

## An Analysis of Dangerous Sexual Offender Assessment Reports: Recommendations for Best Practice

Dominic J. Doyle\*, James R. P. Ogloff and Stuart D. M. Thomas

*Centre for Forensic Behavioural Science, School of Psychology and Psychiatry, Monash University and Victorian Institute of Forensic Mental Health (Forensicare), Clifton Hill, Victoria, Australia*

The operation of Australia's preventive detention legislation depends upon forensic clinician assessments of risk for future sexual offending. However, to date, no information is available regarding how such assessments of risk are being conducted. This study provides the first descriptive analysis of the risk assessment practices of mental health professionals conducting assessments under preventive detention legislation around Australia. Eighty-six forensic evaluation reports on 56 sex offenders subject to preventive detention proceedings were obtained and analysed. Overall, the findings are mixed. Positively, valid structured risk assessment tools were commonly utilized. Also, there was good agreement between experts on the final risk assessment outcome, suggesting a consensus in relevant areas relating to risk assessment. However, a number of concerning results were also found (e.g., some evaluators adopted invalid risk assessment methodologies; others incorrectly applied and interpreted otherwise valid risk tools). Taken together, the findings suggest that the standard of practice of risk assessment must be raised. Recommendations for best practice are proposed.

**Key words:** best practice; legislation; preventive detention; risk assessment; sex offenders.

“Better get a lawyer son. Better get a real good one.”

(Cruel Sea, 1995)

The assessment of risk for future violence is central to many decisions made within the criminal justice system (McSherry & Keyzer, 2009; McSherry, Keyzer, & Freiberg, 2006). Decisions with regard to bail applications, sentencing, parole, and conditions of release from custody may all be affected by the offender's perceived level of risk for violence (Glazebrook, 2010). Accordingly, the law often turns to clinicians – particularly psychiatrists and psychologists – for opinions on the level of risk for violence

posed by an offender (Ogloff & Davis, 2005). Recently, however, clinician opinions of risk for future violence have been afforded an even greater responsibility following the enactment of unique legislation targeting sexual offenders.

In recent years, a number of Australian states have introduced laws allowing for the continued detention or community supervision of sex offenders whose sentences have expired but who are still considered to be dangerous (Doyle & Ogloff, 2009; Sentencing Advisory Council, 2006). The enactment of these preventive detention measures continues an international trend in the proliferation of

---

Correspondence: Dominic J Doyle, Centre for Forensic Behavioural Science, Monash University, 505 Hoddle Street, Clifton Hill Victoria 3068, Australia. Email: dominic.doyle@med.monash.edu.au

legislation designed to reduce risks of sexual recidivism (Mercado & Ogloff, 2007; Vess, 2009b). Most recently, the government in Victoria has introduced legislation expanding the scope of the post-sentence community supervision provisions for sex offenders to allow for their ongoing detention in prison (Victoria, Legislative Assembly, 2005).

In contrast to traditional criminal justice principles, preventive detention legislation is predicated not on the crimes previously committed by an offender and tied to a finding of guilt, but rather on the offender's *risk* to commit other sexual crimes in the future. This shift in focus from previous offending to the risk of future offending has placed the clinical risk assessment as the paramount consideration in preventive detention proceedings (Glazebrook, 2010). Indeed, under the legislation, psychologists and psychiatrists are required to prepare reports that assess the level of risk or likelihood that the offender would commit further sexual offences if not detained in prison or supervised for an extended period upon release. Moreover, the courts are statutorily required to take into account this clinical assessment of risk in deciding whether to preventively detain or supervise the nominated offender (Doyle & Ogloff, 2009). While the experts' reports are not necessarily dispositive (Freckelton & Selby, 2009), very often the court's decision of whether to impose the order turns critically upon the clinician's opinion of risk for future sexual violence (Ogloff & Doyle, 2009).

The law's dependence on risk assessment for the operation of this legislation places a considerable burden on the clinician and raises expectations that are perhaps impossible to attain. Accordingly, concerns held by mental health professionals regarding the validity and precision of risk assessment approaches and technologies have intensified (Ogloff & Doyle, 2009), and, in turn, raised doubts about the

appropriateness of this legislation (Doyle & Ogloff, 2009).

However, irrespective of these misgivings, numerous preventive detention proceedings have occurred nationally and hundreds of risk assessment reports have been tendered to the courts. Yet, to date, nothing is known about how clinicians go about the task of assessing risk for future sexual violence in Australia for no systematic information has heretofore been analysed and published.

The current study presents the findings of a descriptive analysis of risk assessment reports prepared by mental health professionals pursuant to Australia's preventive detention legislation. It is vitally important to establish an understanding of clinicians' risk assessment practices within this high-stakes legal context to ensure that legal decision-makers are provided with the highest quality of expert opinion on risk and to preserve and reinforce professional standards.

This article will first provide an overview of current approaches to risk assessment. While a comprehensive review of the risk assessment literature is beyond the scope of this article, some contemporary approaches will be highlighted. Second, some of the theoretical and practical issues that limit the precision of risk assessment will be outlined. Finally, the descriptive analysis of Australian forensic clinicians' dangerous sexual offender assessment reports will be presented. A number of recommendations for best practice in the assessment of risk for future sexual violence will be proposed.

### ***Contemporary Approaches to Sex Offender Risk Assessment***

Within the last 15 years, substantial research efforts to develop and enhance risk assessment technologies have resulted in the development of numerous formal tools for assessing risk for future sexual violence

(Douglas & Skeem, 2005; Ogloff & Daffern, 2004). These can be divided into two broad camps: actuarial models and structured professional judgement (SPJ).

#### *Actuarial prediction*

Actuarial tools comprise variables that have been found to have a statistical relationship to subsequent offending (Ogloff & Davis, 2005). The final actuarial model consists of the combination of risk factors that demonstrated the strongest statistical relationship to the predicted outcome (i.e., sexual offending). The Static-99 (Hanson & Thornton, 1999), a 10-item instrument, is one of the most popular actuarial tools for the prediction of future sexual offending (Hanson, Morton, & Harris, 2003). Numerous validation studies, and, more recently, a meta-analysis of the accuracy of risk assessment instruments, demonstrate that this instrument reliably provides assessments of risk with a moderate degree of accuracy (Barbaree, Seto, Langton, & Peacock, 2001; Hanson & Morton-Bourgon, 2007; Långström, 2004).

#### *Structured professional judgement*

SPJ instruments consist of empirically informed professional guidelines to assist clinicians to develop an assessment of risk (Hart, Kropp, & Laws, 2003). Similar to actuarial prediction, SPJ tools also consist of risk factors derived empirically and rationally from the research literature. However, in contrast to the actuarial model, rather than summing the items in a mechanical fashion, clinicians formulate a structured clinical opinion of low, moderate, or high risk (Davis & Ogloff, 2008). The SPJ approach takes into account both historical and dynamic risk factors, and allows clinicians to utilize their professional judgement within a structured framework, so that idiosyncratic but

important characteristics of the individual that pertain to risk are considered.

The Sexual Violence Risk-20 (SVR-20; Boer, Hart, Kropp, & Webster, 1997) and the Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003) are examples of sexual risk instruments based on the SPJ model. Due to their somewhat recent development, the SPJ approach has only been evaluated in a handful of studies; although this research has generally been quite promising (e.g., Craig, Browne, & Stringer, 2004; de Vogel, de Ruiter, van Beek, & Mead, 2004; Hanson & Morton-Bourgon, 2007; Macpherson, 2003).

Despite the often heated debate in the literature regarding the relative merits and predictive superiority of actuarial and SPJ methods (see Harris, Rice, & Quinsey, 2008; Hart, Michie, & Cooke, 2007; Quinsey, Harris, Rice, & Cormier, 2006), both approaches have comparable predictive validity (Hanson & Morton-Bourgon, 2007).

As noted, the field of risk assessment has advanced considerably in recent years. Indeed, the effect size for violence risk assessment is now superior to that of many other medical and psychological practices (Davis & Ogloff, 2008). Nevertheless, despite these advances, the assessment of risk is a complex task and there remain theoretical and practical limitations on effective prediction in the individual case (Mullen & Ogloff, 2009).

#### *A Cautionary Tale*

A full critique of the issues that limit the reliability and validity of risk assessment is beyond the scope of this article (see Ogloff & Doyle, 2009). However, some salient concerns will be briefly reviewed; for the limitations of risk assessment are as relevant as the very outcome of the assessment itself. From the outset, the practical issue of the base rate of sexual reoffending serves to curtail the precision

of risk assessment (Wollert, 2006). As explained in greater detail elsewhere (see Mullen & Ogloff, 2009; Ogloff & Davis, 2005), the less common the future behaviour under prediction in the population, the less accurate the predictions. And, contrary to the popular opinion, sexual reoffending is not a high-frequency occurrence (Doyle & Ogloff, 2009). Indeed, sexual recidivism research consistently finds that, as a group, most sex offenders do not go on to sexually reoffend (e.g., a meta-analysis of sexual recidivism including 30,000 sex offenders found an average recidivism rate of 13.7% over 5–7 years) (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2005). Given this relatively low base rate of sexual recidivism and the practical difficulties in predicting a low base rate event, attempting to predict who will commit further serious sexual offences will inevitably be accompanied by false accusations (Doyle & Ogloff, 2009).

A significant issue that compromises the validity of actuarial instruments concerns the unreliability of applying the group-based risk evaluation of an actuarial tool to the assessment of risk in the individual case (Berlin, Galbreath, Geary, & McGlone, 2003; Hart et al., 2007; Mullen & Ogloff, 2009). For instance, if an offender scores 6 on the Static-99 instrument he is considered to be in the “high-risk” category, 52% of whom (in the original sample) were known to reoffend throughout a 15-year follow-up. However, the instrument cannot specify whether the “high” risk offender belongs to the 52% of people in this category who sexually reoffended, or to the 48% of people who did not (Berlin et al., 2003). Therefore, an individual’s score on the actuarial tool fails to be a reliable guide to the individual’s specific risk to sexually reoffend, for the simple reason that actuarial methods are not designed to assign levels of risk to individuals but to groups (Mullen & Ogloff, 2009).

Other considerations that arguably weaken the validity of actuarial models of risk assessment have been detailed elsewhere (see Doyle & Ogloff, 2009; Ogloff & Doyle, 2009), and include the tools’ insensitivity to the impact of age on risk, the lack of parallel between the legal question to be answered and the evaluative results of the instruments, and the limited data validating the instruments’ use locally.

Importantly, however, these limitations also emphasize that when mental health professionals are asked to provide opinions of risk for future violence to courts, or other decision-making bodies, it is incumbent upon the clinician to be clear to state the limitations to the science upon which their findings rest (Mullen & Ogloff, 2009).

### *The Present Study*

The number of publications concerned with the role of risk assessment under preventive detention legislation is considerable and continues to grow (i.e., Keyzer, Pereira, & Southwood, 2004; McSherry, 2005; McSherry et al., 2006; Mercado & Ogloff, 2007; Mullen & Ogloff, 2009; Ogloff & Doyle, 2009; Scott, 2008; Sentencing Advisory Council, 2006; Smallbone & Ransley, 2005; Vess, 2009a, 2009b; Wood & Ogloff, 2006). Despite this attention being paid to the issue, to date, there has been no empirical evaluation regarding *how* such assessments of risk are being conducted under these laws.

Commonly, investigations of clinical practice are achieved by surveying the professionals (e.g., Allan, Martin, & Allan, 2000; Martin, Allan, & Allan, 2001; Mercado, Elbogen, Scalora, & Tomkins, 2001). However, a limitation to this methodology is its vulnerability to self-report biases. A more objective assessment of how forensic clinicians go about the task of risk assessment is achieved via an analysis of the actual reports that they tender to the court. While content analyses of reports produced

under equivalent legislation in the United States have been conducted (see Amenta, 2005; Levenson, 2004), the present study represents the first analysis of reports prepared by mental health professionals conducting assessments of risk for sexual recidivism under Australia's preventive detention legislation. This research seeks to develop a greater understanding of the methodology and standard of practice among forensic clinicians providing expert evidence to the courts within this particular legal context.

We believe this investigation is important for a number of reasons. First, this research will enable a measure of the quality of expert opinion on risk being provided to legal decision-makers. Second, if it is found that the practice of risk assessment is inadequate then this will need to be immediately remedied given that compromised risk assessments have significant implications for public safety, the civil liberties of offenders, and the integrity of the professions to which the evaluators are ascribed. In light of these considerations, the value of an analysis of the state of forensic practice in this legal area is clear.

## Method

### *Sample*

Eighty-six forensic evaluation reports on 56 sex offenders subject to preventive detention legal proceedings were made available to the researchers. The sample consisted of 27 (31.4%) reports from Victoria, 33 (38.4%) reports from Western Australia, and 26 (30.2%) reports from New South Wales. The reports ranged in date from May 30, 2005 to February 2, 2009. Consistent with statutory language all subjects in New South Wales and Western Australia were examined by two forensic evaluators, though at the time of data collection in three cases only one psychiatric report was made available

to the researchers. In Victoria two applications were accompanied by two assessment reports. In summary, the researchers were provided access to 60 reports produced in relation to 30 offenders and 26 reports written in relation to 26 offenders.

### *Report Coding Procedure*

Data were recorded via review of the aforementioned forensic evaluation reports. Based upon a comprehensive review of the literature, the authors developed a coding manual detailing the scoring criteria for variables of interest. The manual was modelled on two other coding instruments used in previous report analyses (Amenta, 2005; Skeem, Golding, Cohn, & Berge, 1998) and underwent several iterations and a pilot analysis. This analysis resulted in a refining of the coding instrument to ensure that consistency of information was obtained across reports (a copy of the data collection form is available from the authors upon request).

### *Data Collection Procedure*

To obtain access to the forensic reports, written requests were sent to the Chief Justices of the Supreme Courts of Queensland, Western Australia, and New South Wales. In Victoria, although both the Supreme and County Courts are eligible to hear post-sentence proceedings, the lower court has heard the vast majority of post-sentence applications. Accordingly, a request was only sent to the Chief Judge of the County Court of Victoria. Although authorities in Queensland declined participation, all other states consented. Upon receipt of the reports, raw, de-identified data were transcribed onto the coding manual. Unique identifiers were assigned to the offender, the reports, and the evaluator. Full ethical approval from Monash University was received.

The reports typically included psychosocial, criminal history, diagnosis, and risk assessment information. Specifically, the information of interest to the researchers were organized around three central themes. First, the rater coded (a) the examiners' general assessment practices (e.g., number and length of interviews), (b) whether the examiner included the referral source and purpose of the assessment, and (c) whether the examiner described notifying the defendant about the purpose and confidentiality of the evaluation and documented their consent to proceed.

Second, the rater coded (a) the types and number of risk assessment methodologies and instruments employed, (b) the manner in which risk assessment results were communicated, (c) the method of communicating the final opinion of risk, (d) the nature of any statements of limitations pertaining to the practice of risk assessment, and (e) if examiners identified factors associated with risk outside of a structured assessment, whether such factors have either robust or equivocal support in the literature on sexual recidivism.

Third, for the purposes of assessing inter-rater reliability, the rater coded (a) Axis I and Axis II diagnoses, (b) risk scores on the Static-99 and PCL-R, and (c) the final risk rating provided by each examiner.

### **Analyses**

Similar to previous analyses of psychiatric and psychological reports both domestically (Allan et al., 2000; Martin et al., 2001) and internationally (Amenta, 2005; Heilbrun & Collins, 1995; Heilbrun, O'Neill, Strohmman, Bowman, & Philipson, 2000; Larkin & Collins, 1989; Petrella & Poythress, 1983; Skeem et al., 1998), the findings of this study were predominantly descriptive. As the first empirical analysis of clinicians' risk assessment practices under Australia's preventive detention legislation, this was deemed appropriate.

Inter-rater reliability was examined on the total score of the Static-99 using the intra-class correlation coefficient (ICC) two-way model for continuous variables (Bartko, 1966; Shrout & Fleiss, 1979). The PCL-R was inconsistently reported as a numerical value, category, or percentile between evaluators. Therefore, all scores and percentiles were transformed into a categorical rating of low (less than 20), medium (between 20 and 30), or high (more than 30). Thus, the PCL-R inter-rater reliability was assessed using kappa coefficients.

DSM-IV (American Psychiatric Association, 2000) diagnoses and final risk rating were dichotomous variables and were also assessed using kappa coefficients. Kappa coefficients were computed when a  $2 \times 2$  table was attained. Given the relatively small number of cases included in the reliability analyses, and the disproportionate impact on kappa values this can have, levels of agreement (%) are also provided. While different interpretations of reliability coefficients exist, this study, consistent with Levenson's (2004) approach, adopted a higher standard given the seriousness of the decisions being made in this legal context. For this study, a reliability coefficient below .60 is considered poor, .60 to .74 is considered fair, and .75 to 1.0 is considered good (Bloom, Fischer, & Orme, 1999).

## **Results**

### **Participants**

#### *Report authors*

Twenty-three mental health professionals authored 86 reports. Sixteen psychiatrists authored 60 (69.8%) reports, with 14 (87.5%) psychiatrists indicating that they had a specialization in the forensic field. Seven psychologists authored 26 (30.2%) reports. Of those with psychology training, four (57.1%) evaluators had received post-graduate qualifications (i.e., Doctor of

Psychology), while the highest level of qualification for three<sup>1</sup> (42.9%) evaluators was Honours or Graduate Diploma in Psychology. The number of reports per evaluator ranged from 1 to 9. Sixteen evaluators (70%) had five or fewer reports included in the sample and seven evaluators (30%) had between five and nine reports.

#### *Offenders under evaluation*

The demographic, clinical, and criminal characteristics of those subject to these forensic evaluations have been reported elsewhere (Doyle, Ogloff, & Thomas, in press). Briefly, their mean age was 44.7 years ( $SD = 14.2$ ), and 10 (17.9%) were known to be in a relationship at the time of the legal proceedings. The majority ( $n = 45$ , 80.4%) had less than a high school education. Almost two thirds ( $n = 35$ , 62.5%) were reported to have had a substance abuse problem throughout their lifetime, and more than two thirds ( $n = 38$ , 67.9%) were currently diagnosed with an Axis I disorder. Thirty-five (62.5%) received a paraphilia diagnosis, the most common of which was a diagnosis of paedophilia ( $n = 29$ , 51.8%). More than half ( $n = 30$ , 53.6%) were diagnosed with a personality disorder, with antisocial personality disorder ( $n = 20$ , 35.7%) being the most common.

The vast majority of the offenders ( $n = 51$ , 91.1%) had previous convictions for sexual offences prior to their index offence, with a mean number of prior sentencing dates of 2.91 ( $SD = 2.7$ ). There mean age at conviction for first sex offence was 23.8 years ( $SD = 11.0$ ), with over one third ( $n = 20$ , 35.7%) committing their first sexual offence prior to the age of 18 years. Twenty-five (44.6%) offenders had a history of offending against male victims, while the majority ( $n = 47$ , 83.9%) had a history of sexual offences against female victims. Most offenders had a history of offending outside the family ( $n = 53$ ,

94.6%) and a history of violent sexual offending ( $n = 33$ , 58.9%). The majority ( $n = 51$ , 91.1%) also had prior convictions for non-sexual offences.

#### **Report Writing Characteristics**

##### *General assessment practices*

All reports indicated the number of interviews conducted with the offender; the average was 1.62 ( $SD = 0.85$ ) but ranged from 1 to 5. A majority of reports ( $n = 79$ , 91.9%) were based on either one or two interviews. The total length of the interviews was noted in 67 (77.9%) reports, with a mean length of 234.81 minutes (3.9 hrs) ( $SD = 129.24$ ), ranging from 90 minutes (1.5 hrs) to 645 minutes (10.75 hrs).

##### *Inclusion of information*

Almost two thirds ( $n = 56$ , 65.1%) of reports included a statement that identified the authority that requested the evaluation, while 30 reports (34.9%) omitted this information. The purpose of the assessment was clearly articulated in 17 (19.8%) reports. The majority ( $n = 66$ , 76.7%) indirectly referred to the reason for assessment by reference to the legislation, while 3 (3.5%) reports omitted any reference to the purpose of the assessment. All reports indicated that the author had engaged in a review of collateral information.

##### *Documentation of notification*

Less than two thirds of the reports ( $n = 53$ , 61.6%) included a statement that the offender was notified regarding the limits to confidentiality. Consent to participate in the assessment was documented in 51 (59.3%) reports, while 39 (45.3%) reports included a statement that the offender was told the nature and purpose of the evaluation. That the offender understood the information contained within the notification was documented in 46 (53.5%) reports.

**Risk Assessment Practices**

*Risk assessment methods*

Table 1 presents the type, frequency, and combination of risk assessment methods employed by the evaluators. Multiple methods of risk assessment were regularly utilized, with 2 ( $n = 35, 40.7\%$ ) and 3 ( $n = 35, 40.7\%$ ) methods being the most common. The methods of unstructured clinical judgement and adjusted actuarial, were, respectively, used in 18 (21%) and 24 (27.9%) reports. The most common combination of methods comprised actuarial and SPJ ( $n = 25, 29.1\%$ ).

*Risk assessment tools*

A range of risk assessment instruments derived from actuarial, adjusted, and SPJ methodologies were employed by the evaluators (see Table 2). The Static-99 was clearly the most frequently used risk

assessment tool. The SPJ tools of the SVR-20 and RSVP were utilized in 44 (51.1%) reports.

*Reporting and interpreting the Static-99*

Table 3 presents frequency data across a number of dimensions relevant to reporting Static-99 results and interpreting the tool's probability estimates. Of those

Table 1. Frequency and type of risk assessment methods employed by evaluators.

Risk assessment method	N	%
Actuarial Alone	3	3.5
Actuarial + Dynamic <sup>a</sup>	2	2.3
Actuarial + SPJ <sup>b</sup>	25	29.1
Actuarial + Dynamic + SPJ	5	5.8
Actuarial + Adjusted + SPJ	4	4.7
Clinical Judgement Alone	5	5.8
Clinical Judgement + Actuarial	8	9.3
Clinical Judgement + Actuarial + Dynamic	1	1.2
Clinical Judgement + Actuarial + Empirically Guided	1	1.2
Clinical Judgement + Actuarial + Adjusted + SPJ	3	3.5
Empirically Guided Alone	1	1.2
Empirically Guided + Actuarial	4	4.7
Empirically Guided + Actuarial + Adjusted	17	19.8
Empirically Guided + Actuarial + SPJ	7	8.1

<sup>a</sup>Dynamic refers to tools that, though consisted of dynamic variables, were not used to adjust the assessment of risk based on historical factors. <sup>b</sup>SPJ is an abbreviation for Structured Professional Judgement.

Table 2. Frequency and type of risk assessment tools employed by evaluators.

Risk assessment tool	N	%
Static-99	79	91.9
RSVP	31	36.0
SONAR	23	26.7
SVR-20	13	15.1
3-Predictor Model	9	10.5
RRASOR	4	4.7
SORAG	1	1.2
HCR-20	12	14.0
PCL-R <sup>a</sup>	46	53.5

<sup>a</sup>Although not designed to be a risk assessment tool, the PCL-R has been reliably associated with both violent and sexual recidivism.

Table 3. Reporting and interpreting Static-99 results.

Components of Static-99 reporting	N <sup>a</sup>	% <sup>b</sup>
<b>Inclusion of Static-99 results</b>		
Total Static-99 score	52	65.8
Probability estimate	66	83.5
Base rate data	7	10.6
<b>Interpretation of the probability estimate</b>		
Group-based risk	38	57.6
Individual's risk	10	15.2
Unclear/contradictory	15	27.3
<b>Errors in reporting probability estimates</b>		
Incorrect recidivism percentages	14	21.2
Uncollapsed recidivism percentages	5	7.6
<b>Impact of offender's age on Static-99 result</b>	15	19

<sup>a</sup>Number of reports that included the described information. <sup>b</sup>Refers to the percentage of reports that provided the relevant information relative to those reports that used the Static-99 tool.



reports that used the Static-99 risk assessment, the majority included the probability estimate ( $n = 66, 83.5\%$ ). A number of reports contained errors in the reporting of the evaluative results ( $n = 19, 28.8\%$ ).

*Stating the limitations of risk assessment*

The type and frequency of statements provided by assessors regarding the limitations of the Static-99 risk tool specifically, and the practice of risk assessment generally are listed in Table 4.

Static-99 specific statements of limitations were provided in 46 (58.2%) reports, ranging from 0 to 5. The limitations most commonly stated concerned the difficulties in applying the group estimate of risk to the individual case ( $n = 36, 45.6\%$ ).

A general statement of the limits to the practice of risk assessment was provided in 39 (45.3%) reports, ranging from 0 to 3. The statement most

commonly provided concerned the limited accuracy of risk assessment ( $n = 29, 33.7\%$ ).

*Risk factors*

In 52 (60.5%) reports, evaluators identified factors outside of a formal (i.e., instrument-based or empirically guided) risk assessment procedure, that they considered to be associated with an elevated risk for reoffence. Table 5 lists those risk factors identified by evaluators that do have empirical support for being associated with recidivism risk. Prior sex offences ( $n = 27, 31.4\%$ ), deviant sexual preferences ( $n = 31, 36.0\%$ ) and lack of social/familial/community support ( $n = 22, 25.6\%$ ), were the individual risk factors most commonly identified by evaluators and supported by the sexual recidivism literature to be associated with risk of sexual reoffending.

Additionally, factors that were identified by evaluators external to a formal risk assessment procedure as being associated with risk, but that do not have strong empirical support, are presented in Table 6. Thirteen (56.5%) evaluators wrote 28 (32.6%) reports within which they identified such “risk” factors. Minimizing culpability ( $n = 17, 19.7\%$ ), and denial ( $n = 9, 10.5\%$ ) were the two factors most commonly believed to be associated with sexual recidivism that are currently lacking in empirical support.

*Communication of final risk rating*

A final opinion of risk was provided in 79 (91.9%) reports. Of those reports to include a final opinion, all but one ( $n = 78, 98.7\%$ ) utilized a categorical method of risk communication (i.e., high, moderate, low). The majority of assessments concluded the offender posed a high risk ( $n = 64, 74.4\%$ ) of sexual reoffending. Five reports (5.8%) provided a risk rating

Table 4. Frequency of statements of limitations provided for Static-99 and general risk assessment.

Statement of limitation	N	%
<b>Static-99 risk assessment</b>		
Moving from group to individual estimations of risk	36	45.6
Absence of dynamic factors	23	29.1
Not validated on Australian sex offenders	18	22.8
Accuracy of the instrument	13	16.5
Ethical issues regarding its use in the legal context	11	13.9
Dissimilar definition of sex offence between Static-99 and legislation	6	7.6
Not validated on indigenous sex offenders	3	3.8
<b>General risk assessment</b>		
Accuracy of risk assessment	29	33.7
Limitations to the science of risk prediction	18	20.9
Not entirely objective process	9	10.5
Tools not validated for use in Australia	2	2.3

Table 5. Frequency of empirically supported risk factors.

Factors	N	%
<b>Demographic factors</b>	4	4.7
Age	1	1.2
Marital status	2	2.3
Employment history	3	3.5
<b>Criminal history</b>	22	25.6
Total number of prior offences	8	9.3
History of rule violation	13	15.1
<b>Sexual criminal history</b>	29	33.7
Prior sex offences	27	31.4
Stranger victims	6	7.0
Extrafamilial victims	9	10.5
Early onset of sexual offending	1	1.2
Male victims	11	12.8
Diverse sexual crimes	10	11.6
<b>Sexual deviance</b>	32	37.2
Deviant sexual preferences	31	36.0
Sexual preoccupations	5	5.8
<b>Personality disorder</b>		
Psychopathy	2	2.3
Antisocial personality disorder	6	7.0
Any personality disorder	1	1.2
<b>Treatment history</b>	7	8.1
Failure to complete treatment	3	3.5
Failure to participate in treatment	4	4.7
<b>Dynamic factors</b>	41	47.7
Sexual attitudes tolerant of sexual violence	7	8.1
Intimacy deficits	14	16.3
Lack of appropriate adult sexual relationship	11	12.8
Impulsivity	12	14.0
Substance abuse	14	16.3
Circumstances post release (e.g., release plans)	8	9.3
Lack of social/familial/community support	22	25.6
Psychological problems (i.e., negative mood)	5	5.8

of “very high.” Risk ratings of moderate-high, moderate, and moderate-low were, respectively, noted in 5 (5.8%), 2 (2.3%), and 3 (3.5%) reports. A risk rating of “likely” was provided in one report (1.2%).

In addition to describing the subject as “high risk” for future sexual offending, 16 (18.5%) reports also described the offender as being “some risk,” a “significant risk,” a “virtually certain risk,” and “very,”

Table 6. Frequency of identified risk factors with equivocal support.

Factors	N	%	d <sup>a</sup>
Minimizing culpability	17	19.7	.00
Denial	9	10.5	-.02
Low treatment motivation	8	9.3	-.02
Victim empathy	7	8.1	-.01
Low intelligence	7	7.0	.04
Prior history of violent offending	6	7.0	.01
Victim of sexual abuse	3	3.5	.02
Degree of force used	2	2.3	.00
Degree of sexual contact	1	1.2	-.16
Adverse childhood environment	1	1.2	.00

<sup>a</sup>The standardized mean difference statistic is taken from Hanson and Morton-Bourgon’s (2004) updated meta-analysis of predictors of sexual recidivism. According to Cohen (1988), d values of .20 are considered small. The value of d is approximately twice as large as the correlation coefficient calculated from the same data.

“significantly,” and an “unacceptably” high risk.

**Inter-Rater Reliability**

*Risk scores, diagnoses, and final risk ratings*

All inter-rater reliability results are displayed in Table 7. Analysis of the Static-99 produced an ICC coefficient of .85. While the level of agreement was moderate (65%), the inter-rater reliability (Pearson  $r = .81$ ) was high. Static-99 scores differed in seven cases; in all but one the difference in Static-99 scores did not correspond with a difference in the associated risk rating. The PCL-R was used by both evaluators in only 10 cases with good levels of agreement (70%), but poor reliability ( $\kappa = .46$ ).

Reliability of Axis I diagnoses ranged from good to excellent. Paedophilia demonstrated excellent reliability ( $\kappa = .93$ ), as did psychotic disorders ( $\kappa = 1.0$ ). The reliability of Axis II diagnoses were poor, except for psychopathy ( $\kappa = 1.0$ ).

The level of agreement between evaluators on the final risk rating was very good (84.7%).

Table 7. Inter-rater reliability coefficients.

Measures and variables	<i>N</i>	Agreement %	ICC	Kappa <sup>a</sup>
<b>Risk assessment instruments</b>				
Static-99	20	65.0	.81	
PCL-R	10	70.0		.46
<b>Diagnoses</b>				
Paedophilia	30	96.7		.93
Other paraphilias	30	96.7		
Multiple paraphilias	30	93.3		.76
Psychotic disorder	30	100.0		1.0
Other Axis I disorder	30	96.7		.78
Antisocial PD	30	83.3		.65
Psychopathy	30	100.0		1.0
Personality disorder NOS	30	73.3		.19
Other personality disorder	30	86.7		-.053
<b>Final risk rating</b>				
	24	84.7		
High <sup>b</sup>	24	79.2		.42
Moderate <sup>c</sup>	24	79.2		.32
Moderate low		95.8		

<sup>a</sup>Kappa values were only available when a 2 × 2 table could be attained. <sup>b</sup>Very high and high ratings were merged to form a single rating of high for the purposes of cross-tabulation. <sup>c</sup>Moderate-high and moderate ratings were merged.

## Discussion

The assessment of an offender's risk for further sexual offending is central to preventive detention proceedings. This investigation provided the first descriptive analysis of forensic evaluators' risk assessment practices and the reliability of risk assessment outcomes in these legal matters. Despite the descriptive nature of this study, some important conclusions can be drawn, and practical recommendations made.

Taken together, the findings of this investigation are mixed. That the majority of clinicians employed valid structured tools to assess future SVR is encouraging, and indicates a significant translation of empirical research into clinical practice. Encouraging too was that there was good agreement between the experts on the final risk assessment outcomes, suggesting a consensus in relevant areas relating to risk assessment. Despite these relatively positive findings, however, a number of disconcerting results were also found. For example, some evaluators adopted invalid risk assessment methodologies. Others incorrectly

applied and interpreted otherwise valid risk tools. Also, the limits that constrain the science of risk assessment were all too infrequently communicated. Given that these legal proceedings involve fundamental questions of individual liberty and public safety, these are egregious errors.

Overall, the findings suggest that the standard of practice of risk assessment must be raised. In what follows the results across the domains of interest will be discussed. Recommendations for best practice will be proposed.

## Report Writing and Assessment Practices

Under the code of ethics applicable to the professions of psychology and psychiatry (Australian Psychological Society, 2007; Royal Australian and New Zealand College of Psychiatrists, 1998), practitioners are obligated to provide the examinee with a notification outlining the nature and purpose of an assessment, the limits to confidentiality that pertain to the assessment, and obtain the person's informed

consent to proceed with the interview. However, a significant proportion of reports failed to document that the various constituents of the notification had taken place. We do not presume that failure to document the notification equates to failure to provide the notification. Nevertheless, the careful documentation of the notification is advised so that the fulfilment of the ethical obligation to notify is formally recorded and the assessor is protected from claims to the contrary.

Contrary to general principles of forensic report writing (e.g., Allnutt & Chappow, 2000), a number of reports failed to identify the authority (e.g., Supreme Court) that requested the evaluation and clearly articulate the reason for the referral. For clarity, the authors recommend that the referrer and the purpose of the assessment be clearly documented.

### ***Risk Assessment Practices***

#### *Methods of risk assessment*

Despite ongoing debate among experts regarding the relative merits of various sex offender risk assessment methods, some broad points of agreement are being reached. These emerging points of agreement are that empirically validated actuarial measures best form the foundation of risk assessment while a structured consideration of dynamic risk factors assist in formulating the nature of the risk presented by the offender and a management strategy to reduce such risks (Vess, 2009b). A significant number of reports approached the task of risk assessment in this way, combining actuarial and SPJ methods. It is positive that their expert opinion has been grounded in the best risk assessment methods available.

Less encouraging was the finding that a number of clinicians utilized an unstructured clinical judgement approach in their assessment of SVR. Simply, the empirical evidence does not support unaided clinical

judgement as a valid method of risk assessment (Hanson & Morton-Bourgon, 2007; Ogloff & Davis, 2005). Some may advocate an unstructured approach to risk assessment is necessary when no relevant structured tools are available (e.g., when required to assess risk for sexual recidivism in women offenders). However, in such cases experts are cautioned from providing an opinion that is without empirical foundation.

Several reports, problematically, presented an opinion on risk based solely upon the results of an actuarial method. The actuarial approach provides a valid, yet incomplete assessment of risk. Even the instrument's authors advise that the Static-99 is not comprehensive because it "neglects whole categories of potentially relevant variables" (Hanson & Thornton, 1999, p. 18). Clinicians are cautioned from relying exclusively upon an actuarial method at the expense of a more comprehensive, multimodal risk assessment procedure.

A significant number of reports utilized the adjusted-actuarial method as part of their risk assessment. While the consideration of dynamic risk variables is relevant to risk assessment (Douglas & Skeem, 2005), the empirical validity of the adjusted-actuarial approach is far from established. Indeed, the development of the dynamic risk instrument SONAR in Hanson and Harris' (2000) research contained limitations such as invalid items, while its refinement in a later study (Hanson, Harris, Scott, & Helmus, 2007) had little effect on its capacity to add incremental validity to actuarial predictions. While the method holds some promise for evaluating changes to an offender's risk, further and better research is needed to justify the use of this method in a legal context.

#### *Risk assessment tools*

As noted, a risk assessment should be based upon the best available methodology (Mercado & Ogloff, 2007). Clearly, meeting

this requirement necessitates the use of the best available risk assessment tools. Positively, the results showed that the majority of clinicians used valid and reliable structured tools, across actuarial and SPJ methodologies.

The Static-99 was utilized in almost all reports, which reflects its status as a well-validated tool with reliably moderate degrees of accuracy (Hanson & Morton-Bourgon, 2007). However, used only in half of the reports were SPJ tools, the RSVP and SVR-20. Although SPJ tools are relatively recent, they have been validated in a number of studies with promising results (Craig et al., 2004; de Vogel et al., 2004; Macpherson, 2003). Indeed, Hanson and Morton-Bourgon's (2007) meta-analysis of the accuracy of risk assessment instruments revealed that the strongest single predictor of sexual recidivism was a measure of SPJ (i.e., de Vogel et al., 2004). The SPJ approach, unlike actuarial tools, can assist the clinician in the formulation of the nature of the risk posed by the offender; it has also been recommended for those wishing to understand their cases in depth (Hanson & Morton-Bourgon, 2007).

Last, a number of assessors utilized the dynamic risk tool SONAR (Hanson & Harris, 2000; Hanson & Harris, 2001). As discussed previously, this risk instrument, and the adjusted-actuarial method to which it belongs, lacks the sufficient empirical base to justify its use in preventive detention proceedings.

#### *Communicating Static-99 results*

The utility of a risk assessment tool is realized only when it is correctly administered, accurately interpreted, and its results are effectively communicated. Concerningly, the results showed that the utility of the Static-99 was too often undermined by inadequate reporting of outcome information, erroneous reporting of its results, and its incorrect interpretation.

First, omitted from a significant number of reports were the probability percentages of recidivism associated with the offender's risk score and the samples' recidivism base rate upon which the probability percentages were determined. This information is imperative to understanding and contextualizing the tool's risk rating; its omission disallows judicial decision makers the necessary information to fully understand the descriptive and relative nature of the Static-99's rating in relation to future sex offending risk.

Several reports also expressed the probability estimates associated with an offender's Static-99 score as indicating the offender's *specific* risk of reoffending (i.e., "There is a 4 in 10 chance that [the offender] will reoffend within 5 years"). More reports still were unclear about this relationship between the probability estimate and the offender's specific reoffence risk (i.e., "[the offender's] risk is quantified as a 40% likelihood of reoffending over 5 years from a sample of similar offenders in Canada and the United Kingdom"). The probability estimate associated with the offender's score refers to the recidivism percentages of a *group* of sexual offenders. To apply this group-based percentage to the individual is wrong. Given the gravity of the decision to be made by court partly, but necessarily, based upon the expert's risk assessment, this incorrect interpretation of the Static-99 risk tool is a glaring error.

A number of mistakes were also made in the direct reporting of the probability estimates themselves. Contrary to the tool's manual (Harris, Phenix, Hanson, & Thornton, 2003), a handful of evaluators applied the uncollapsed recidivism percentages to offenders whose risk score was higher than 6. In so doing, the offender's risk for reoffending has been erroneously inflated. In other reports, recidivism percentages were incorrectly quoted. For example, a number of reports stated the 15-year recidivism estimate for the high-risk

category as “54%” when the instrument’s manuals and publications note this as “52%” (Hanson & Thornton, 1999). In some reports, the probability estimates were rounded up (i.e., from 39% to 40%). While these may be relatively minor errors, given the role that the exact numerical probability estimate may play in the judiciary’s decision on whether the offender meets the threshold level of risk to warrant an order (see *RJE v Secretary to the Department of Justice*, 2008; *TSL v Secretary to the Department of Justice*, 2007), it is a fundamental requirement that the correct percentages are communicated to the courts.

Finally, users of the Static-99 considered the effect of the offenders’ age on the validity of the actuarial assessment on very few occasions. This is contrary to the research evidence. The literature indicates that actuarial instruments insufficiently capture the decline in recidivism risk associated with advanced age (Ogloff & Doyle, 2009). Indeed, adjusted age-related probability estimates have been available for some years (Hanson, 2005), thus clinicians have available to them empirically-supported guidelines to consider the impact of offenders’ age on their risk.

It has been argued that the incorrect use of a recognized risk measure is potentially worse than not using a measure at all for “erroneous and misleading conclusions may be drawn that appear to have the weight of scientific research behind them and therefore carry an undeserved weight in legal proceedings” (Vess, 2009b, p.186). The standard of practice in the use of the Static-99 must be raised.

#### *Stating the limitations to risk assessment*

Limitations associated with risk assessment tools or the risk assessment enterprise more generally, were infrequently stated. There a number of pertinent limitations to the

assessment of risk for future sexual violence, as outlined here and elsewhere (Ogloff & Doyle, 2009). Failure to convey the limits to the technology upon which one’s expert opinion rests, contravenes the professional’s ethical obligations and invites the potential for the court to accord undeserved weight to the risk outcome than is warranted (Glazebrook, 2010).

#### *Risk factors*

A number of reports identified factors outside of a formal risk assessment procedure the author believed to be associated with an increased risk for sexual recidivism. A majority of such factors were static and dynamic factors themed around the offender’s sexual criminal history, sexual deviance, and maladaptive interpersonal functioning and social supports. These factors have consistently been identified as empirically associated with sexual recidivism (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2005), thus reflecting an emerging concordance of opinion on risk. However, conversely, a number of reports also included risk factors that were identified by evaluators as being associated with risk that have equivocal or no empirical support (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2005). For example, a number of reports identified factors such as denial, a lack of victim empathy, low treatment motivation, and history of violent offending, to indicate elevated risk, when in fact such factors do not have a well-accepted empirically supported relationship to the outcome being predicted (Hanson & Bussière, 1998). Simply, one cannot justify the use of such risk factors.

#### *Inter-Rater Reliability*

The inter-rater reliability of risk assessment scores on the Static-99 was good, though somewhat lower than previous research

(Bartosh, Garby, Lewis, & Gray, 2003; Harris, 2003; Levenson, 2004). This indicates that clinicians will consistently provide comparable assessments of the likelihood of sexual reoffence risk based on this instrument. The inter-rater reliability for Axis I paraphilic diagnoses was high; thus clinicians are reliably identifying the important psychopathological constructs that are linked to reoffence risk. Finally, the level of agreement between the evaluators regarding the final risk rating was good. This result is significant for it indicates that there is a consensus in relevant areas relating to risk assessment.

### *Communication of Final Risk Rating*

Almost all reports communicated a final opinion on risk using the categorical method (i.e., high, medium, and low). A uniform approach to risk communication is clearly helpful for the task the court is engaged in. Last, some reports included additional descriptors such as “some risk,” “virtually certain risk,” “unacceptably high,” or “very high” in addition to their conclusion of high risk. These additional statements are ambiguous, potentially misleading, and likely to contribute an unnecessary element of confusion to considerations of risk.

### *Recommendations for Best Practice*

The findings of this investigation suggest there is substantial room for improvement in how clinicians assess risk for sexual violence and communicate their findings. In order to strengthen the reliability and validity of expert opinion in this area and preserve professional standards, a number of recommendations can be put forward.

- (1) The use of an unstructured clinical judgement approach to the assessment of risk is invalid, and therefore has no probative value and

should not be relied upon exclusively in these assessments.

- (2) Clinicians are cautioned against adjusting actuarially derived risk ratings based on dynamic risk variables until more and better research provides an empirically defensible reason for doing so. Dynamic risk factors are useful in understanding offending in an individual but as yet have limited validity in predicting risk over the long-term.
- (3) The actuarial and SPJ methods are valid and complementary approaches to risk assessment. Actuarial assessments are wisely used to anchor the risk assessment, given the empirically robust relationship between static risk factors and future sexual violence. The SPJ method complements this approach by incorporating dynamic and idiographic risk information into a comprehensive evaluation of the possible nature of future sexual violence and provides targets for risk management. It would be appropriate for clinicians to utilize both when assessing risk for future sexual violence in applied assessments such as preventive detention proceedings.
- (4) The types of risk assessment tools recommended for use are simply those that have the greatest evidence base. Accordingly, the Static-99 remains the most reliable and best validated actuarial measure. The RSVP and its predecessor the SVR-20 are very promising SPJ tools and will enable the clinician to understand their cases in much greater detail. The PCL-R is also a valid measure of the construct of psychopathy, which, given its relationship to SVR, requires evaluation in an assessment of risk for future sexual violence.

- (5) The results of a properly conducted risk assessment must be effectively communicated. When reporting the results of the Static-99 clinicians are advised to qualify comparative categorical labels such as high risk with the associated probability estimates of recidivism. Additionally, clinicians must report the correct probability percentages, consider the effect of the offender's age on the validity of the actuarial result, and be very careful to not assign the probability estimate of recidivism to the specific offender.
- (6) The base rate of recidivism associated with an actuarial instrument's test sample must also be communicated. This is because the evidentiary value of the offender's probability of recidivism is dependent on the base rate.
- (7) Clinicians must know and convey the limitations to the state of knowledge in the field of risk assessment. Failure to do so violates ethical obligations and potentially gives the court the wrong impression about the predictive ability of the available technology.
- (8) When communicating final opinions on risk, evaluators should employ the categorical method (e.g., high, medium, low). Defined conventionally, medium/moderate risk would be equal to the base rate of recidivism associated with the individual offender, high risk is significantly higher than the base rate, and low risk is significantly lower than the base rate. Evaluators are also cautioned from describing an offender's risk as "unacceptable," "significant," or "likely." Such terms invite unnecessary ambiguity to the process of determining an offender's risk potential.
- (9) Forensic clinicians need to be reminded that research has identified

numerous factors to be empirically associated with sexual reoffence risk. To identify factors that lack such empirical support is unjustifiable. Clinicians are recommended to remain well versed in the contemporary research literature on risk assessment. Keeping up-to-date with scientific advances and debates within the field will protect the clinicians' opinion on risk from serious criticism and a scathing cross-examination.

#### *Limitations and Future Directions*

This present study has limitations that should be considered when interpreting these findings. First, the sample size was smaller than expected following Queensland's decision not to participate in the study. Therefore, the results apply only to those clinicians preparing reports under this legislation in New South Wales, Western Australia, and Victoria. Second, the disciplines of psychiatry and psychology were not evenly represented and some evaluators authored more reports in the sample than others. Accordingly, some of the results may be more relevant to a particular discipline, or author. Nevertheless, the sample of reports is a valid representation of reports tendered in these matters across the participating jurisdiction. Another limitation was the small number of cases available for the inter-rater reliability analyses.

Additional research is required in this area. An extension of this analysis to Queensland is warranted to provide a truly national assessment of the standard of practice of risk assessment in this legal context. Further, expanding research into the areas of clinicians' decision-making processes in evaluating risk and how outcomes derived from multiple risk assessment tools and methods are integrated into a final opinion on risk, is needed.



### Conclusion

This investigation described the risk assessment practices of forensic clinicians following an analysis of sex offender assessment reports submitted under Australia's preventive detention legislation. As an analysis of clinicians' actual practices, via their reports, this research represents a methodological advancement of previous investigations into the clinical practices of mental health professionals.

The results of this investigation were mixed. On the one hand, the findings indicated that there is a consensus in relevant areas relating to risk assessment resulting in regular agreement between clinicians with regard to the final risk assessment outcome. Furthermore, the consistent use of structured risk tools is an encouraging sign that the hard science of the field is translating into applied practice. However, beyond these relatively positive findings, more concerning results were found. Too many clinicians used unreliable methods of risk assessment, erroneously reported the results of a risk instrument, and failed to effectively communicate risk assessment outcomes. Too few clinicians stated the limitations that pertain to the science of risk assessment. In short, the standard of the practice of risk assessment for future sexual offending must be raised. Clinical modesty and professional rigour are required.

The implications of these findings extend beyond recommendations for improving clinical practice. Preventive detention proceedings involve fundamental questions of human rights and community safety. Under this legislation courts are faced with the unenviable task of balancing the human rights of offenders with the risk to community safety posed by such offenders. A less than competent risk assessment and report unnecessarily complicates this task with potentially deleterious consequences for the public and the offender.

It is hoped that this research will be used by clinicians, judges, and the legal fraternity to raise the level of practice in this area, so that the court can be assisted in its efforts to achieve that balance between the civil liberties of offenders and the right of the public to be kept safe from undue risk.

### Note

1. However, two psychologists were approaching completion of their doctoral degrees in psychology.

### References

- Allan, A., Martin, M., & Allan, M. (2000). Assessments for the courts: A survey of Australian psychologists. *Psychiatry, Psychology, and Law*, 7, 150–158.
- Allnutt, S.H., & Chaplow, D. (2000). General principles of forensic report writing. *Australian and New Zealand Journal of Psychiatry*, 34, 980–987.
- Amenta, A.E. (2005). *The assessment of sex offenders for civil commitment proceedings: An analysis of report content*. Unpublished doctoral dissertation, University of Missouri, Columbia.
- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- Australian Psychological Society. (2007). *Code of Ethics*. Melbourne: Australian Psychological Society.
- Barbaree, H.E., Seto, M.C., Langton, C.M., & Peacock, E.J. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior*, 28, 490–521.
- Bartko, J.J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports*, 19, 3–11.
- Bartosh, D.L., Garby, T., Lewis, D., & Gray, S. (2003). Differences in the predictive validity of actuarial risk assessments in relation to sex offender type. *International Journal of Therapy and Comparative Criminology*, 47, 422–438.
- Berlin, F.S., Galbreath, N.W., Geary, B., & McGlone, G. (2003). The use of actuarials at civil commitment hearings to predict the likelihood of future sexual violence. *Sexual Abuse: A Journal of Research and Treatment*, 15, 377–382.

- Bloom, M., Fischer, J., & Orme, J.G. (1999). *Evaluating practice: Guidelines for the accountable professional* (3rd ed.). Boston: Allyn and Bacon.
- Boer, D.P., Hart, S.D., Kropp, P.R., & Webster, C.D. (1997). *Manual for the Sexual Violence Risk-20*. Vancouver, BC: The British Columbia Institute Against Family Violence.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Craig, L.A., Browne, K.D., & Stringer, I. (2004). Comparing sex offender risk assessment measures on a UK sample. *International Journal of Offender Therapy and Comparative Criminology*, 48, 7–27.
- Davis, M.R., & Ogloff, J.R.P. (2008). Key considerations and problems in assessing risk for violence. In D.C.R. Zukauskienė (Ed.), *Psychology and law: Bridging the gap* (pp. 191–210). Aldershot, UK: Ashgate.
- de Vogel, V., de Ruiter, C., van Beek, D., & Mead, G. (2004). Predictive validity of the SVR-20 and Static-99 in a Dutch sample of treated sex offenders. *Law and Human Behavior*, 28, 235–251.
- Douglas, K.S., & Skeem, J.L. (2005). Violence risk assessment. Getting specific about being dynamic. *Psychology, Public Policy, and Law*, 11, 347–383.
- Doyle, D.J., & Ogloff, J.R.P. (2009). Calling the tune without the music: A psycho-legal analysis of Australia's post-sentence legislation. *Australian and New Zealand Journal of Criminology*, 42, 179–203.
- Doyle, D.J., Ogloff, J.R.P., & Thomas, S.D.M. (in press). Designated as dangerous: Characteristics of sex offenders subject to post-sentence orders in Australia. *Australian Psychologist*.
- Freckelton, I.R., & Selby, H.M. (2009). *Expert evidence: Law, practice, procedure and policy* (4th ed.). Sydney: Thomson-Reuters.
- Glazebrook, S. (2010). Risky business: Predicting recidivism. *Psychiatry, Psychology, and Law*, 17, 88–120.
- Hanson, R.K. (2005). *The validity of Static-99 with older sexual offenders*. Ottawa: Public Safety and Emergency Preparedness Canada.
- Hanson, R.K., & Bussière, M.T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66, 348–362.
- Hanson, R.K., & Harris, A. (2000). *The Sex Offender Need Assessment Rating (SO-NAR): A method for measuring change in risk levels*. Ottawa: Public Safety and Emergency Preparedness Canada.
- Hanson, R.K., & Harris, A.J.R. (2001). A structured approach to evaluating change among sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 13, 105–122.
- Hanson, R.K., Harris, A.J.R., Scott, T.-L., & Helmus, L. (2007). *Assessing the risk of sexual offenders on community supervision: The dynamic supervision project*. Ottawa: Public Safety Canada.
- Hanson, R.K., & Morton-Bourgon, K.E. (2004). *Predictors of sexual recidivism: An updated meta-analysis (User Report 2004–02)*. Ottawa: Public Safety and Emergency Preparedness Canada, Retrieved September 24, 2008 from [http://ww2.pssp.gc.ca/publications/corrections/pdf/200402\\_e.pdf](http://ww2.pssp.gc.ca/publications/corrections/pdf/200402_e.pdf).
- Hanson, R.K., & Morton-Bourgon, K.E. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, 73, 1154–1163.
- Hanson, R.K., & Morton-Bourgon, K.E. (2007). *The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis (User Report 2007–01)*. Ottawa: Public Safety and Emergency Preparedness Canada.
- Hanson, R.K., Morton, K.E., & Harris, A.J.R. (2003). Sexual offender recidivism risk: What we know and what we need to know. *Annals of the New York Academy of Sciences*, 989, 154–166.
- Hanson, R.K., & Thornton, D. (1999). *Static-99: Improving actuarial risk assessments for sex offenders (User Report 1999–02)*. Ottawa: Department of the Solicitor General of Canada.
- Harris, A., Phenix, A., Hanson, R.K., & Thornton, D. (2003). *Static-99 coding rules revised*. Ontario: Solicitor General Canada.
- Harris, G.T. (2003). Men in his category have a 50% likelihood, but which half is he in? Comments on Berlin, Galbreath, Geary, and McGlone. *Sexual Abuse: A Journal of Research and Treatment*, 15, 389–392.
- Harris, G.T., Rice, M.E., & Quinsey, V.L. (2008). Shall evidence-based risk assessment be abandoned? *British Journal of Psychiatry*, 192, 154.
- Hart, S.D., Kropp, P.R., & Laws, D.R. (2003). *The risk for sexual violence protocol (RSVP): Structured professional guidelines*

- for assessing risk of sexual violence. British Columbia, Vancouver: Mental Health, Law, and Policy Institute, Simon Fraser University and British Columbia Institute on Family Violence.
- Hart, S.D., Michie, C., & Cooke, D.J. (2007). Precision of actuarial risk assessment instruments. *British Journal of Psychiatry*, *199*, s60–s65.
- Heilbrun, K., & Collins, S. (1995). Evaluations of trial competency and mental state at time of offense: Report characteristics. *Professional Psychology: Research and Practice*, *26*(61–67).
- Heilbrun, K., O'Neill, M.L., Strohmman, L.K., Bowman, Q., & Philipson, J. (2000). Expert approaches to communicating violence risk. *Law and Human Behaviour*, *24*, 137–148.
- Keyzer, P., Pereira, C., & Southwood, S. (2004). Pre-emptive imprisonment for dangerousness in Queensland under the Dangerous Prisoners (Sexual Offenders) Act 2003: The constitutional issues. *Psychiatry, Psychology, and Law*, *11*, 244–253.
- Långström, N. (2004). Accuracy of actuarial procedures for assessment of sexual offender recidivism risk may vary across ethnicity. *Sexual Abuse: A Journal of Research and Treatment*, *16*, 107–120.
- Larkin, E.P., & Collins, P.J. (1989). Fitness to plead and psychiatric reports. *Medicine, Science, and Law*, *29*, 26–32.
- Levenson, J.S. (2004). Reliability of sexually violent predator civil commitment criteria in Florida. *Law and Human Behaviour*, *28*, 357–368.
- Macpherson, G. (2003). Predicting escalation in sexually violent recidivism: Use of the SVR-20 and PCL: SV to predict outcome with non-contact recidivists and contact recidivists. *Journal of Forensic Psychiatry and Psychology*, *14*, 615–627.
- Martin, M., Allan, A., & Allan, M.M. (2001). The use of psychological test by Australian psychologists who do assessments for the courts. *Australian Journal of Psychology*, *53*, 77–82.
- McSherry, B. (2005). Indefinite and preventive detention legislation: From caution to an open door. *Criminal Law Journal*, *29*, 94–110.
- McSherry, B., and Keyzer, P. (2009). *Sex offenders and preventive detention*. Sydney: The Federation Press.
- McSherry, B., Keyzer, P., & Freiberg, A. (2006). *Preventive detention for 'dangerous' offenders in Australia: A critical analysis and proposals for policy development*: Report to the Criminology Research Council. Retrieved 19 October 2009, from <http://www.criminologyresearchcouncil.gov.au/reports/200405-03.pdf>.
- Mercado, C.C., Elbogen, E.B., Scalora, M., & Tomkins, A. (2001). Judgments of dangerousness: Are sex offenders assessed differently than civil psychiatric patients? *Psychiatry, Psychology, and Law*, *8*, 146–153.
- Mercado, C.C., & Ogloff, J.R. (2007). Risk and the preventive detention of sex offenders in Australia and the United States. *International Journal of Law and Psychiatry*, *30*, 49–59.
- Mullen, P.E., & Ogloff, J.R.P. (Eds.). (2009). *Assessing and managing the risk of violence towards others* (Vol. 2). Oxford, UK: Oxford University Press.
- Ogloff, J., & Daffern, M. (2004). *The dynamic appraisal of situational aggression (DASA)*. Melbourne: Centre for Forensic Behavioural Science, Monash University.
- Ogloff, J.R.P., & Davis, M.R. (2005). Assessing risk for violence in the Australian context. In D. Chappell & P. Wilson (Eds.), *Issues in Australian crime and criminal justice* (pp. 301–338). Chatswood, NSW: LexisNexis, Butterworths.
- Ogloff, J.R.P., & Doyle, D.J. (2009). A clarion call: Caution and humility must be the theme when assessing risk for sexual violence under post-sentence laws. *Sexual Abuse in Australia and New Zealand*, *2*, 59–69.
- Petrella, R.C., & Poythress, N.G. (1983). The quality of forensic evaluations: An interdisciplinary study. *Journal of Consulting and Clinical Psychology*, *51*, 76–85.
- Quinsey, V.L., Harris, G.T., Rice, M.E., & Cormier, C.A. (2006). *Violent offenders: Appraising and managing risk* (2nd ed.). Washington, DC: American Psychological Association.
- RJE v Secretary to the Department of Justice [2008] VSCA 265.
- Royal Australian and New Zealand College of Psychiatrists. (1998). *Code of ethics* (2nd ed.). Melbourne: RANZCP.
- Scott, R. (2008). Case commentary: Risk assessment and sentencing of serious sex offenders. *Psychiatry, Psychology, and Law*, *15*, 188–201.
- Sentencing Advisory Council. (2006). *High-risk offenders: Continued detention and*

- supervision options*. Melbourne, Victoria: Sentencing Advisory Council.
- Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428.
- Skeem, J., Golding, S., Cohn, N., & Berge, G. (1998). The logic and reliability of expert opinion on competence to stand trial. *Law and Human Behaviour*, 22, 519–547.
- Smallbone, S., & Ransley, J. (2005). Legal and psychological controversies in the preventive incapacitation of sexual offenders. *University of New South Wales Law Journal*, 28, 299–305.
- TSL v Secretary to the Department of Justice [2006] VSCA 199.
- Vess, J. (2009a). Fear and loathing in public policy: Ethical issues in laws for sex offenders. *Aggression and Violent Behaviour*, 14, 264–272.
- Vess, J. (2009b). Risk assessment of sexual offenders for extended supervision orders in New Zealand: Basic principles and current practice. *Journal of Child Sexual Abuse*, 18, 174–189.
- Victoria, Legislative Assembly. (2005). Parliament of Victoria Hansard, 24 February 2005. Retrieved 21 March 2010, from <http://tex.parliament.vic.gov.au/bin/texhtml?form=VicHansard.adv>
- Wollert, R. (2006). Low base rates limit expert certainty when current actuarials are used to identify sexually violent predators: An application of Baye's theorem. *Psychology, Public Policy and Law*, 12, 56–85.
- Wood, M., & Ogloff, J. R. (2006). Victoria's Serious Sex Offenders Monitoring Act 2005: Implications for the accuracy of sex offender risk assessment. *Psychiatry, Psychology and Law*, 13, 182–198.

Copyright of Psychiatry, Psychology & Law is the property of Routledge and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.