme members likely to benefit. But this policy it would be open to criticism for selecting 'easy' offenders, and neglecting offenders who are in greater need of help.

266
Getting research evidence into policy and practice:

This final section makes some observations on how the emerging evidence base can link more strongly to the probation service's work with offenders.

---

the need for a research culture, and for research and development capacity.

As Sheldon (1998) has observed of the overlapping field of social care, the research base for work with offenders is relatively small given the scale of the endeavour, the economic and social consequences of crime for the community, and the pace of change in this area. McIvor (1995) notes:

'research and evaluation have had limited impact on probation policy and practice' (Page 209)

Research and development is not a primary concern of criminal justice agencies. Research capacity to explore and evaluate 'what works' exists in the Home Office Research and Statistics Directorate, and in a few university departments. The probation service directly employs a handful of researchers, and the Prison Psychological Service makes some further contribution to the research evidence base (see for example Thornton et al, 1984.) However when the Home Office Research and Statistics Directorate surveyed probation services for examples of well-evaluated programmes which had shown some evidence of effectiveness, only four examples were discovered (three of which are included in the current research, see Ellis and Underdown, 1998.) The paucity of high-quality evaluation of services reflects the reality that the probation service is concerned primarily with delivering services, rather than with ensuring they are effective.

The issue of capacity is important. The research reported in this chapter was able to be carried out only because the organisation decided to employ designated research staff, and to make a programme evaluation a priority. Without research capacity,
whether generated via professional culture or organisational resources, high-quality research, which is embedded in the organisation, is unlikely to be carried out.

In health care, the importance of research and development is well acknowledged. Most hospital doctors undertake research as part of their training, and in the course of their specialist duties. Senior doctors may devote most of their time to research. Significant efforts are given to dissemination of findings as to treatment effectiveness (for example by the Cochrane Collaboration, and the NHS Centre for Reviews and Dissemination.) In other professions which have a science-base, such as clinical psychology, a significant proportion of practitioners carry out research. A number of authors have highlighted the need for practitioner-research in probation, and social work, including McIvor (1995), Cheetham et al (1992) and Evritt et al (1992.) McIvor (1995) point to the importance of an organisation supporting practitioner-evaluation, and the need for some 'pay-off' for the practitioner carrying out the research. There remains little of a research culture in the probation service. Without a culture which respects, understands and contributes to research evidence, there is less chance of effective interventions being identified and adopted.

- the problematic relationship between social science research and social policy;

It is not straightforward to link evidence produced by social research to positive practice. Pitts (1992), in a challenging and influential paper on this matter, observes that social research often undermines the 'common sense' views of the world required by policy-making and decision making. He goes on to observe that it is only findings which support time-limited, individualised interventions with an exclusive focus on offending that are admitted to the policy debate, and that perspectives which address the personal, social, economic and racial factors which increase vulnerability to involvement in crime are excluded. (Though this takes no account of policy focused research by Stewart et al, 1993, Mair and May, 1997, Aubrey and Hough, 1997.) Pitts (1992) goes on to declare
'the problem of a hybrid discourse... is that it must inevitably be conceptually bankrupt. Those elements of theory which are enslaved will be appropriated because they address the same object, in the same uncritical manner, as the legal/common sense discourse which has enslaved them.' page...

This is an argument against incorporating research findings piecemeal into the policy process. Rather than being an argument against evidence-based policy-making, it is an argument for theory-based policy making. However policy-making is rarely straightforward. Even where there are clear policy implications arising from research findings, the mechanisms whereby policy is informed are complex and non-linear, as Tizzard (1988) sets out. This is particularly the case in the field of offender-intervention, where there may be conflicting ideas as to the aim of sentences vis a vis punishment, retribution, deterrence, reparation, denunciation and rehabilitation.

- **the importance of reducing offending as an aim for the probation service.**

The current research shows that interventions which are responsive to probationers' offending-related needs can be effective in reducing offending. This underlines a moral imperative on the part of probation officers, (and of the probation service) to take measures likely to be effective in reducing offending (see for example McWilliams and Pease, 1990). These ideas are developed by Raynor and Vanstone (1994), to encompass newer findings as to effective methods. Raynor and Vanstone (1994) conclude that there is a moral imperative not to support offenders in choices likely to result in further offending, but amongst the necessarily constrained choices available to the offender, to support offenders' participating in interventions that teach skills which help avoid further offending.

This thesis has avoided discussion of ethical debates around coercion and control (which are far sharper where the stakes are higher, in the area of dangerous
offenders, and of mentally disordered people dealt with outside the criminal justice system, see for example Hinton, 1983.) These debates are likely to persist (see for example Neary, 1992a, 1992b.) But research evidence as to what is effective can change these debates, as Raynor and Vanstone (1994) show in their reworking of Bottoms and McWilliams' (1978) non-treatment paradigm for probation. Furthermore, if rehabilitation, defined as helping restore people to full citizenship (McWilliams and Pease, 1990) is a moral imperative, then so too must be identifying methods which are effective in supporting this, and ensuring that they are effectively implemented.

As Stewart and Stewart (1993) show, most offenders supervised by the probation service are unemployed, and face a variety of often compounding social problems of which their supervising probation officers are very aware. Coupled with this awareness, the analysis of crime provided by deviancy theorists, primarily Taylor et al (1974), had a profound influence over the probation service (though latterly 'left realist' approaches may have become more influential over probation officers' thinking.) This led many probation officers to see their role as being to help their clients address social problems, rather than challenge their offending. However the probation service will not eliminate causes of crime rooted in social injustice. As Bottoms and McWilliams (1978) drawing a public health analogy note: civil engineers have had more responsibility for ending disease than have doctors. (Though Wilson, 1995, describes an 'etiological crisis' amongst policy makers, based on the observation that post-war rising standards of living were associated with rising rather than falling crime rates. And Andrews, 1995, presents evidence that indicators of social deprivation are weaker predictors of offending than are individual psychological factors.)

The probation service may have a role in advocating on behalf of its clients, and pointing up the impact of social policy on a macro-level, as was done by the Association of Chief Officers of Probation, following the publication of Stewart and Stewart's (1993) report on the social circumstances of younger offenders. But what must define the role of the probation service day-to-day in its work with offenders,
is the application of methods likely to help offenders away from offending which harms their victims, themselves, their families and their communities.

Were probation to abandon reducing offending as an aim, it would be liable to become a surveillance arm of community correctional services, as Petersillia et al (1994) describe having happened in the United States. More importantly for service-users, were the probation service to move away from reducing offending as a primary aim, no other agency would be in a position to take on this task (the voluntary sector agencies active in this area lack capacity and expertise; the prison service is overloaded with the task of holding offenders in custody.) So offenders would be denied help.

This research has shown, in common with many other studies, that reducing offending is easier to state as an aim than to achieve in practice. For individual practitioners this is a difficult aim to pursue (probation work has been described as the social work profession's 'Russian front' by McDonald et al, 1992.) But reducing offending should remain a central aim of the probation service's work with offenders. Some ways of working with offenders clearly are effective in reducing offending, and there are promising directions for research to find out more about these methods and how they are effective. It is not yet possible to detail what methods are likely to be best with which offenders, and given the complexity of this task, this is very unlikely ever to become a matter of formula. But what is already known should give those who work with offenders grounds for optimism, in that they can be effective, and that still more effective ways of working can be found.
References


Association of Chief Officers of Probation (1994). *Probation Works and Everybody*
Benefits. London: ACOP.


Centre for Evidence-Based Social Care (1998). *What is Evidence-Based Social Care?* Quarterly Newsletter of the Centre for Evidence Based Social Services, Issue 1. Exeter: Centre for Evidence-Based Social Services.


Probation Service.


Home Office.


Inner London Probation Service (1993). Three year plan for the probation service,


Jeffers, S., 1995. *Black and ethnic minority offenders’ experience of the probation service*, University of Bristol, School of Advanced Urban Studies.


Lab, S. P. and Whitehead, J. T. (1990). *From "nothing works" to "the appropriate works": the latest stop on the search for the secular grail*. Criminology, 28 No. 3. pages 405-417.


Office Special Conferences Unit.


Unpublished paper distributed at 'What Works?' conference, Salford University.


