Summary
Lapse into old mistakes

Measuring recidivism to help determine the rehabilitative effect
of penal interventions

In criminology, measuring recidivism is a well-known method of
establishing the preventive effect of penal interventions. According to
Radzinowicz (1945), the first British professor of criminology, there are
measurements of this kind dating from as early as the 19th century.
Recidivism studies have been conducted in the Netherlands too, but the
earliest we know of stem from the 1970s. Due to developments in public
sector fund accounting methods, interest in reconviction rates is now
again on the increase. The Dutch government is looking for clear indica-
tors to gauge the success of implemented policy measures. In the area of
penal law – a sector involving expenditure amounting to billions of euros
each year – recidivism, or rather the absence thereof, is readily embraced
as a valid and efficient outcome indicator.
Under what circumstances do recidivism statistics provide proof of the
efficacy of policy interventions? And how can this construct best be
measured? In the Netherlands there is no tradition of recidivism research.
The methods used in past evaluation studies vary widely, the measure-
ments were predominantly small-scale, and most studies were conducted
at a time when the penal interventions under investigation were still at a
developmental stage. Recidivism research in the Netherlands has lacked
a sound methodological basis and there is consequently no comprehen-
sive view on the outcome of penal interventions – until recently, that is,
since the WODC has meanwhile developed a 'Recidivism Monitor'. This
is a project in which reconviction rates of virtually all offenders who
have come into contact with the Dutch judiciary are routinely computed.
The development of this instrument and its applications in evaluation
research form the central themes of this dissertation.

Chapter 1: Recidivism

Chapter 1 introduces four domains in which the concept of recidivism
plays an important role. From a political point of view recidivism – that
is repeat crime in the broadest sense of the word – is directly linked to
public safety. More than half of the registered crimes are committed by
repeat offenders. Halting them would lead to a considerable drop in crime
rates, thus helping citizens feel safer. The Dutch government has set
targets to lower reconviction rates among juvenile offenders and ex-pris-
oners by ten per cent by the year 2010.
A second outlook on the concept of recidivism entails a managerial
approach. For judicial agencies, recidivism serves as a possible outcome
indicator. Reconviction rates can be used within the organisation as input

data in a system of performance measurement. From a legal perspective, preventing recidivism is of course one of the established functions of the administration of penal law. Next to retribution and general deterrence, preventing reoffending is one of the objectives of imposing a sanction. Some offender programmes aim to reduce the risk that participants will reoffend through treatment; other interventions are supposed to trigger other mechanisms like deterrence or incapacitation.

In criminology in general, measuring recidivism is an empirical test. Whatever the theory on the development or discontinuation of criminal behaviour, collecting recidivism data will prove or disprove the hypotheses that were deduced from it. As Tarling said, recidivism research is 'one of the key tools of criminology'.

Chapter 2: Recidivism research

Chapter 2 discusses some of the landmarks in the history of international recidivism research. Data on reoffending can be applied merely descriptively, in research on criminal careers, in etiological research and in evaluation studies. Here we are primarily concerned with evaluation research, in which the recidivism data are used to determine the success of a planned intervention in circumstances that otherwise would be conducive to sustained criminal behaviour. Empirically the preventive effect can only be established by drawing a comparison between those who underwent the intervention and those who did not. When comparing the groups, the upfront similarities between the two will have to be maximised in order to exclude alternative explanations of any difference found in the outcome measure. In observational studies this is done using techniques of statistical control. The comparison is between persons who – based on a set of predictors – have the same a priori risk of reoffending. In quasi-experimental studies, participants in the group undergoing the intervention are matched to similar participants in a control group. In randomised controlled studies or 'true experiments', participants are assigned at random to the experimental or the control group. Most scholars agree that this latter research procedure yields the best comparison. A randomised controlled trial (rct) is considered the 'gold standard' of evaluation research.

Maltz (1984) mentions two other methodological issues that dictate the form of the recidivism research: the operational definition of recidivism; and the length of the follow-up period. There is a distinction between de facto and de jure recidivism, between the actual reoffending behaviour and what ends up in the records of the police department or the judiciary. An unknown number of offences escape the attention of the authorities, yet most recidivism studies use official data registered by the police departments or the Prosecutor's Office (PPS). The particular array of

offences that is included in the operational definition of recidivism has an important impact on the outcome of the research. So too has the length of the observation period. While there is no single ideal time frame for all recidivism studies, it is important to balance the observation period when comparing results because the longer the follow-up, the higher the recidivism rate. Recidivism can be analysed as an event that takes place in time. 'Survival' or 'failure time' models are statistical techniques that use the time elapsing before an event takes place as the dependent variable. Partanen (1969) was the first to apply this type of analysis in a criminological context. Cox (1972) introduced a regression technique for time-to-event data. This technique can be used to control for composition differences when comparing the recidivism rates of dissimilar offender groups.

Another important innovation on the international field of recidivism research has been the introduction of meta-analysis. In meta-analysis the results of a number of evaluation studies on some type or types of intervention are combined. The outcome of each study – here in terms of recidivism rates in the experimental and control group – is expressed as an overall effect size. By designing a common measure the methodological differences of the primary studies are eliminated. This enables the researcher to make more general statements as to the effectiveness of the intervention programmes being investigated. The technique of meta-analysis and the experimental tradition it stems from have met with much criticism, notably from sociologists like Pawson & Tilley (1997) who favour a 'realistic' evaluation approach. In their view the outcome of an intervention must not be isolated from the actual context in which it triggers mechanisms that may or may not alter the conduct of individual offenders. Both methods can be seen as complementary, however, as is elaborated in chapter 5.

Chapter 3: Measuring recidivism

Chapter 3 gives a detailed description of the 'Recidivism Monitor'. Another way to work around methodological differences in evaluation research is to standardise the measurements altogether. In a number of countries recidivism research is nowadays conducted periodically and on a national basis. In the Netherlands the WODC has developed procedures to routinely compute reconviction rates for a number of important offender groups using fixed criteria yielding comparable results. Data for the national research are drawn from the OBJD (Onderzoeks- en Beleidsdatabase Justitiële Documentatie) data base, a copy of the official registry of adjudicated criminal cases. The OBJD is updated with details of new penal cases every three months. The OBJD is maintained purely for research purposes. The data are retained indefinitely; they never become out of date.

Using the OBJD as central data source means that the Recidivism Moni-
tor takes into account only those offences that are handled by the Public
Prosecutions Service. No statements are made concerning either unde-
tected offences or offences that have not been prosecuted. Minor offences
(mostly traffic offences) are not generally taken into consideration. Nor
are criminal cases ending in acquittal, dismissal by reason of unlikeli-
hood of conviction or any other technical decisions. Only court cases in
which the accused has been found guilty and cases settled by the public
prosecutor by means of a discretionary dismissal or a 'transaction' are
included, as are cases on which a decision is still pending. Based on
this definition, the prevalence and frequency rates of several forms of
recidivism are drawn up, varying with the seriousness and type of crime
involved.

The reconviction rates describe the known recidivism in a specific
offender group. To interpret or explain the findings the data from the
OBJD will have to be linked to information on social-demographic and
personal factors that are theoretically relevant. Aside from judicial data
the OBJD registers only a few items of personal data: gender, date of birth
and nationality. These background characteristics are used as control
variables, to make better comparisons and estimate the net trend in the
recidivism rate. They do not in themselves explain why these rates are as
high as they are. For an understanding of the outcome of the Recidivism
Monitor, reference must be made to criminological theory and thoughts
on registration practices.

Chapter 4: Recidivism in the Netherlands

Chapter 4 presents some outcomes of the Recidivism Monitor for four
major offender groups: adults and juveniles who have been convicted of a
crime, former prisoners and former inmates of juvenile detention centres.
For these groups data are available on offenders who were convicted or
released between 1997 and 2004. Fluctuations in the composition of the
offender groups and the fact that the figures may be affected by changes
in registration practices make it difficult to keep a clear perspective on
the development of reconviction rates in the respective offender popula-
tions. That is why adjusted rates are computed. The observed rates are
corrected by means of a parametric survival model (Royston, 2001) which
estimates the influence of the recorded background characteristics on the
likelihood of reconviction and checks the data for significant registration
effects. The adjusted rates show the trends in the level of recidivism as if
there have been no changes in the composition of the offender popula-
tions, in the willingness to report a crime or in the national clear-up rate.
The diagram shows the percentages of recidivists within the successive
cohorts of the four offender populations. The ex-prisoners and former

inmates of juvenile detention centres have the highest rates, but then these are relatively high-risk groups. After checking for two possible registration effects and allowing for changes in the offenders' personal characteristics, it emerges that in three of the four populations the 2-year reconviction rate has increased by several percentage points across the total study period. Only among juvenile detainees has the prevalence of general recidivism declined. Why this is the case is to some extent open to debate. We know that during the study period the juvenile detention centres became progressively more heavily populated by detainees with a relatively low a priori risk of reoffending – more girls and more first offenders – but the figures in the above diagram already take these fluc-tuations into account. The juveniles who were institutionalised may also have become less 'recidivism-prone' in some hidden, unobserved ways. Due to a massive expansion of their combined capacity, juvenile detention centres have become less selective in admitting young offenders. Juveniles who would have been given a community service sanction some years ago are now institutionalised. That may be the reason why reconviction rates in this sector have fallen while in the other sectors of the administration of Dutch penal law they have risen.

Adjusted reconviction rates within 2 years after imposing the sanction or release from the institution
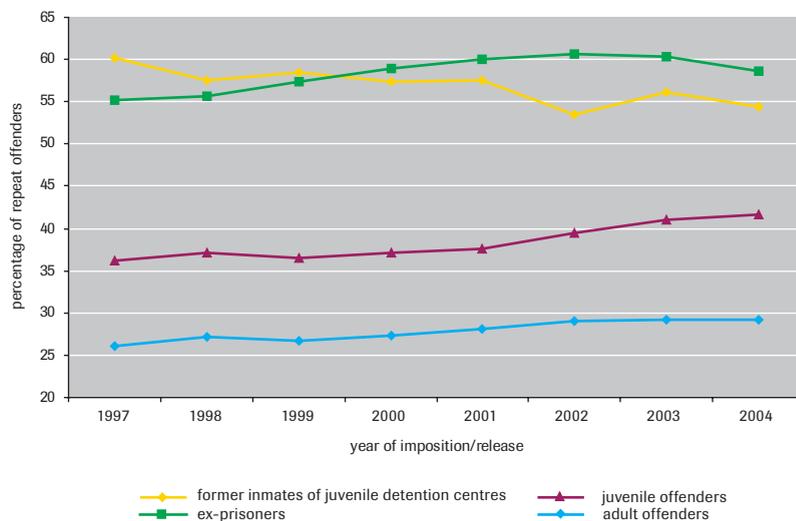
Chapter 5: Evaluation research

In chapter 5 we return to the broader context of evaluation research. Figures on recidivism do not in themselves demonstrate the effectiveness of penal interventions. More is required to establish the effectiveness of offender programmes. Measuring recidivism is only one part – the last link – in a series of evaluative activities. This 'chain of evaluation' consists of three phases: the phase in which the plans of the offender programme are thoroughly investigated, a phase where the processes underlying the implementation of the programme are scrutinised and a phase where the products of the planned activities are determined. In the case of offender programmes that are explicitly aimed at offender rehabilitation, measuring recidivism is – in the end – the obvious thing to do. But the outcome evaluation has to be preceded by a qualitative analysis of the mechanisms that are set in motion by the programme and which, at an individual level, are supposed to lead to a reduced risk of reoffending. Evaluating penal interventions means that we want to know not only whether a particular offender programme works, but also why and in what way the programme is successful. These last two questions are addressed in the plan and process phases of the evaluation chain.

The 'evaluation chain' is of course a compromise between the theory driven, realistic evaluation suggested by Pawson and Tilly and the experimental approach supported by the Campbell Collaboration. An experimental approach does not stand in the way of a qualitative analysis of the causal mechanisms that govern the impact the interventions have on the lives and conduct of the participants. Setting up an experiment in order to compare outcomes is not incompatible with such a qualitative analysis. On the contrary, in the chain of evaluations the outcome of an experiment should corroborate the evidence that has already been accumulated. If the plan evaluation proved that the underpinning of the programme is in line with what is known about 'what works', if the process evaluation has shown that the programme is administered in the way intended and if preliminary measurements of the impact of the programme on the lives of the participants have proved promising, we may assume with steadily growing confidence that the programme will in fact reduce recidivism. And if by any chance the final outcome measurement should show we were mistaken, we will adjust our theories of successful intervention programmes and try again.

Chapter 6: Toward a program of recidivism research

Chapter 6, the final chapter, outlines the themes of the dissertation and discusses the future development of the Recidivism Monitor. Broadly speaking there are three levels on which recidivism statistics can be used

to help evaluate criminal policy. On each level the problem of causation, the internal validity of the research, is addressed differently. Firstly, there is the level of national overviews of reconviction rates. Reconviction rates like those in the above diagram can contribute to a general understanding of the outcome of penal interventions taking place in the Netherlands. Although the trends revealed neither prove nor disprove that the policy measures taken on this area have been successful, they can serve to substantiate speculation about the effects of these measures by anyone who has a theory. In the end the significance of the basic outcomes of the Recidivism Monitor is determined in the realm of public and scientific debate. To help interpret the findings, the prediction models used to adjust the observed rates will be supplemented by indicators of socioeconomic developments. Unemployment figures and school dropout rates (or any other factor that is theoretically relevant) can be added to the model to see whether they account for the rise or fall of the reconviction rates.

A second application of recidivism statistics relates to organisational structures. In the Netherlands the police force is divided into 25 operational districts, the PPS and the courts into 19 geographical jurisdictions and the probation service into 10 regions. Prison and remand centres are organised in 20 clusters and there are at least 13 juvenile detention centres. In every sector the reconviction rates can be broken down by region to serve as input for a benchmark against which the performance of the units are compared to establish which practices work best. The problem here of course is that some units have to deal with more difficult offenders than others. Some of the differences in personal characteristics can be corrected however, by using prediction models. For each unit the expected level of recidivism can be computed based on the recorded characteristics of the offender group processed. This rate can then be compared to the actual reconviction rate to see whether the unit performs better or worse than might be expected. Care is still needed to guard against 'over-interpreting' the meaning of the reconviction rates. Not every unit is geared to constraining recidivism. However, provided there is agreement as to what extent the reduction of recidivism is an organisational goal, comparisons of expected and actual reconviction rates can be an informative part of a benchmark for like units working on like tasks. The third and possibly most decisive or conclusive evaluative use of recidivism statistics is the comparison of reconviction rates in an outcome study into the effect of a specific offender programme. The best way to make such a comparison is to organise a randomised controlled trial. A quasi-experimental design can be helpful too, especially in situations in which much is known about the circumstances that normally lead to recidivism. In both cases however a considerable number of observations is needed to be able to interpret a difference in reconviction rates as a statistically significant effect of the intervention under investigation.

In the Netherlands the sample sizes are often fairly small, which amplifies the need for a qualitative evaluation strategy preceding an outcome study using recidivism figures and statistical inferences. Careful plan and process evaluation also produces evidence of the workings of an intervention. If we are to lower recidivism among offenders we cannot afford to neglect whatever indication of effectiveness we may find.