Psychology, Crime & Law

Risk and need assessment in British probation: the contribution of LSI-R

Risk and need assessment in British probation: the contribution of LSI-R

PETER RAYNOR

Department of Applied Social Sciences, University of Wales, Swansea, UK

Abstract
From 1996 until about 2000 the Canadian Level of Service Inventory – Revised (LSI-R) was in use in a number of probation services in England and Wales, and it is still in use in the Jersey Probation and After-Care Service. This article reviews what has been learned about risk and need assessment in British probation through the use of LSI-R, drawing on data collected for a Home Office study and for evaluative research in Jersey. Particular areas of interest are accuracy, differences between male and female offenders, the comparative effectiveness of probation and community service, the apparent counterproductive impact of probation on low-risk offenders, and the efficacy of risk-related change measurement. The conclusion points out the wide-ranging advantages of risk/need assessment for probation services, and discusses why services in England and Wales have been slow to benefit from this.

Keywords: *Probation (England and Wales), probation (Jersey), risk, risk/need assessment, reconviction, gender*

Introduction

The context of this paper is the increasing concern with the concept and consequences of risk which has characterized correctional services, and particularly the Probation Service, in Britain over approximately the past decade. In part this reflects wider changes in society, and in the nature of perceived social threats, which lie beyond the scope of this paper (see, for example, Beck, 1992; Hudson, 2003; Kemshall, 2003). However, it seems safe to assume that probation services will remain concerned about risk for the foreseeable future. This concern takes at least two forms which have often been confused with each other. One concerns the protection of potential victims from very harmful acts by dangerous offenders: that is, with the prevention of actions which are statistically fairly rare, but have very serious consequences if they do occur (Her Majesty's Inspectorate of Probation, 1995; Maguire, Kemshall, Noaks, & Wincup, 2001). The other is concerned with assessing the probability of reconviction, particularly among those who persistently commit the most common kinds of crime. Here the risks of reoffending are typically much higher but the crimes are of a more everyday kind, the "bread and butter" of penal policy. This paper is mainly about the

Correspondence: Professor Peter Raynor, Department of Applied Social Sciences, Vivian Tower, University of Wales, Swansea SA2 8PP, UK. E-mail: p.raynor@swansea.ac.uk

latter form of risk, and particularly about the emergence there of the technique of risk/need assessment (Bonta, 1996), which estimates risk through an assessment of criminogenic needs in order to encourage a rehabilitative approach to reducing offending by addressing offenders' needs.

Broadly speaking, the salience of risk as an issue for British probation staff dates from the 1991 Criminal Justice Act and the requirement in the subsequent National Standards that probation officers should routinely carry out risk assessments (Home Office, 1992). No reliable method for doing this was prescribed or indeed known at that time; however, in the same year a statistical analysis of the criminal records of 13,711 offenders was used to develop the basis of the first national reconviction predictor for use in probation services in England and Wales (Copas, 1992). This predictor, based on age, sex, previous convictions, sentencing history and current offence ("static" risk factors) was the immediate ancestor of OGRS, the Offender Group Reconviction Scale, which is still used and regularly updated (Home Office, 1996; Taylor, 1999). As an actuarial instrument based on centrally recorded data it has proved reliable and practical, but as it includes no information on risk factors that correctional agencies might try to change ("dynamic" risk factors) it cannot be used to assess need or to plan or evaluate supervision.

The Inspectorate's report on *Dealing with Dangerous People* (Her Majesty's Inspectorate of Probation, 1995) identified a number of problems in probation officers' approaches to risk assessment, and the following year saw the publication of a Home Office research review (Kemshall, 1996); an assessment, case management and evaluation instrument (ACE) developed by the Oxford Probation Studies Unit (Roberts, Burnett, Kirby, & Hamill 1996); and the introduction to several probation areas of the Canadian Level of Service Inventory – Revised (LSI-R; Andrews & Bonta, 1995) through the work of the Cognitive Centre Foundation in South Wales. In the following year the Home Office published an evaluation of some simple assessment scales for probation officers which were not particularly satisfactory (Aubrey & Hough, 1997), and in 1998, following a number of small-scale pilot studies (Raynor, 1997, 1998a), the Home Office responded to the widespread adoption of LSI-R and ACE by commissioning an evaluation of both, which was published 2 years later (Raynor, Kynch, Roberts, & Merrington, 2000). The purpose of this paper is to review some of the findings that have emerged from the deployment of LSI-R in England and Wales up to that time, and in Jersey in the Channel Islands continuously from 1996 to the present. The data are drawn from the Home Office research study (HORS 211, Raynor et al., 2000), including some re-analysis; from probation areas that participated in that study; and from a series of evaluative studies being undertaken in the Jersey Probation and After-Care Service (Heath, Raynor, & Miles, 2002; Miles & Raynor, 2004; Raynor & Miles, 2001).

The LSI-R is a well-known 54-item risk and need assessment instrument, used in many countries (now including Scotland and Ireland) and recognized by the American Psychological Association. Based originally on work carried out by Don Andrews with probation officers in Ontario at the end of the 1970s (Andrews, 1982) and extensively refined over a number of years, it has recently been issued in a new "case management" version, the LS/CMI (Andrews, Bonta, & Wormith, 2004). The version used in Britain has been that published in 1995 (Andrews & Bonta, 1995), which is still the best-known and most widely used variant. This version was particularly well supported by research, including a meta-analysis of recidivism predictors which found it to be the best of the risk assessment scales reviewed (Gendreau, Little, & Goggin, 1996). More recently, Hollin (2002) has described it as having "the strongest research pedigree". Another recent meta-analysis

(Gendreau, Goggin, & Smith, 2002) lists no less than 49 studies of the predictive efficacy of LSI-R, of which the British study (Raynor et al., 2000) used the largest sample (948). This has now been exceeded in the Jersey study reported here and in Miles and Raynor (2004). Other research in Britain has concerned the use of LSI-R in prisons (Hollin, Palmer, & Clark, 2003). This paper, however, is particularly concerned with its potential as part of a strategy for improving probation practice or supporting "what works". Consequently it focuses on four issues that have been of particular interest in probation circles in recent years: predictive validity; differences between male and female offenders; comparing the effectiveness of different community sentences; and the evaluation of supervision through measuring risk-related change. Some of these findings have been the outcome of substantial targeted research whilst others have been a by-product of routine use in monitoring service delivery; however, together they serve as examples of how methods of this kind can help to inform more effective practice.

Predicting reconviction

In the Home Office funded evaluation published as HORS 211 (Raynor et al., 2000) the overall accuracy of prediction was measured as the "percentage correctly predicted", previously used by Copas (1992), Lloyd, Mair, and Hough (1994) and May (1999). This calculation involves taking the range of predictor values yielded by a sample, dividing them into "high" and "low" at a point corresponding to the proportions actually reconvicted or not reconvicted, then treating all "high" scores as predicting reconviction and all "low" scores as predicting non-reconviction. Reconvicted high scorers and non-reconvicted low scorers then count as "correct" predictions. (For example, for a group of offenders with known predictor scores and a known reconviction rate of 50%, the top 50% of scores would be counted as "high" and predicting reconviction, and the bottom 50% would be counted as predicting non-reconviction. In this example, a perfect predictor would score not 100% correct but 75% correct, since high scores actually indicate a range of probabilities between 50% and 100% and low scores a range between 0% and 50%. Random prediction would be expected to score 50%.) Table I summarizes the data collected from pilot areas in England and Wales for HORS 211, together with the most recent similar data from Jersey (Miles & Raynor, 2004).

Table I. LSI-R scores, reconvictions and prediction.

	n	Mean LSI-R score	% reconvicted in 1 year	% correctly predicted
England and Wales				
Men	785	20.0	43.9	65.5[*]
Women	163	21.2	35.0	65.0[*]
All	948	20.2	42.4	65.4[*]
Jersey				
Men	1170	16.9	26.8	69.4[*]
Women	210	15.7	9.0	86.7[*]
All	1380	16.7	24.1	71.6[*]

[*]$p < 0.001$, based on significance of difference between mean LSI-R scores of those reconvicted and not reconvicted. In the full England and Wales sample relevant statistics are: for reconvicted offenders mean LSI-R = 24.13 (SD 8.9) and for unreconvicted offenders 17.25 (SD 9.1). Jersey equivalents are 21.4 (SD 9.3) and 15.25 (SD 8.6).

For the 948 LSI-R assessed offenders who had all other relevant data available in HORS 211 the percentage correctly predicted (using reconviction of a standard list offence within 12 months as the criterion) was 65.4%, compared to 67.1% achieved in the same sample by OGRS2 (the current version of the Offender Group Reconviction Scale. The corresponding figure for ACE, using a different group of 903 probation cases, was 61.5%). The difference between mean LSI-R scores for those reconvicted and not reconvicted was highly significant ($p < 0.001$). As recommended by Lloyd et al. (1994) this simple measure was supplemented by examining predicted and actual reconviction rates across the range of predictor values. The proportions reconvicted in each quintile of the LSI-R score distribution were 18%, 31%, 44%, 55% and 68%, respectively. A "percentage correctly predicted" was also calculated using the subset of eight items which are included in the screening version of LSI-R (a simplified version designed for rapid use to identify those requiring a fuller assessment: Andrews & Bonta, 1998). This proved almost as accurate as the full LSI-R at 65.2%.

Although the LSI-R, based on both static and dynamic risk factors, did not perform quite as well in this study as OGRS2 based on static factors only, it has the advantage of helping in the identification of needs and targeting of services, which static-only predictors cannot. The detailed information on needs which is set out in HORS 211 (Raynor et al., 2000) has been supplemented by a number of area studies, and the emerging patterns of need are not surprising for a probation population (compare, for example, Mair & May, 1997). They also show expected differences: for example, problems with employment and money are more prevalent in samples drawn from economically deprived urban areas (Raynor, 1997). A more detailed discussion of prediction, including which items and combinations of items are most strongly associated with reconviction, is provided in HORS 211; however, it should be clear from the findings outlined above that for a simple instrument, LSI-R provided a useful form of risk and need assessment for those probation areas that experimented with it.

For Jersey, Table I contains information on 1380 offenders initially assessed between Autumn 1996 and the end of June 2001. The average LSI-R scores are somewhat lower than in England and Wales and the reconviction rates are considerably lower (probably reflecting the social and cultural characteristics of Jersey society – see Heath et al., 2002). However, the "percentage correctly predicted" is high. To some extent (and particularly for women: see below) this is a consequence of much lower reconviction rates, which create a higher probability of "correctly predicting" non-reconviction (for example, for women in Jersey an even higher correct prediction rate of 91% could have been achieved simply by predicting that none would reconvict). However, these figures indicate that for practical purposes LSI-R can help to distinguish offenders with a low risk of reconviction from those with a higher risk, and can therefore assist in the concentration of resources on the latter, where they are more likely to make a positive difference (Andrews et al., 1990). They also indicate that calibration of LSI-R, i.e. what actual risk of reconviction corresponds to a particular LSI-R score, is likely to vary in different jurisdictions, so that local evidence-based calibration is likely to be needed: however, the general association with reconviction (i.e. that higher scorers are more likely to reconvict than lower scorers) appears to be consistently present across the different areas and jurisdictions represented in these studies. Table II shows correlations (r) between LSI-R scores and reconvictions in both jurisdictions, and between OGRS2 scores and reconvictions in England and Wales.

Table II. Correlations (*r*) of LSI-R scores and OGRS2 scores with reconviction.

Score type	*n*	*r*
England and Wales		
All LSI-R	948	0.353
LSI-R men	785	0.361
LSI-R women	163	0.336
All OGRS2	948	0.393
OGRS2 men	785	0.397
OGRS2 women	163	0.342
All LSI-R "Screening Version" (SV)	805	0.317
LSI-R SV men	660	0.311
LSI-R SV women	139	0.343
Jersey		
All LSI-R	1380	0.287
LSI-R men	1170	0.285
LSI-R women	210	0.297

All significant: $p < 0.001$.

Male and female offenders

Risk and need assessment is not simply a practical way of estimating risk and targeting services. It also has the potential to provide a gradually accumulating evidence base in relation to a number of issues of concern in probation services. As the first example of this, reanalysis of data from HORS 211 allows some interesting comparisons to be made between the male offenders ($n = 785$) and the female offenders ($n = 163$) in the sample, and similar data are now available for Jersey ($n = 1170$ and $n = 210$, respectively).

Recent years have seen lively controversy about how far the results of research on samples of male or mostly male offenders can be applied to women (see, for a summary of the issues, Gelsthorpe, 2001). It is argued, for example, that women typically have lower reconviction rates, so that the application of risk predictors derived from male offenders may overpredict reoffending and lead to more severe sentencing (Shaw and Hannah-Moffatt, 2000; Hudson, 2002); that women and men offend for different reasons, so that little generalization from one to the other is possible (Kendall, 2002); and that the "what works" movement in general has failed to pay sufficient attention to women. Other research has suggested that a number of risk factors and criminogenic needs are similar (for example, Howden-Windell & Clark, 1999; Simourd & Andrews, 1994) and that the broad principles of effective work with offenders, such as the principles of targeting risks and needs (Dowden & Andrews, 1999), are as applicable to women as to men. One study concluded that "the concepts underlying the LSI appear robust enough to bridge the boundaries of gender ..." (Coulson, Ilacqua, Nutbrown, Giulekas, & Cudjoe, 1996, p. 437).

Whilst it appears generally acknowledged that the needs of women offenders have been the focus of less research than those of men, some researchers have broadly accepted the risk/needs framework but attempted to clarify where the evidence points to differences in needs. For example, Gelsthorpe (2001) points to evidence that women offenders show higher levels of poverty, victimization, mental health problems, self-harm and problems linked to family relationships. A meta-analysis by Hubbard and Pratt (2002) points to school and family relationships and histories of physical and/or sexual assault as risk factors which predict female offending rather than male offending. Howden-Windell and Clark (1999) also point to relationships which support offending, and substance abuse. One recent review draws on this and other similar material to reach "a tentative conclusion ... that there seem

to be some similarities but also some differences in what might constitute the most significant dynamic risk factors for men and women" (Porporino, Van Dieten, & Fabiano, 2003, p. 2). The authors go on to develop a model of women's acquisitive offending as a consequence of "cumulative social and emotional disadvantage", and argue that anti-social beliefs and attitudes play a smaller role in explaining women's offending than in explaining offending by men.

Table I reanalyses the HORS 211 sample to show average LSI-R scores, reconviction rates and percentages correctly predicted for men and women separately, and includes a similar analysis of new data from Jersey. In both areas, for both men and women, the percentage correctly predicted is high, with statistically significant differences between the mean scores of those who reconvict and those who do not. This suggests that at least some of the same risk factors serve to distinguish between reconvicted and non-reconvicted offenders in both groups with about the same degree of reliability. However, the *level* of reconviction risk represented by a given LSI-R score is substantially lower for women than for men. In the HORS 211 sample the reconviction rate for women was some seven percentage points lower while their mean LSI-R score was actually higher. In Jersey, women had slightly lower average LSI-R scores but dramatically lower reconviction rates. This suggests that there are grounds for anxiety about gender-specific overprediction unless the instrument is appropriately re-calibrated for use with women offenders. However, one would expect it, after re-calibration, to be about as reliable in use with women as with men.

There are also clear indications that the amount of re-calibration should depend on the particular characteristics of the population to which it is being applied. Table III (based on Miles & Raynor, 2004) compares the LSI-R scores and 1-year reconviction rates for men and women in Jersey in four risk bands, based on the quartile distribution of LSI-R scores for all offenders in that area. This shows that the difference between the reconviction rates of men and women is greater in Jersey than in England and Wales, and is present across the whole risk range. This may reflect the fact that Jersey (effectively an autonomous microstate with its own government and legal system) is a smaller community which retains, in many respects, the characteristics of a rural and non-industrial society (Heath et al., 2002). Whatever the precise reasons, findings such as these strongly suggest that cultural or economic difference between communities can affect the relationship between gender, LSI-R scores and risk.

The data generated by LSI-R in British probation are also interesting in relation to questions about differences in needs. Here it is important to distinguish (which most writers on this subject do not) between questions of *relevance* (e.g. do the same risk factors tend to

Table III. LSI-R scores and reconvictions: Jersey quartile risk bands.

Quartiles (based on full Jersey sample)	*n*	% reconvicted in 1 year
Men		
1 (LSI-R = up to 9)	275	13.1
2 (LSI-R = 10–15)	296	17.9
3 (LSI-R = 16–22)	292	30.1
4 (LSI-R = 23 and over)	307	44.6
Women		
1 (LSI-R = up to 9)	66	1.5
2 (LSI-R = 10–15)	54	7.4
3 (LSI-R = 16–22)	45	8.9
4 (LSI-R = 23 and over)	45	22.2

increase the probability of reconviction among both men and women?), questions of *weight* (e.g. does a given risk factor increases reconviction to the same extent for both men and women?) and questions of *incidence* (do men and women who offend show different distributions of risk factors, so that the aetiology and reasons for their specific offences typically differ?). The data reviewed above are consistent with the hypothesis that a similar range of risk factors is *relevant* to assessing risks of reconviction for men and women, although their *weights* or the levels of reconviction with which they are associated are consistently lower for women. However, reanalysis of the HORS 211 data showed clear differences in the *incidence* of particular risk factors. When average scores on LSI-R components were compared as percentages of the possible maximum scores, the risk factors for which incidence among women was 10 or more percentage points higher than incidence among men were (in descending rank order) problems with employment, finances, drug misuse and family relationships. The risk factor for which men's scores most exceeded women's scores was criminal history. This pattern is broadly consistent with the arguments advanced by Porporino et al. (2003), and suggests that far from being irrelevant or mistaken when applied to women offenders, careful use of the risk/needs model can make a useful contribution to the evidence-based exploration of differences.

The impact of community sentences

The third area identified at the beginning of this paper as possibly open to illumination through risk/needs data was the relative efficacy of community sentences, particularly at low and moderate risk levels. The beneficial impact of appropriate programmes of supervision at higher risk levels is now amply attested by a number of systematic reviews (see, for example, Andrews et al., 1990; McGuire, 2002), but the situation at lower risk levels is not always so clear, and many recipients of community sentences are at these lower levels. Specifically, questions have been raised about the relative efficacy of probation orders and community service orders (recently renamed community rehabilitation orders and community punishment orders, respectively, but probably more familiar to readers under their traditional names), and about the impact of probation orders on low-risk offenders. This section of the paper considers what can be learned about these questions from the available data on LSI-R in British probation.

A number of studies which have compared expected and actual rates of reconviction have suggested that for offenders with similar initial risks of reconviction, actual reconviction rates following community service orders are typically lower than those following probation orders (examples are Lloyd et al., 1994 and Raynor & Vanstone, 1997). Such findings have often been quoted in support of proposals to increase investment in the development of community service orders, most recently in the form of Enhanced Community Punishment (National Probation Service, 2002). However, these arguments have always been open to the criticism that because the expected reconviction rates are calculated using predictors based on static risk factors such as age, sex and criminal history, they could be concealing higher levels of need among those made subject to probation orders, and this could render comparisons based on static risk factors unfair. There is evidence that people with more needs or problems (dynamic risk factors) are more likely to be sentenced to probation orders, probably because sentencers recognize the rehabilitative component in probation orders and probation officers writing pre-sentence reports tend to select people with more problems when proposing a probation order (Raynor, 1998b; May 1999). This raises the possibility that comparisons of the outcomes of probation and community service might

lead to different conclusions if initial levels of dynamic risk factors were taken into account in addition to the static risk factors which have customarily been used.

Table IV is based on offenders subject to pre-sentence assessments using LSI-R in Gloucestershire from 1997 to March 1998, and provides an example of how expectations based on static factors only (OGRS2 in this case) and expectations based on combined static and dynamic assessment (LSI-R) can point to different conclusions. At first sight, comparing the OGRS2 scores of the community service group with those of the probation group (which are not significantly different) would lead us to expect about the same reconviction rate in both groups. In fact it is 7% lower in the community service group, looking like yet another study in which community service performs rather better than expected. However, if we look at the LSI-R scores we see a much larger gap between the community service and probation groups, reflecting the greater incidence of dynamic risk factors in the latter. This difference is significant ($p < 0.001$). Drawing on the approximate conversion tables derived from the LSI-R probation service pilot studies (Raynor, 1998a; Raynor et al., 2000) this would lead us to expect a reconviction rate for probation orders around 14% higher than community service. In practice it is only 7% higher, which now looks like quite a good result for probation.

Similarly, the comparison of outcomes between probation orders and prison, when considered on the basis of static factors only, suggests that prisons should be producing about 16% more reconvictions than probation, corresponding to the difference in the OGRS2 scores (which is significant: $p < 0.01$). The actual difference in reconviction rates is only 7%, which looks like another relatively poor outcome for probation. However, the LSI-R assessments for these two groups are much closer and not significantly different, suggesting little difference in expected reconviction rates. The observed difference of 7% could now be seen as a rather better result for probation. Again, the fact that different groups of offenders can have different balances of static and dynamic risk factors (in this case, longer criminal records among the prisoners) can lead to quite misleading conclusions if only static factors are considered. Bearing in mind the small numbers and the inherent limitations of predictors, this example from Gloucestershire is suggestive rather than conclusive, but it shows how apparent findings can actually be reversed if a more comprehensive form of risk assessment is used. Matching of samples on dynamic as well as static risk factors has recently been proposed as a technique which could lead to more accurate conclusions in the evaluation of prison-based cognitive skills programmes (Cann, Falshaw, Nugent, & Friendship, 2003). The figures in Table IV strongly suggest that similar considerations apply in community sentences, and should be taken into account in future studies. (The final section of this paper will explore some reasons why this has not yet happened.)

The second example of how risk/needs data can be used to reformulate some old problems is perhaps less encouraging for probation. This concerns the impact of probation on low-risk offenders. It is well known that the proportion of offenders on probation who could be described as "low risk" has been growing since the early 1990s (Raynor, 1998c).

Table IV. Probation, community service and prison in Gloucestershire: risk assessments and reconviction.

Sentence	Number	Mean OGRS2	Mean LSI-R	% reconvicted
Probation	202	51.5	22.4	49
Community service	186	46.8	16.1	42
Prison	52	67.6	23.9	56

In 1991 11% of new probationers had no previous convictions; by 2001 this had risen to an astonishing 27% (Home Office, 2002). The Chief Inspector of Probation has described this as the "silting up" of probation caseloads with low-risk offenders (Morgan, 2002, 2003). One reason for concern about this is that a large reconviction study some years ago (Walker, Farrington, & Tucker, 1981) found that first offenders sentenced to probation were twice as likely to be reconvicted as first offenders who were fined. Again this study, although arguably consistent with the "risk principle" (Andrews et al., 1990), can be criticized on the grounds that the matching of groups of offenders on criminal history only might disguise the greater needs of those sentenced to probation. However, the Jersey LSI-R data strongly suggest that even when matching includes dynamic factors, probation does not look like the best option for low-risk offenders (some possible reasons for this are suggested in Raynor, 2004).

Table V compares the outcomes of different community sentences for those with LSI-R scores from 10 to 15 (this is the low/medium Jersey quartile; there are virtually no probationers in the lowest quartile). It is striking that *for this low-risk group*, the reconviction rate for probation is much higher than community service or fines in spite of the similar initial LSI-R scores. [The difference in outcome between probation and community service is significant at 12 months ($p < 0.05$), and the differences between probation and community service and between probation and fines are both significant at 24 months ($p < 0.05$).] In other words, Walker's finding about low-risk probationers still fits the Jersey data even after dynamic factors are included. The rather more encouraging findings from Gloucestershire were based on data from all risk groups, and the Jersey probation outcome data are more encouraging for high-medium and high risk groups (Miles & Raynor, 2004), but probation officers are now discouraged from proposing probation for low-risk offenders in Jersey. Unfortunately, no such study has yet been carried out in England and Wales; if it had been, and had led to similar findings, the problem of "silting up" might have been more energetically addressed. Such findings, of course, always require replication and amplification before major policy conclusions can be drawn, but they already provide enough evidence to point to the potential benefits of risk/need methods in exploring these issues.

Risk-related change measurement

One important claimed advantage of risk/need assessment methods is that because they are partly or wholly dynamic, they can be used for repeat assessment to measure changes in dynamic factors, and consequently risk, which occur during supervision. This offers, for example, the prospect of evaluating programmes without waiting for follow-up reconviction studies before any conclusions can be drawn. This potential for risk-related change measurement (or "dynamic predictive validity", Bonta, 2002) is demonstrated if it can be shown that changes over time in LSI-R scores are in fact related to changes in the probability of reconviction. Until recently such evidence was available only from suggestive

Table V. Low/medium risk community sentences in Jersey.

Sentence	Number	Mean LSI-R	% reconvicted within 12 months	% reconvicted within 24 months
Community service	67	11.8	10	21
Fine	69	12.3	13	20
Probation	92	12.6	24	39

Table VI. Risk-related change.

Group	n	Mean LSI-R score on first assessment	% reconvicted in 12 months	Significance*
England and Wales				
Low start: decreasing	42	13.6	26	
Low start: increasing	31	15.2	55	0.013
High start: decreasing	47	27.7	55	
High start: increasing	37	27.6	78	0.027
Jersey				
Low start: decreasing	69	15.0	29	
Low start: increasing	29	14.8	59	0.006
High start: decreasing	84	28.3	54	
High start: increasing	21	27.4	76	0.06

*Significance here is based on a chi-square test of the difference in reconvictions between increasers and decreasers.

findings in relatively small-scale Canadian studies (Andrews & Robinson, 1984; Motiuk, Bonta, & Andrews, 1990) but the research carried out on LSI-R in Britain has added substantially to this. Table VI summarizes data on risk-related change from those offenders in the HORS 211 sample who had repeat assessments, and from more recent research in Jersey on offenders undertaking programmes between autumn 1996 and June 2001.

In Table VI this information is presented separately for low and high initial scorers to illustrate how it operates across the range. The first column divides each sample (England and Wales, and Jersey) into four groups: those with below average initial scores whose scores decreased further when reassessed after a period of supervision ("low start decreasing"); those with below average initial scores whose scores increased on reassessment ("low start increasing"); those with above average initial scores whose scores decreased on reassessment ("high start decreasing") and those with above average initial scores whose scores increased on reassessment ("high start increasing"). The point of considering high and low starters separately is to guard against the possibility that changes in scores might simply represent regression towards the mean, and average initial scores for each group are given in the second column to show that there is little difference in initial scores between decreasers and increasers in each initial risk group. It is then easily seen that among both high and low starters, those whose scores increased were reconvicted at a noticeably higher rate than those whose scores decreased. In the England and Wales group this difference is significant for both low and high starters. In the Jersey group, where the overall proportion of "increasers" is lower, the difference is significant for low starters and near-significant for high starters.

Table VII brings together all the increasers from both samples and compares their reconviction rate with that for all decreasers: again the difference is highly significant, giving support to the dynamic validity of the LSI-R in these applications. [The HORS 211 study also showed good risk-related change measurement by the all-dynamic assessment

Table VII. Direction of change and reconvictions (all cases, $n = 360$).

LSI-R scores:	Reconvicted	Not reconvicted
Increased during supervision	79 (67%)	39 (33%)
Decreased during supervision	102 (42%)	140 (58%)

Significance (chi-square): $p < 0.001$.

instrument ACE (Roberts et al., 1996), but overall predictive validity was rather less for ACE than for LSI-R.] In the Jersey Probation and After-Care Service the LSI-R has been routinely used in repeat assessments since 1996, and most of their offending behaviour programmes show statistically significant improvement in scores (Miles & Raynor, 2004).

Conclusion

These examples show the broad potential contribution of risk/need assessment to British probation. As well as helping to throw light on some long-standing problems, this approach to assessment offers considerable possibilities in the field of real-time evaluation, allowing the measurement of risk-related change during supervision and helping in the identification of effective and ineffective practices without the long delays required to undertake reconviction studies. There is, of course, a price to be paid by practitioners in adopting such methods: those who value the traditional autonomy, indeterminacy and subjectivity of individual "clinical" judgement find that the use of a standard method for assessment challenges their assumptions in a number of ways. These challenges have recently been explored in a particularly useful series of papers by Gwen Robinson (Robinson, 1999, 2001, 2002, 2003). However, there seems little doubt that these methods must play an important role in the development of rational penalties that reduce the tendency to offend, and the National Probation Service has been right to adopt a form of risk/need assessment as standard practice.

What might perhaps be questioned is the route taken to achieve this. The reader will probably have noticed that all the most recent data in this paper come not from England or Wales but from Jersey. This is because the Jersey Probation and After-Care Service has used LSI-R continuously since 1996, whereas its use in England and Wales declined after 1999. At that time, at least half the probation areas in England and Wales were gaining experience with risk and need assessment through the use of either LSI-R or ACE (Raynor et al., 2000), but early in 1999 the Home Office decided to design its own risk/need instrument. Rather oddly, this decision was announced by circular (Home Office, 1999; Robinson, 2001) only days before the Home Office was due to receive an interim report on the evaluation of LSI-R and ACE which contained the first comprehensive assessment of the accuracy and efficacy of these instruments in probation services in England and Wales. The new instrument, which was to be developed instead of the previously announced tendering exercise, was to be ready by August 2000. This underestimation of development time (it is still not fully in general use at the time of writing) may have been due to a failure to appreciate what is involved in the development and testing of such an instrument. Most of us, if we need a house, will try to buy or rent one, broadly suitable for our needs, which has already been built. Only those with plenty of time and resources, who can afford to delay moving in for several years, will buy a field and start digging the foundations for a mansion.

What eventually emerged from a very substantial (and presumably costly) design and development effort was the Offender Assessment System, OASys (OASys Development Team, 2001; at that time the team had 16 members). OASys is a very comprehensive and strongly research-based assessment instrument, informed by detailed study of others including LSI-R and ACE. An internal Home Office study also collected information on practitioners' concerns about LSI-R and ACE to pave the way for the new instrument (Aye-Maung & Hammond, 2000), but no such study of the new instrument has been published yet. OASys has attracted criticism from probation practitioners for being over-complex and too time-consuming at a time of increasing workloads, and was one factor in

the industrial action in many probation areas over workloads in 2002–2003. The official Home Office estimate for the time taken to complete an OASys assessment is now two and a half hours (personal communication from National Probation Directorate, 2003), whereas experienced practitioners in the 1990s reported to the Cognitive Centre that LSI-R typically added between 10 and 15 minutes to the time taken to prepare a pre-sentence report (Raynor, 1997).

The National Probation Directorate's research on the main pilot studies of OASys has not been published at the time of writing, although substantially complete in 2002 (Clark, Garnham, & Howard, 2002). However, a number of public statements by the Home Office (for example, Mason, 2003) have indicated that in the pilot studies it achieved a level of accuracy fractionally higher than that previously established for LSI-R in Britain (for all practical purposes about the same), and like LSI-R it is slightly less accurate than OGRS2 (see also Merrington, 2004). There is at the time of writing no published information on its capacity to act as a risk-related change measure. What it clearly does provide is a far greater quantity of information, and while questions might be asked about the cost-effectiveness of collecting all this in all cases, the research potential is considerable, provided that practitioners are sufficiently convinced of its value to collect the information reliably. In the meantime, no general approach to risk/need assessment has been in place in England and Wales to underpin the rapid introduction of probation programmes based on "what works" principles, and the initial targets for completion of programmes were based on negotiations with the Treasury in 1999 rather than on any measurement of the need for them or the numbers of offenders likely to benefit. These targets have been difficult to meet; in fact, in most areas they have not been met (Her Majesty's Inspectorate of Probation, 2003) and they have recently been revised downwards. Earlier implementation of risk/need assessment might well have avoided some of these problems. However, the means for this implementation are now in place, and we can expect England and Wales to begin to reap the benefits that risk/need assessment has already brought to other jurisdictions, provided that an appropriate balance between comprehensiveness and practicality can eventually be struck.

Acknowledgements

References

Andrews, D. (1982). *The Level of Supervision Inventory (LSI): The first follow-up*. Toronto: Ontario Ministry of Correctional Services.
Andrews, D., & Bonta, J. (1995). *The Level of Service Inventory – Revised manual*. Toronto: Multi-Health Systems.
Andrews, D., & Bonta, J. (1998). *The Level of Service Inventory – Screening version*. Toronto: Multi-Health Systems.
Andrews, D., Bonta, J., & Wormith, J. S. (2004). *Level of Service/Case Management Inventory (LS/CMI): An offender assessment system. User's manual*. Toronto: Multi-Health Systems.

Andrews, D., & Robinson, D. (1984). *The Level of Supervision Inventory: second report*. Report to Research Services. Toronto: Ministry of Correctional Services.

Andrews, D., Zinger, I., Hoge, R., Bonta, J., Gendreau, P., & Cullen, F. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, *28*, 369–404.

Aubrey, R., & Hough, M. (1997). *Assessing offenders needs: Assessment scales for the Probation Service*. Research Study 166, London: Home Office.

Aye-Maung, N., & Hammond, N. (2000). *Risk of reoffending and needs assessments: The user's perspective*. Research Study 216, London: Home Office.

Beck, U. (1992). *Risk society: Towards a new modernity*. London: Sage.

Bonta, J. (1996). Risk–needs assessment and treatment. In A. Harland (Ed.), *Choosing correctional options that work* (pp. 18–32). London: Sage.

Bonta, J. (2002). Offender risk assessment: Guidelines for selection and use. *Criminal Justice and Behavior*, *29*, 355–379.

Cann, J., Falshaw, L., Nugent, F., & Friendship, C. (2003). *Understanding what works: Accredited cognitive skills programmes for adult men and young offenders*. Findings 226, London: Home Office.

Clark, D., Garnham, N., & Howard, P. (2002). *An evaluation and validation of the Offender Assessment System (OASys)*. Draft report to the Home Office (unpublished).

Copas, J. B. (1992). *Statistical analysis for a risk of reconviction predictor*. Report to the Home Office, Warwick: University of Warwick (unpublished).

Coulson, G., Ilacqua, G., Nutbrown, V., Giulekas, D., & Cudjoe, F. (1996). Predictive utility of the LSI for incarcerated female offenders. *Criminal Justice and Behavior*, *23*, 427–439.

Dowden, C., & Andrews, D. (1999). What works for female offenders: A meta-analytic review. *Crime and Delinquency*, *45*, 438–452.

Gelsthorpe, L. (2001). Accountability: difference and diversity in the delivery of community penalties. In A. Bottoms, L. Gelsthorpe, & S. Rex (Eds.), *Community penalties: Change and challenges* (pp. 146–167). Cullompton, UK: Willan.

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the 'unparalleled' measure of offender risk? A lesson in knowledge cumulation. *Criminal Justice and Behavior*, *29*, 397–426.

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult recidivism. *Criminology*, *34*, 575–607.

Heath, B., Raynor, P., & Miles, H. (2002). What Works in Jersey: The first ten years. *VISTA*, *7*, 202–208.

Her Majesty's Inspectorate of Probation (1995). *Dealing with Dangerous People: the Probation Service and Public Protection*. London: Home Office.

Her Majesty's Inspectorate of Probation (2003). *Annual Report 2002/3*. London: Home Office.

Hollin, C. (2002). Risk–needs assessment and allocation to offender programmes. In J. McGuire (Ed.), *Offender rehabilitation and treatment* (pp. 309–332). Chichester: Wiley.

Hollin, C., Palmer, E., & Clark, D. (2003). The Level of Service Inventory – Revised profile of English prisoners. *Criminal Justice and Behavior*, *30*, 422–440.

Home Office (1992). *National standards for the supervision of offenders in the community*. London: Home Office.

Home Office (1996). *Guidance for the Probation Service on the Offender Group Reconviction Scale*. Probation Circular 63/1996, London: Home Office.

Home Office (1999). *Effective Practice Initiative: a joint risk/needs assessment system for the Prison and Probation Services*. Probation Circular 16/1999, London: Home Office.

Home Office (2002). *Probation statistics. England and Wales 2001*, London: Home Office.

Howden-Windell, J., & Clark, D. (1999). *Criminogenic needs of female offenders: A literature review*. Report to H. M. Prison Service, London (unpublished).

Hubbard, D., & Pratt, T. (2002). A meta-analysis of the predictors of delinquency among girls. *Journal of Offender Rehabilitation*, *34*, 1–13.

Hudson, B. (2002). Gender issues in penal policy and penal theory. In P. Carlen (Ed.), *Women and punishment* (pp. 21–46). Cullompton, UK: Willan.

Hudson, B. (2003). *Justice in the risk society*. London: Sage.

Kemshall, H. (1996). *Reviewing risk: A review of research on the assessment of risk and dangerousness: implications for policy and practice in the Probation Service*. Report to the Home Office Research and Statistics Directorate. Birmingham: University of Birmingham.

Kemshall, H. (2003). *Understanding risk in criminal justice*. Maidenhead: Open University Press.

Kendall, K. (2002). Time to think again about cognitive–behavioural programmes. In P. Carlen (Ed.), *Women and punishment* (pp. 182–198). Cullompton, UK: Willan.

Lloyd, C., Mair, G., & Hough, M. (1994). *Explaining reconviction rates: A critical analysis*. Research Study 136, London: HMSO.

Maguire, M., Kemshall, H., Noaks, L., & Wincup, E. (2001). *Risk management of sexual and violent offenders: The work of public protection panels*. Police Research Series 139, London: Home Office.

Mair, G., & May, C. (1997). *Offenders on probation*. Research Study 167, London: Home Office.

Mason, R. (2003). Letter to Cognitive Centre Foundation, 3 September.

May, C. (1999). *Explaining reconviction following a community sentence: the role of social factors*. Home Office Research Study 192, London: Home Office.

McGuire, J. (2002). Integrating findings from research reviews. In J. McGuire (Ed.), *Offender rehabilitation and treatment* (pp. 3–38). Chichester: Wiley.

Merrington, S. (2004). Assessment tools in probation: Their development and potential. In R. Burnett, & C. Roberts (Eds.), *What works in probation and youth justice* (pp. 46–69). Cullompton, UK: Willan.

Miles, H., & Raynor, P. (2004). *Community sentences in Jersey: Risks, needs and rehabilitation*. St Helier: Jersey Probation and After-Care Service.

Morgan, R. (2002). Something has got to give. *HLM – the Howard League Magazine*, 20(4), 7–8.

Morgan, R. (2003). Foreword. *Her Majesty's Inspectorate of Probation Annual Report 2002/2003*. London: Home Office.

Motiuk, L., Bonta, J., & Andrews, D. (1990). *Dynamic predictive criterion validity in offender assessment*. Paper presented to the Canadian Psychological Association, Ottawa.

National Probation Service (2002). *Enhanced community punishment*. Application to Joint Accreditation Panel (unpublished).

OASys Development Team (2001). *Offender Assessment System User Manual*, London: Home Office.

Porporino, F., Van Dieten, M., & Fabiano, E. (2003). *A Women's Programme for acquisitive female offenders: theory and applications manual*. Ottawa: T3 Associates.

Raynor, P. (1997). *Implementing the Level of Service Inventory – Revised in Britain: Report 2: Initial results from five probation areas*. Dinas Powys: University of Wales, Swansea and Cognitive Centre Foundation.

Raynor, P. (1998a). *Implementing the Level of Service Inventory – Revised in Britain: Report 3: Risk and need assessment in the five pilot areas*. Dinas Powys: University of Wales, Swansea and Cognitive Centre Foundation.

Raynor, P. (1998b). Attitudes, social problems and reconvictions in the STOP probation experiment. *Howard Journal*, 37, 1–15.

Raynor, P. (1998c). Reading probation statistics: a critical comment. *VISTA*, 3, 181–185.

Raynor, P. (2004). Opportunity, motivation and change: some findings from resettlement research. In R. Burnett, & C. Roberts (Eds.), *What works in probation and youth justice* (pp. 217–233). Cullompton, UK: Willan.

Raynor, P., & Miles, H. (2001). *Risks, Needs and Reoffending: evaluating the impact of community sentences in Jersey*. Report to the Jersey Probation Service. Swansea: University of Wales, Swansea.

Raynor, P., & Vanstone, M. (1997). *Straight Thinking On Probation (STOP): The Mid Glamorgan Experiment*. Probation Studies Unit Report 4, Oxford: University of Oxford Centre for Criminological Research.

Raynor, P., Kynch, J., Roberts, C., & Merrington, M. (2000). *Risk and Need Assessment in Probation Services: an Evaluation*. Research Study 211, London: Home Office.

Roberts, C., Burnett, R., Kirby, A., & Hamill, H. (1996). *A system for evaluating probation practice*. Probation Studies Unit Report 1, Oxford: University of Oxford Centre for Criminological Research.

Robinson, G. (1999). Risk management and rehabilitation in the probation service: collision and collusion. *Howard Journal*, 38, 421–433.

Robinson, G. (2001). Power, knowledge and What Works in probation. *Howard Journal*, 40, 235–254.

Robinson, G. (2002). Exploring risk management in the Probation Service: Contemporary developments in England and Wales. *Punishment and Society*, 4, 5–25.

Robinson, G. (2003). Technicality and indeterminacy in probation practice: a case study. *British Journal of Social Work*, 33, 593–610.

Shaw, M., & Hannah-Moffatt, K. (2000). Gender, diversity and risk assessment in Canadian corrections. *Probation Journal*, 47, 163–172.

Simourd, L., & Andrews, D. (1994). Correlates of delinquency: a look at gender differences. *Forum on Corrections Research*, 6, 26–31.

Taylor, R. (1999). *Predicting reconviction for sexual and violent offenders using the revised Offender Group Reconviction Scale*. Research Findings 104, London: Home Office.

Walker, N., Farrington, D., & Tucker, G. (1981). Reconviction rates of adult males after different sentences. *British Journal of Criminology*, 21, 357–360.