s by
and
nis-
ion
ıgs;
ıily
will
pe-
nal

RANDY K. OTTO AND KEVIN S. DOUGLAS, EDITORS

HANDBOOK OF VIOLENCE RISK ASSESSMENT

" 2010

Routledge
Taylor & Francis Group
New York London

12
Static-99
An Actuarial Tool to Assess Risk of Sexual and Violent Recidivism Among Sexual Offenders

DANA ANDERSON and R. KARL HANSON

Introduction

In 1988, Joseph Fredericks raped and murdered 11-year-old Christopher Stephenson. In 1994, Jesse Timmedequas raped and murdered Megan Kanka in New Jersey. These horrific crimes were both committed by men who had previously been convicted of sexual offenses, prompting new, widespread legislative initiatives to address the risk of sexual offenders in the community ("Christopher's Law" in Ontario; "Megan's Law" across the United States). Fortunately, sexual murder is rare, and, contrary to popular opinion, only a minority of sexual offenders are known to commit another sexual offense after being caught. The observed recidivism rate is 10% to 15% after five years, increasing to approximately 20% after 10 years (Hanson & Bussière, 1998; Harris & Hanson, 2004) Even considering that many offenses go undetected, it is difficult to argue that all sexual offenders will inevitably reoffend. Some will and others will not.

This chapter will describe one approach to evaluating the recidivism risk of individual sexual offenders, namely, Static-99 (Hanson & Thornton, 1999, 2000). Static-99 is the most widely used measure for the assessment of risk of recidivism among sexual offenders in Canada and the United States (Archer, Buffington-Vollum, Stredny, & Handel, 2006) and is used in jurisdictions as diverse as Taiwan, Israel, Finland, and Singapore. It is commonly used for diverse purposes, including treatment planning (McGrath, Cumming, & Burchard, 2003), community supervision (Interstate Commission for Adult Offender Supervision, 2007), and civil commitment evaluations (Jackson & Hess, 2007).

The purpose of this chapter is to review the development and validation of Static-99 and its application in applied risk assessment. We describe how it could be used on its own, as well as integrated into comprehensive evaluations involving both static and dynamic (changeable) risk factors.

The Development of the Static-99

Rapid Risk Assessment for Sexual Offense Recidivism and Structured Anchored Clinical Judgment Scale

Static-99 was developed by R. Karl Hanson and David Thornton to assess sexual offenders for their risk of sexual and violent recidivism. It was created from the Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR) (Hanson, 1997) and Thornton's Structured Anchored Clinical Judgment (SACJ-Min) scale (described in Grubin, 1998). The RRASOR was developed with the goal of creating "an easily administered scale that was likely to be valid for a range of settings" (p. 4). The RRASOR items were obtained from Hanson and Bussière's (1998)

meta-analytic study of factors related to sexual offense recidivism. In that study, they found that the following easily scored factors were among those most strongly correlated with sexual reoffending: prior sex offenses, stranger victims (of sex offenses), prior offenses (of any type), age, never married, nonrelated victims (of sex offenses), and male victims (of sex offenses). After examining the variables' association with sexual recidivism using stepwise regression, Hanson (1997) retained four variables for the RRASOR: age (of offender), previous sex offenses, presence of unrelated victims of sexual crimes, and presence of male victims of sexual crimes. A rating of 0 or 1 was assigned to 3 of the 4 items, and a rating of 0 to 3 points was assigned to the item "prior sex offenses." The items were then cross-validated on another sample. The total possible score on the RRASOR ranges from 0 to 6. Hanson (1997) concluded that the RRASOR could reasonably be used as a screening tool,* but cautioned that certain important variables were neglected and that the RRASOR is not intended to be used in isolation. Many variables known to predict sexual reoffending, such as phallometrically assessed deviant sexual preferences and psychopathy, were not included in the RRASOR. The RRASOR was developed as a brief, efficient tool, based on items that were easily scored from commonly available records.

David Thornton's Structured Anchored Clinical Judgment Scale (Grubin, 1998; Hanson & Thornton, 2000) utilized a stepwise approach to assessing risk of sexual and violent recidivism. The first step is comprised of historical variables, where one point is awarded for each of the following items: current sexual offense, previous sex offense conviction, current nonsexual violent offense conviction, previous nonsexual violent offense conviction, and three or more previous convictions (of any type). The number of points at the first step determines risk level. The second step includes aggravating factors (male victims of sex offenses, stranger victims of sex offenses, noncontact sex offense, substance abuse, previously in care, never married, deviant sexual arousal, high [>25] score on the Psychopathy Checklist–Revised); the presence of two or more aggravating factors raises the risk level one category. The third stage of this risk assessment method includes variables related to current behavior and treatment completion. Only the items in the first two steps that were subject to cross-validation, referred to as the SACJ-Min, were included in the development of the Static-99.

Despite their similarity, exploratory analyses indicated that RRASOR and SACJ-Min both made incremental contributions to the prediction of sexual recidivism. Consequently, Hanson and Thornton decided to combine the two instruments by retaining 10 nonredundant items: age, intimate relationship history (ever lived with a lover for 2+ years), index (current) nonsexual violent offense conviction, prior nonsexual violent offense conviction, prior sexual offending, general criminal history (sentencing dates), noncontact sexual offense conviction, extrafamilial sex offense victim(s), stranger victim(s) of sex offenses, and male victim(s) of sex offenses (see Figure 12.1).

The items were selected based on their empirical relationships with recidivism, and not on the basis of construct validity. Nevertheless, perusal of the items suggested that they could be grouped into five general categories: sexual deviance, which is indicated by male victims, offender never having lived common-law (or been married) for 2 or more years, and the presence of noncontact sexual offense(s); range of potential victims, indicated by the presence of extrafamilial victims and stranger victims; persistence of sexual offending, indicated by prior sex offenses (and this item is given the highest weighting in the Static-99); antisociality, as indicated by nonsexual violence (index and prior), and number of previous sentencing dates; and age.

* Subsequent research has supported the use of the RRASOR as a reasonable screening tool (average d of .60, based on 11,031 offenders in 34 different studies; see Hanson & Morton-Bourgon, 2009). We do not recommend its use, however, given Static-99's superior predictive validity.

Construction Samples

The first three samples described below were used to develop the scale, and the fourth was used to explore the validity of the findings. Specifically, various combinations of RRASOR and SACJ-Min items were tested in the Pinel, Millbrook, and Oak Ridge samples. Several combinations of items yielded equivalent predictive accuracy, and the 10-item version selected was deemed simpler than the other options. The predictive accuracy of the 10-item version, now called Static-99, was then tested on a completely new sample from Her Majesty's Prison Service that was not part of the development process. Given that there was no shrinkage between the developmental samples and the cross-validation samples, all available offenders were combined to compute

Question Number	Risk Factor	Codes	Score
1	Young	Aged 25 or older	0
		Aged 18 — 24.99	1
2	Ever Lived With	Ever lived with lover for at least two years?	
		Yes	0
		No	1
3	Index Nonsexual Violence — Any Convictions?	No	0
		Yes	1
4	Prior Nonsexual Violence — Any Convictions?	No	0
		Yes	1
5	Prior Sex Offenses	Charges / Convictions	
		None / None	0
		1 to 2 / 1	1
		3 to 5 / 2 to 3	2
		6+ / 4+	3
6	Prior Sentencing Dates (excluding index)	3 or less	0
		4 or more	1
7	Any Convictions for Noncontact Sex Offenses	No	0
		Yes	1
8	Any Unrelated Victims	No	0
		Yes	1
9	Any Stranger Victims	No	0
		Yes	1
10	Any Male Victims	No	0
		Yes	1
	Total Score	Add up scores from individual risk factors	

	POINTS	Risk Category
Suggested Nominal Risk Categories	0,1	Low
	2,3	Moderate–Low
	4,5	Moderate–High
	6+	High

Figure 12.1 Static-99 tally sheet.

the recidivism rate tables. The Oak Ridge sample was not included in the recidivism rate calculations because the survival times (dates of reoffense) were not available for that sample. Recidivism was defined as reconviction data in the three samples used to construct the Static-99 recidivism rate tables in the manual (the use of reconviction data as the definition of recidivism is further discussed in the "Validity" section of this chapter). Institute Philippe Pinel is a maximum-security psychiatric facility near Montreal, Quebec, Canada. The sex offenders used in this study were released between 1978 and 1993. The average age at release for this sample was 36 years. Although many offenders in this sample would have undergone some level of treatment, it is assumed that treatment methods would be relatively ineffective, given the recent advances in treatment approaches with sex offenders, as well as improvements in identifying appropriate treatment targets (see, for example, Hanson & Harris, 1998). For our purposes, they have been considered mainly untreated.

Information regarding recidivism was obtained for 344 offenders, comprised of 70% child molesters. The average years of follow-up data in this sample was four. The case files of the offenders were reviewed to obtain information regarding the predictor variables (i.e., items on Static-99), and records from the Royal Canadian Mounted Police (R.C.M.P.), the federal police service in Canada, were obtained from 1994 to determine recidivism rates up to that time. Some information from the case files in this sample could not be obtained; therefore, some information was extrapolated, aggregated, or omitted, depending on what information was available (see Hanson & Thornton, 1999, for details). Fifteen percent of this sample had recidivated with a new sexual crime, and 21% had recidivated with a violent (including sexual) offense.

Millbrook Correctional Centre was a maximum-security provincial jail located in central-eastern Ontario, Canada. This sample, consisting of 191 child molesters, was comprised of admissions between 1958 and 1974 and followed for an average period of 23 years. The average age of these offenders at release was 33 years. Approximately half of the offenders completed a brief treatment program, which had no observable effect on recidivism rates (Hanson, Steffy, & Gauthier, 1993). Clinical and correctional files were used to obtain information on the predictor variables, but some coding of variables was omitted due to the lack of information on file. As with the sample above, recidivism information (i.e., the criterion variable) was obtained from records from the R.C.M.P. Thirty-five percent of these offenders had reoffended with a new sexual crime and 44% had reoffended with a violent (including sexual) offense.

There were 142 offenders drawn from the population at Oak Ridge Division of the Penetanguishene Mental Health Centre. Oak Ridge is a maximum-security psychiatric hospital north of Toronto, Ontario, Canada. This sample was comprised of approximately 49% child molesters who were referred for psychiatric assessments related to legal proceedings (e.g., fitness to stand trial) or who were referred to the facility by courts, other correctional facilities, or the mental health system. The offenders in this sample were referred to the Centre between 1972 and 1993. The average length of follow-up of the releases was 10 years. Information regarding new convictions was obtained from the R.C.M.P. records. As well, new sexual offense recidivism information, regardless of charges laid, was obtained from mental health records. As with the above samples, some coding of variables was omitted from some cases. The average age at release of this sample was 30 years. Within the follow-up period, 35% of this sample had reoffended with a new sexual crime and 57% had reoffended with any type of violent (including sexual) crime.

The fourth sample, used for the validation of Static-99, was comprised of 531 offenders, approximately 61% of whom were child molesters, released from Her Majesty's Prison Service in England and Wales in 1979. All offenders were tracked for a fixed follow-up period of 16 years. This is considered an untreated sample, and information for all of the predictor variables was available. The average age at release for this sample was 34 years. The sexual reoffending rate of

this sample within the follow-up period was 25%, and 37% of this sample had reoffended with any type of violent (including sexual) crime.

Reliability

In psychological measurement, reliability of a test can defined as the extent to which the observed scores are attributable to the construct of interest ("true score") rather than measurement error. There are several different methods of indexing reliability, including internal consistency, rater agreement, or test-retest consistency. The Static-99 items were derived from empirical data and not from any particular theory that would suggest what the items were measuring. Consequently, the Static-99 items were not expected to "hang together," even though it is possible to interpret item content post hoc. For example, convictions for noncontact offenses and male victims can be considered indicators of sexual deviance; however, these items were not chosen specifically because of their relationship to sexual deviance but for their stand-alone association with the outcome variable of interest—recidivism. Consequently, internal consistency is not an appropriate measure of reliability for Static-99.

For the purpose of the reliability on Static-99, our concern is rater reliability, or the extent to which observers or raters of the Static-99 agree on the scores for the same cases. In both research studies and applied evaluations, the rater reliability has been high (see Harris, Phenix, Hanson, & Thornton, 2003). Barbaree, Seto, Langton, and Peacock (2001), for example, found a correlation of .90 for raters retrospectively coding files for research. When comparing 55 cases rated by independent evaluators for civil commitment hearings in California, Hanson (2001) found an average 91% of agreement on item ratings, an item kappa of .80, and an intraclass correlation for total scores of .87. With high-risk federal offenders in Canada, Looman (2006) found inter-rater reliability of .90 (n = 26) and G. T. Harris et al. (2003) found an intraclass correlation of .87 (n = 10). Hanson, Harris, Scott, and Helmus (2007) found high levels of agreement for the scores computed by trained probation officers in a prospective study (intraclass correlation of .91). Rater reliability in the .90 range means that raters will rarely disagree by more than one point (SD = 1.97; the 95% confidence interval for standard error of measurement is 1.22).

Although inter-rater reliability levels reported have been generally high, this is not always the case. In a recent study, Ducro and Pham (2006) found an oddly low intraclass correlation coefficient of .63. Consequently, it is important to outline some considerations and guidelines for maintaining acceptable levels of agreement between raters. First, it is imperative to have adequate records—particularly for criminal history. The less complete the records, the more raters are required to extrapolate. In the Ducro and Pham (2006) study, standardized criminal history records were not available, and the coders had to infer the criminal history variables from the not-so-clear accounts provided in the clinical records of patients at a forensic psychiatric facility.

Second, procedures are needed to ensure that competency in the scoring criteria is achieved and maintained. In most cases, this involves structured training along with formal procedures for quality assurance (for example, peer review, scoring "parties"). The Static-99 scoring manual is reasonably thorough, but there are many subtle distinctions that are easily confused or forgotten. Consequently, even experienced raters (including the authors) need to regularly review the manual to avoid rater drift. The level of reliability places a limit on the validity, or predictive accuracy, of the tool. If raters disagree on the scoring, then the instrument will not provide the maximum potential level of predictive accuracy within that sample.

Validity

As with reliability, different types of validity are relevant for different types of tests. As noted above, the items chosen for the Static-99 were not chosen to assess latent constructs, but were

chosen based on each item's direct association with recidivism. Notwithstanding the way in which the items were chosen, some exploration of construct validity has been conducted for the items on the Static-99.

A general consensus is developing that the major factors contributing to recidivism among sexual offenders are sexual deviance and antisociality. Hanson and Bussière (1998) summarized the factors related to sexual recidivism based on the data from several studies included in their meta-analysis. Measures of sexual deviancy produced the strongest predictors of sexual offense recidivism, and to a lesser extent, criminal lifestyle variables also predicted sexual recidivism. The factors of antisocial lifestyle and sexual deviancy held as the grouping labels for variables predicting sexual recidivism in an updated meta-analysis (Hanson & Morton-Bourgon, 2004, 2005). Roberts, Doren, and Thornton (2002) examined the structure underlying the relationships among various actuarial risk assessment instrument items and concluded that the components assessed by the instruments (including Static-99) were an antisocial/violence component and a pedophilic deviance/sexual repetitiveness component. Similarly, Barbaree, Langton, and Peacock (2006b) found that most actuarial instrument items load onto factors representing criminality or sexual deviance (which was reflected in four factors representing different aspects of sexual deviance).

Regardless of the types of factors or underlying dimensions that appear to be assessed by the Static-99, evaluators should be most concerned with the tool's predictive accuracy. There have been at least 42 studies (n = 13,288) examining the extent to which Static-99 accurately rank orders the offenders in terms of relative risk to reoffend in independent replication studies (Hanson & Morton-Bourgon, 2007). On average, these studies found predictive accuracy (average d = .70, 95% confidence interval of .64 to .76) similar to that observed in the developmental samples (d = .78). The standardized mean difference (d), is the difference between the recidivists and the nonrecidivists compared to the extent to which individuals within each of those groups vary from each other. The value of d is approximately double that of the correlation coefficient calculated from the same data. Another common method of reporting predictive accuracy is the area under the receiver operating characteristic curve (AUC) (Rice & Harris, 2005). This statistic is equivalent to the common language effect size, and is interpreted as the probability that a randomly selected recidivist will have a higher score on a test than a randomly selected nonrecidivist. In the case of Static-99, the AUCs are typically around .70. Similar levels of predictive accuracy have been found for diverse samples from Canada (16 studies), the United States (10 studies), and the United Kingdom (10 studies) (see Hanson & Morton-Bourgon, 2007), as well as in other European countries (Bengtson & Långström, 2007; Ducro & Pham, 2006; Endrass, Urbaniok, Held, Vetter, & Rossegger, in press; Rettenberger & Eher, 2006; Sjöstedt & Långström, 2001).

The variability of findings across studies, however, is more than would be expected by chance. In some studies, the predictive accuracy is substantially higher than in the developmental samples (e.g., Thornton, 2002), and in some cases, substantially lower (e.g., Långström, 2004; Tough, 2001). The reason for this variation is not fully understood, and it could plausibly be related to artificial differences in research methods and data sources, as well as to real differences in the relative predictive accuracy across populations.

Although Static-99 has demonstrated a robust ability to rank order offenders across samples, it is also important to consider the stability of the predicted recidivism rates. Even when evaluators can confidently identify offenders ranked in the top 10% of risk, decision-makers would be interested to know whether the absolute recidivism rates of the high risk group is 5%, 50%, or 95%. Determining the real recidivism rates is a difficult issue, because not all offenses are detected. Stability in the observed rates, however, would provide some basis upon which to extrapolate.

There has been relatively little research examining the stability of the absolute recidivism rates, and the results remain equivocal. In the developmental samples, Static-99 was able to differentiate a group of high-risk offenders, whose observed long-term recidivism rate was approximately 50%, from low-risk offenders, who observed recidivism rate was approximately 10%. Doren (2004a) concluded that the 5-year recidivism percentages, or probability estimates, were reasonably well replicated in seven studies for both the RRASOR and Static-99, with the exception of a total Static-99 score of four. Doren's (2004a) analysis, however, did not fully consider the impact of variations in base rates and variations in the relative risk associated with different risk categories (Mossman, 2006). In a new collection of eight studies, Helmus (2007) found that higher risk offenders (i.e., those with a score of 6 or over) reoffended sexually at a lower rate after 5 and 10 years than what would be expected given the probability estimates in the Static-99 manual. Similarly, Nicholaichuk's (2001) observed rates of recidivism at 10 years were generally lower in the high-risk samples than those observed in the Static-99 construction samples.

The reasons for these differences are not known, although the differences that have been observed are all in the same direction: down; i.e., the observed rates in the recent studies are lower than the rates observed in the development samples. The differences could be due to treatment effects, because most of the offenders in the construction samples either received no treatment or would have received treatment that would be regarded by today's standards as ineffective. Another factor is improved community supervision (residency restrictions, surveillance officers, and so forth). Officers might be "catching" the offenders for minor, nonsexual offenses, prior to sexual recidivism. Regardless of the reasons for the differences, evaluators should be cautious when interpreting and presenting the absolute recidivism rates, because the factors influencing these rates are not fully known.

Evaluators considering the absolute recidivism rates of a specific offender must assess the extent to which he is part of a sample similar to the development samples. If the offender is from a sample that differs substantially on a variable (or variables) known to influence recidivism rates (e.g., offender has completed a credible sex offender treatment program), then the evaluator must either find studies with more appropriate probability estimates or revise the explicit statement made regarding probability of recidivism accordingly (e.g., note the limitations to the current research, describe other factors relevant to recidivism risk, and indicate relevant research that suggests the variable(s) in the particular case that are associated with recidivism and not accounted for by the particular instrument).

Further complicating this issue is that the recidivism estimates are based on the reconviction data in the normative samples. The extent to which the observed recidivism rates approximate the real recidivism rates is a matter of debate in the scientific community. Victimization surveys find that a small proportion of sexual offenses are reported to police, and police records show that a minority of reported sexual offenses result in arrest or conviction. When given conditions of trust, convicted sexual offenders indicate that for each conviction there were 6 to 8 other victims for whom they were never the subject of criminal justice intervention (Groth, Longo, & McFadin, 1982; Weinrott & Saylor, 1991). For some offenses, the detection rate could be expected to be much lower (Russell, 1983a, 1983b).

Evaluators need to remember, however, the distinction between undetected offenses and undetected offenders. Even if the detection rate per offense is low, offenders who commit many of these offenses are likely to get caught (and appear in recidivism statistics). Consequently, the observed short-term recidivism rates are likely to seriously underestimate the true short-term rates, whereas the observed long-term recidivism rates will get closer to the true long-term rates as the length of the follow-up increases. Even with lifetime follow-up and complete records, the observed rates will always underestimate the true rates. The amount of this underestimation is unknown.

Appropriate and Inappropriate Populations

The Static-99 was developed for use with adult males who have been charged with or convicted of an offense that has a sexual motivation. The offense need not be named a sexual offense (i.e., the name of the charge could be one that does not necessarily indicate a sexual offense), but the motivation for the offense must be sexual (e.g., "theft" of women's used shoes). Some official record for an arrest, charge, conviction, breach, etc. must exist in order to use the Static-99. Obtaining "knowledge" of a sexual offense without official records (e.g., self-report, polygraph data) is not sufficient to use the Static-99.

The victim or victims of the sexual offense must be children, nonconsenting adults or adolescents, or others incapable of providing consent, including corpses and animals. The Static-99 is not for use with sex offenders whose only sexual offending involves no specific victim (e.g., prostitution, possession of pornography, illegal but consensual sexual behavior, such as consenting sex between similar-aged peers), nor for men whose only offenses are related to nondisclosure of HIV status. The Static-99 has not been validated for use with females or juveniles; however, some exceptions for juveniles are noted in the manual (e.g., older juveniles whose offending patterns are "adult" in nature) (see A. J. R. Harris et al., 2003, for further explanation).

The Static-99 was developed on adult male offenders released from custody who were followed for a period of time at risk in the community. These men had not been in the community for a lengthy time period following the detection and sanctioning of the sexual offending. Therefore, the Static-99 is not for use with offenders who have more than 10 years of offense-free behavior while "at risk" (i.e., in a community setting). The criteria for being offense-free include no sexual or violent offense charges, and no more than brief periods of incarceration for any reason (see Appendix 1 in A. J. R. Harris et al., 2003).

The Static-99 has shown acceptable levels of validity with Canadian Aboriginal offenders, clergy, offenders with developmental disabilities, offenders with psychiatric disorders, and offenders who have been charged but not convicted of a sexual offense. At the present time, there is no research to indicate that the factors included in the Static-99 are invalid with any of these populations. Static-99 has been demonstrated to work as intended in Westernized democracies (for example, Canada, United States, Europe, Australia, New Zealand); however, it has rarely been studied in different legal systems, and the preliminary studies have found lower-than-expected predictive accuracy for non-European sexual offenders in Sweden (Långström, 2004) and child rapists in Japan (Watanabe, Yokota, Yoshimoto, Ihara, & Fujita, 2007). It is always desirable to conduct a validity study before routinely using Static-99 in a specific jurisdiction; such validity studies are essential, however, when considering implementation in countries with legal traditions substantially different than those typically found in Westernized democracies.

Some evaluators are tempted to use Static-99 for offenders who do not fit into its sampling frame because no other measures have been validated for these special populations (for example, Internet-only offenders, johns). Avoid the temptation. Even if it was possible to score Static-99 (which is not always clear), the probability estimates and nominal risk categories derived from the construction samples would be meaningless when applied to groups outside the intended sampling frame.

Strengths and Limitations

Strengths

The Static-99 is the most widely used actuarial tool to gauge long-term risk potential of sexual and violent recidivism among sexual offenders. We believe there are several reasons for the widespread acceptance and use of this tool: Static-99 provides an explicit list of risk factors and an empirically supported method of combining the risk factors into an overall evaluation; it is as accurate as other measures for the prediction of sexual recidivism; it is relatively easily scored;

and it is freely available. Static-99 and the accompanying manual are available free of charge through the Web site of the office of Public Safety Canada (www.ps.gc.ca). Any individual or organizational user who wishes to obtain a copy may download all the necessary forms and directions. The disadvantage of such free distribution, however, is that access to the materials is no guarantee that they will be used correctly. Training and quality assurance programs are recommended for all users.

Another reason for the widespread implementation is that it can be scored by those without advanced professional degrees using readily available information. In contrast, several of the other tools required judgments and diagnoses that can only be completed by psychologists or psychiatrists (e.g., major mental disorder, psychopathy) and relatively costly evaluation procedures (e.g., phallometric data). The scoring of the Static-99 is straightforward, once the scoring rules have been learned and practiced. Unlike instruments that rely on subjective ratings or that use complicated algorithms, the Static-99 has the appeal of simplicity. The rules for scoring are clearly presented such that organizations have the ability to conduct their own quality assurance tests and to determine the reliability of their staff's ratings.

The widespread usage of the test has also resulted in far more research data being collected on Static-99 than on any other measure commonly used with sexual offenders. This plethora of research results in narrower confidence intervals for the group estimates, which translate into increasingly more precise estimates of a test's validity, again increasing the evaluators' confidence in the results of the tool.

Limitations

One major limitation of Static-99 is its lack of demonstrated construct validity. Evaluators who use more than one actuarial instrument frequently find that the results of different instruments conflict (Barbaree, Langton, & Peacock, 2006a). One approach to resolving such disagreements is to attempt to explain the divergent results by suggesting that the instruments are assessing different aspects of risk (e.g., Doren, 2004b). The problem with such interpretations is that Static-99 was not intended to have construct validity, so it is difficult to tell what it is measuring. Further work is required to establish conceptually meaningful, actuarial risk tools (see Hanson & Thornton, 2003, for an example).

Static-99 is as accurate as any of the other available measures for the prediction of sexual recidivism (Hanson & Morton-Bourgon, 2007). By conventional standards, the effect size is medium to large (Rice & Harris, 2005). Given the serious consequences of individual risk decisions, however, there would be value in improving the level of predictive accuracy. It is unlikely that the accuracy can be substantially improved by including alternate combinations of similar static, criminal history variables (Hanson & Thornton, 2003; Helmus, 2007). Improvements in predictive accuracy will likely come from assessing a broader range of variables.

Although Static-99 can be used on its own, we recommend that comprehensive assessments consider other factors that are consistently related to sexual or violent offense recidivism. Static-99 neglects a range of factors that have been demonstrated to be predictive of sexual recidivism. Most notably absent are direct measures of sexual deviance, such as phallometric testing results. Furthermore, Static-99 does not incorporate any dynamic risk factors (criminogenic needs), even though research indicates that such factors account for variance in recidivism beyond that which is accounted for by the static factors on the Static-99 (e.g., Beech, Friendship, Erikson, & Hanson, 2002; Hanson et al., 2007; Olver, Wong, Nicholaichuk, & Gordon, 2007; Thornton, 2002). When factors are included using a structured, empirically based procedure, they can increase the accuracy of the overall risk evaluation. One empirically based method of including dynamic factors with Static-99 scores is described in the next section.

Another external factor not fully considered in Static-99 is advanced age (Hanson, 2006). Static-99 has one item distinguishing between offenders 18 to 24 (higher risk) and those 25 or older (lower risk). There is accumulating evidence, however, that further age differentiations may increase the accuracy of risk prediction (e.g., Barbaree & Blanchard, 2008; Hanson & Thornton, 2003). In particular, offenders over the age of 60 are consistently at lower risk to reoffend than would be suggested by their Static-99 scores (Hanson, 2006). Research has yet to establish the extent to which the observed reduction in recidivism rates is due to offenders improving as they get older (Barbaree & Blanchard, 2008) or the observation that older sex offenders are generally lower risk than younger sex offenders (Harris & Rice, 2007).

The Sex Offender Needs Assessment Rating (SONAR)

In order to identify dynamic risk factors contributing to recidivism among sexual offenders, Hanson and Harris (1998, 2000) conducted a retrospective study of offenders released on community supervision. They investigated the files of 201 recidivists and 209 nonrecidivists (all sex offenders). They coded the notes taken by the supervising officers prior to the reoffense for the recidivists, and for the same time period for the nonrecidivists, and they interviewed the supervising officers. Following the data collection, Hanson and Harris generated a list of domains that separated the recidivists from the nonrecidivists. Some of these domains were labeled "stable" dynamic risk factors, because these were areas of the offenders' lives that appeared to be chronically troubled (i.e., for a period of 6 months to a year) prior to the reoffense. These domains are best described as persistent patterns of thinking, feeling, and behaving, or the typical processes in an offender that are capable of change, albeit with concerted effort (such as with treatment or through interventions during supervision), and once truly changed, the changes would be expected to endure. Other domains were labeled "acute" dynamic risk factors, because these areas appeared to create problems immediately prior to the commission of the reoffense (that is, within one month of the reoffending). These areas would be considered rapidly changing environmental circumstances or, alternatively, sudden glitches in the otherwise persistent patterns of functioning. This list of stable and acute dynamic risk factors became the Sex Offender Needs Assessment Rating (SONAR) (Hanson & Harris, 2001).

The Stable-2000 and Acute-2000

Combining the SONAR research with the work of other projects in dynamic risk assessment (Sex Offender Treatment Evaluation Project, Beech et al., 2002; Structured Risk Assessment, Thornton, 2002), Hanson and Harris created a more comprehensive list of stable and dynamic risk factors and embarked on a prospective validation study of their new instruments: the Stable-2000 and the Acute-2000.

The prospective validation study involved supervision offices of numerous jurisdictions in Canada and in the United States. Data on the Static-99, Stable-2000, Acute-2000, and other demographic and offense characteristics were collected for approximately 1,000 sex offenders who were new offenders to the caseloads of the participating officers. The Static-99 was scored once, the Stable-2000 was scored every 6 months until the offender had reached the end of his or her supervision period, and the Acute-2000 was scored at every supervision meeting until the offender had reached the end of his or her supervision period.* The study began in 2001, and the results of the study were published in 2007 (see Hanson et al., 2007).

* The project included six female sexual offenders, one of whom reoffended with a nonsexual violent offense.

The Stable-2007 and Acute-2007

The data from the Dynamic Supervision Project (Hanson et al., 2007) found support for Stable-2000 and Acute-2000, but also suggested changes to the Stable and Acute measures were warranted. First, the attitudes section did not add predictive validity and hence was removed. The reason for the lack of incremental validity is unknown. In the Hanson and Morton-Bourgon (2004) meta-analysis, offense-supportive attitudes showed an overall positive association with recidivism, but the effect was only significant for offenders who had entered treatment. Attitudes showed no relationship to recidivism for offenders assessed in adversarial contexts. Consequently, it cannot be immediately assumed that attitudes are inconsequential until issues in the applied measurement of attitudes are resolved. Irrespective of the cause, the attitude section as it was presented in the Stable-2000 did not add value to the measure and was therefore removed.

Aside from the above major change, other minor changes were made to the scoring of some sections. One item (Emotional Identification with Children) proved valid for child molesters but not rapists, and hence the instructions for the Stable-2007 lead the assessor to score the item only for offenders with a child victim. Finally, the method for scoring the measure was simplified, and the interpretive needs categories were changed to empirically derived groupings.

The items on the Acute-2000 remained as they were; however, it appears that these acute factors were not as acute as initially assumed. The ratings on the acute variables did predict recidivism, but the most recent acute measures did not predict recidivism as well as the aggregated ratings on the acute over a longer period prior to the reoffending. This result suggests that ongoing expressions of problematic behavior, as opposed to sudden glitches, better distinguished recidivists from nonrecidivists. These findings led to further suggestions for the administration of the Acute-2007 during the course of supervision.

The study found that stable and acute evaluations incrementally improved the risk predictions above that provided by Static-99. For the prediction of sexual recidivism, the AUC was .77 for the Static-99 alone and .81 for the Static-99 and the Stable-2007 combined. For offenders in the highest and lowest risk categories of the Static-99, the Stable-2007 scores did not suggest major modifications to the risk level For example, offenders who scored low on the Static-99 but high on Stable-2007 would be considered in the overall moderate-low category. However, those who scored in the low-moderate and moderate-high risk categories of the Static-99 more closely approximated the recidivism rates in the adjacent nominal risk category when Stable-2007 scores were in the extreme end. That is, offenders who scored in the low-moderate range of the Static-99 (i.e., a 2 or 3) but scored in the high range of the Stable-2007 recidivated at rates expected of offenders in the moderate-high category, whereas offenders scoring in the same Static-99 risk category but scoring in the low range on the Stable-2007 reoffended at rates expected of those in the low risk categories. Similarly, those scoring in the moderate-high range on the Static-99 (i.e., a score of 4 or 5) with a score in the low range on Stable-2007 reoffended at rates expected of those in a moderate-low risk category. However, those scoring in the same nominal risk category of the Static-99 but scoring in the high range on the Stable-2007 reoffended at rates comparable to those in the high risk categories. Specific recidivism rates for these overall risk/priority categories are offered in the appendices of Hanson et al. (2007). The reader is cautioned, however, on the small numbers within some of these combinations of nominal risk category bins. For example, there were only three individuals who had a Static-99 score of 6 or over (i.e., high-risk category) along with a low score on the Stable-2007, and only one of these individuals recidivated with a new sexual crime (an exhibitionist). This study is important because it demonstrated that the Static-99 could be implemented in a reliable and valid manner by community supervision officers (Helmus & Hanson, 2007). Furthermore, the supervision officers could

be taught a structured approach to evaluating risk factors that improves risk prediction while providing meaningful targets for intervention (see Hanson et al., 2007).

Questions and Future Directions

Selecting an Actuarial Measure

Determining which actuarial measure to use depends on the similarity of the individual offender being assessed to the groups of offenders from which the normative data were derived, the referral question, and the resources available.

If the primary interest is the prediction of general violent recidivism (sexual or nonsexual), the Violence Risk Appraisal Guide (VRAG) and the Sex Offender Risk Appraisal Guide (SORAG) (Quinsey, Harris, Rice, & Cormier, 2005) appear more accurate than Static-99 (Hanson & Morton-Bourgon, 2007). Both the VRAG and SORAG, however, require the administration of the Psychopathy Checklist–Revised, which involves higher user qualifications and greater expenditure of resources (cost, time, etc.).

For the prediction of sexual recidivism, Static-99 is as accurate as any of the available measures, and more accurate than its predecessor, the RRASOR (Hanson & Morton-Bourgon, 2007). Some evaluators continue to use the RRASOR in the belief that it provides a purer measure of sexual deviance, whereas Static-99 taps both sexual deviance and general criminality. Given that there are better methods of assessing the constructs associated with risk (e.g., Stable-2007, Hanson et al., 2007; Thornton's Structured Assessment of Risk and Need, HM Prison Service, 2005), we believe that the use of the RRASOR should be completely discontinued. In every case where the RRASOR would be appropriate, Static-99 provides a better assessment of risk. To the extent that it makes sense to interpret the underlying constructs, Static-99 addresses a wider range of risk factors (both antisociality and sexual deviance), and it has been replicated much more extensively than the RRASOR (that is, in more countries, with more diverse samples, with a greater number of cases overall). Using both Static-99 and RRASOR has the added risk of inflating the perception of consistency, given that the RRASOR results will appear to be an independent confirmation of the Static-99 findings even though they add no new information.

Do We Need to Know Absolute Recidivism Rates?

Projecting absolute recidivism rates is difficult given the observed fluctuations across studies, and the expectation that interventions and global policy initiatives should systematically reduce recidivism. In many contexts, highly accurate estimates of absolute recidivism rates are not needed. For the purpose of supervising a given caseload of sexual offenders released to the community, the prominent concern is rank ordering the offenders in terms of risk in order to determine how the resources might be allocated most efficiently.

In other contexts, however, providing accurate probability estimates is of paramount importance. For example, in cases where legal decisions are made based on some statutorily defined probability of reoffending (for example, dangerous offender sentences in Canada and civil commitment hearings in many jurisdictions of the United States), evaluators need to carefully consider the available evidence and determine which findings are most appropriate for estimating the risk for the individual at hand (although it is recognized that many statutes do not specify a numeric "probability of reoffending," and case law is not unanimous or consistent in this requirement).

This Offender Seems Like an Exception—Can I "Override" the Static-99?

It is not unusual for evaluators to encounter offenders whose scores on the Static-99 does not seem well matched to their initial perception of the risk. For example, an offender's score may

fall into the low-risk range, but his offenses were particularly violent, or there are many victims who do not figure on the official record. We recommend against adjusting the Static-99 findings based on such ad hoc observations.

When Hanson et al. (2007) conducted their prospective validation study of the static and dynamic risk assessment instruments (Static-99, Stable-2000, and Acute-2000), they collected information on "overrides" to the Static-99. That is, they allowed supervising officers who scored the measures to indicate if they felt there were circumstances that justified an adjustment of the overall risk assessment category, and they were asked to record what circumstance(s) they would use to make such an adjustment. There was no agreement about when the overrides should be used, and when they were, the adjusted risk assessment resulted in a slight decrease in predictive accuracy over that provided by Static-99 alone. When, however, the Stable-2007 factors were used to adjust the nominal risk category provided by the Static-99 score, the predictive accuracy was increased. Therefore, the Static-99, as noted throughout this chapter, does not account for all important factors in estimating risk of recidivism. Nonetheless, adjustments to the nominal risk category can be made by incorporating empirically based risk factors, and the current research by Hanson et al. (2007) provides a structured and validated tool for incorporating dynamic risk factors into the overall assessment.

Case Example

The following example suggests a way in which information obtained from the Static-99 and Stable-2007 might be utilized within the context of a risk assessment for an incarcerated individual.

Mr. Dodgy is a 38-year-old male serving a sentence of 5 years for sexual assault with a weapon. The current offense was committed against a previous girlfriend whom he dated for 6 months. His criminal history dates back to age 18 years when he was convicted for break, enter, and theft. Mr. Dodgy has one previous charge for sexual assault, but he pled to assault, and the sexual assault charge was withdrawn. This offense occurred 3 years ago and was perpetrated against the same victim of the current offense. It appears that the previous offense led to the breakup of the relationship. Mr. Dodgy has also been convicted of criminal harassment against the same victim 1 year before the current offense. Mr. Dodgy's previous offenses include assault (mainly against males, with the exception of the one noted above), other property offenses, and one driving while impaired. Mr. Dodgy was married 10 years ago to another woman, but this marriage lasted for approximately 1 year. He states they separated and divorced due to his infidelity and drinking. He has no children, and he has never lived with any woman other than his previous wife. Mr. Dodgy is not currently involved in a relationship, but says he dates "a lot" and enjoys "playing the field." He stated he does not have much luck with longer-term relationships, as women often become "naggy, bitchy, and boring" when there are increases in the commitment. Mr. Dodgy admitted he did not feel comfortable discussing his thoughts on relationships with women because there are too many women in positions of power over him (e.g., his parole officer, most of the facilitators of programs, and the unit psychologist), and he did not believe this was fair in a male correctional facility. Mr. Dodgy reported no difficulties making friends, but stated he prefers to be alone most of the time. He said his relationship with family members is "satisfactory" and said he would spend time with two of the men from work, but notes that one has a significant problem with drugs and alcohol, is frequently out of money, and often approaches Mr. Dodgy for small loans to manage until he receives his next paycheck. Mr. Dodgy characterizes his relationship with the current victim as "poison" and says he realizes he cannot ever see or contact her again. He admits to the current offense, but states that "it was blown out of proportion in court." He pled guilty after a preliminary hearing. Mr. Dodgy has been steadily employed with a roofing company for the past 15 years, and he believes he will be able to return to work for the same company. The most recent community assessment report indicates that his employer is likely to give him a job upon release. He admits to having a drinking problem and states that "most roofers drink on the job and smoke a lot of pot." He denied the use of drugs, including prescription medication; however, file information revealed an institutional charge one month ago resulting from a urinalysis testing positive for THC.

The following paragraphs summarize the findings on the Static-99 and Stable-2007 for Mr. Dodgy within the context of the risk assessment report. The risk assessment was being completed during Mr. Dodgy's period of incarceration to assist with an upcoming hearing with the Parole Board.

. . . The STATIC-99 is an instrument designed to assist in the prediction of sexual and violent recidivism for sexual offenders. This risk assessment instrument was developed by Hanson and Thornton (1999) based on follow-up studies from Canada and the United Kingdom with a total sample size of 1,301 sexual offenders. The STATIC-99 consists of 10 items and produces estimates of future reoffense risk based upon the number of risk factors present in any one individual. The risk factors included are the presence of prior sexual offenses, having committed a current nonsexual violent offense, having a history of nonsexual violence, the number of previous sentencing dates, age less than 25 years old, having male victims, having never lived with a lover for two continuous years, having a history of noncontact sex offenses, having unrelated victims, and having stranger victims. Mr. Dodgy scored a 5 on the STATIC-99. Using the data provided by the normative samples, individuals with these characteristics, on average, sexually reoffended at a rate 33% over 5 years and at 38% over 10 years. The rate for any violent recidivism (including sexual) for individuals with these characteristics was, on average, 42% over 5 years and 48% over 10 years. Based upon the STATIC-99 score, this places Mr. Dodgy in the moderate-high (between the 62nd and the 88th percentile) risk category relative to other adult male sex offenders. This nominal risk level will be combined with the STABLE-2007 Needs score to form a combined risk assessment.

The STABLE-2007 was developed to assess change in intermediate-term risk status, assess treatment needs, and help predict recidivism in sexual offenders. Hanson and Harris (2000; Hanson, Harris, Scott, & Helmus, 2007) developed this risk assessment instrument based on a large prospective study from Canada and the states of Alaska and Iowa with a total sample size of 997 sexual offenders. The STABLE-2007 consists of 13 items and produces estimates of stable dynamic risk based upon the number of stable dynamic risk factors present in any one individual. The risk factors included are the nature of significant social influences, capacity for relationship stability, emotional identification with children, hostility toward women, general social rejection, lack of concern for others, impulsivity, poor problem solving skills, negative emotionality, sex drive and preoccupation, sex as coping, deviant sexual preference, and cooperation with supervision.

Mr. Dodgy scored a 14 out of a possible 24 points on the STABLE-2007, and this score places him as high needs on this assessment instrument. Men without a child sexual victim are scored out of 24 possible points. The STATIC-99 and the STABLE-2007 are then combined into a composite score. Mr. Dodgy scored as a moderate-high risk on the STATIC-99, and this score is combined with Mr. Dodgy's STABLE-2007 score of high needs to produce estimates of sexual, violent, or any criminal recidivism between 1 and 4 years in the community. The complete risk tables are shown in the paper referenced below (Appendices 8 & 9). Compared to other sex offenders on community supervision, men with similar characteristics to Mr. Dodgy on both the static and dynamic risk factors assessed by these instruments fell into the high nominal risk category for sexual recidivism, and men with the same risk profile as Mr. Dodgy were seen to recidivate sexually at a rate of 17.1% over 2 years and 22% over 4 years (Appendix 8). Men with the same risk profile as Mr. Dodgy reoffended violently at rates that indicate Mr. Dodgy falls into the high nominal risk category for violent recidivism, and men with the same risk profile as Mr. Dodgy were seen to recidivate violently at a rate of 25.6% over 2 years and 31.9% over 4 years (Appendix 9). It is noted for the reader that these rates of recidivism are based on the only sample to date examining the combination of the STATIC-99 and STABLE-2007 scores, and hence caution must be exercised when interpreting these rates as estimates of the probability of recidivism.

As well, the nominal risk categories noted above are predicated on recidivism estimates provided by the STATIC-99 and STABLE-2007, which are group estimates based upon charges, reconvictions, and breaches of conditional release derived from groups of individuals with these risk characteristics. As such, these estimates are based upon a group of offenders with the same risk profile as Mr. Dodgy. Mr. Dodgy's risk may be higher or lower than the estimated probabilities, depending on other risk factors not measured by these instruments. Given the documentation provided and the information provided by Mr. Dodgy in our current interview, there are no apparent factors known or suspected to be associated with recidivism in Mr. Dodgy's case that would warrant an adjustment to the risk levels indicated in this report. The specific factors presenting as a focus for intervention (i.e., those that are highlighted as problematic according to the STABLE-2007 evaluation)

are negative social influences, difficulties maintaining intimate relationships, hostility toward women, sexual preoccupations, lifestyle impulsivity and lack of cooperation with supervision. . . .

Conclusion

The Static-99 has proven to be a useful means of categorizing offenders according to their risk of recidivism. Although it can be used on its own, the most useful assessments are those that use a structured, empirically based method for combining static risk factors with measures of criminogenic needs (dynamic risk factors). The Static-99 can help gauge the long-term potential for sexual and violent recidivism among men who have already been charged with or convicted for a sexually motivated offense that involves a targeted victim or group of victims. The Static-99 is not valid with every sex offender. Also, the Static-99 neglects factors that have been repeatedly associated with risk for sexual and violent recidivism, and in the context of providing a comprehensive risk assessment, such factors should be considered. Promising measures of dynamic risk factors are the Stable-2007 and Acute-2007, which accounted for variance in recidivism beyond that accounted for by the Static-99 in the Dynamic Supervision Project (Hanson et al., 2007). The assessor must be cognizant of the limitations to any tool being used to assist in the estimate of potential for recidivism and therefore must be aware of the extent to which the individual offender being assessed is similar to the groups of offenders for whom the test has been validated, and the assessor must evaluate whether the tool assists in answering the referral question (or the extent to which the tool contributes meaningful information to address the reason for the risk assessment being conducted).

> There are very few human beings who receive the truth, complete and staggering, by instant illumination. Most of them acquire it fragment by fragment, on a small scale, by successive developments, cellularly, like a laborious mosaic.
>
> **Anaïs Nin**

References

Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment, 87,* 84–94.

Barbaree, H. E., & Blanchard, R. (2008). Sexual deviance over the lifespan: Reductions in deviant sexual behavior in the aging sex offender. In D. R. Laws & W. T. O'Donohue, (Eds.). *Sexual deviance: Theory, assessment and treatment* (pp. 37–60). New York: Guilford Press.

Barbaree, H. E., Langton, C. M., & Peacock, E. J. (2006a). Different actuarial risk measures produce different risk rankings for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 18,* 423–440.

Barbaree, H. E., Langton, C. M., & Peacock, E. J. (2006b). The factor structure of static actuarial items: Its relation to prediction. *Sexual Abuse: A Journal of Research and Treatment, 18,* 207–226.

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.

Beech, A., Friendship, C., Erikson, M., & Hanson, R. K. (2002). The relationship between static and dynamic risk factors and reconviction in a sample of U.K. child abusers. *Sexual Abuse: A Journal of Research and Treatment, 14,* 155–167.

Bengtson, S., & Långström, N. (2007). Unguided clinical and actuarial assessment of re-offending risk: A direct comparison with sex offenders in Denmark. *Sexual Abuse: A Journal of Research and Treatment, 19,* 135–153.

Doren, D. M. (2004a). Stability of the interpretive risk percentages for the RRASOR and Static-99, *Sexual Abuse: A Journal of Research and Treatment, 16,* 25–36.

Doren, D. M. (2004b). Toward a multidimensional model for sexual recidivism risk. *Journal of Interpersonal Violence, 19,* 835–856.

Ducro, C., & Pham, T. (2006). Evaluation of the SORAG and the Static-99 on Belgian sex offenders committed to a forensic facility. *Sexual Abuse: A Journal of Research and Treatment, 18,* 15–26.

Endrass, J., Urbaniok, F., Held, L., Vetter, S., & Rossegger, A. (in press). The accuracy of the Static-99 in predicting recidivism in Switzerland. *International Journal of Offender Therapy and Comparative Criminology.*

Groth, N., Longo, R., & McFadin, J. B. (1982). Undetected recidivism among rapists and child molesters. *Crime & Delinquency, 28,* 450–58.

Grubin, D. (1998). *Sex offending against children: Understanding the risk.* Police Research Series Paper 99. London Home Office.

Hanson, R. K. (1997). The development of a brief actuarial risk scale for sexual offense recidivism. (User Report 97-04). Ottawa, ON, Canada: Department of the Solicitor General of Canada.

Hanson, R. K. (2001). *Note on the reliability of Static-99 as used by the California Department of Mental Health evaluators.* Unpublished report. Sacramento, CA: California Department of Mental Health.

Hanson, R. K. (2006). Does Static-99 predict recidivism among older sexual offenders? *Sexual Abuse: A Journal of Research and Treatment, 18,* 343–355.

Hanson, R. K., & Bussière, M. T. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology, 66*(2), 348–362.

Hanson, R. K., & Harris, A. J. R. (1998). *The Sex Offender Need Assessment Rating (SONAR): A method for measuring change in risk levels.* (User Report 2000-01). Ottawa, ON, Canada: Department of the Solicitor General of Canada.

Hanson, R. K., & Harris, A. J. R. (2000). Where should we intervene? Dynamic predictors of sexual offense recidivism. *Criminal Justice and Behavior, 27,* 6–35.

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*(2), 105–122.

Hanson, R. K., Harris, A. J. R., Scott, T., & Helmus, L. (2007). Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project. User Report, Corrections Research. Ottawa, ON, Canada: Public Safety Canada.

Hanson, R. K., & Morton-Bourgon, K. (2004). Predictors of sexual recidivism: An updated meta-analysis. Corrections User Report No. 2004-02. Ottawa, ON, Canada: Public Safety and Emergency Preparedness Canada.

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.* Corrections User Report No 2007-01. Ottawa, ON, Canada: Public Safety and Emergency Preparedness Canada.

Hanson, R. K., & Morton-Bourgon, K. E. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment, 21*(1), 1–21.

Hanson, R. K., Steffy, R. A., & Gauthier, R. (1993). Long-term recidivism of child molesters. *Journal of Consulting and Clinical Psychology, 61,* 646–652.

Hanson, R. K., & Thornton, D. (1999). *Static-99: Improving actuarial risk assessments for sex offenders.* User Report 99-02. Ottawa, ON, Canada: Department of the Solicitor General of Canada.

Hanson, R. K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior, 24*(1), 119–136.

Hanson, R. K., & Thornton, D. (2003). Notes on the development of Static-2002. User Report 2003-01. Ottawa, ON, Canada: Department of the Solicitor General of Canada.

Harris, A. J. R., & Hanson, R. K. (2004). Sex offender recidivism: A simple question. Corrections Users Report No. 2004-03. Ottawa, ON, Canada: Public Safety and Emergency Preparedness Canada.

Harris, A. J. R., Phenix, A., Hanson, R. K., & Thornton, D. (2003). Static-99 coding rules: Revised 2003. Ottawa, ON, Canada: Department of the Solicitor General of Canada.

Harris, G. T., & Rice, M. E. (2007). Adjusting actuarial violence risk assessments based on aging or the passage of time. *Criminal Justice and Behavior, 34,* 297–313.

Harris, G. T., Rice, M. E., Quinsey, V. L., Lalumiere, M. L., Boer, D., & Lang, C. (2003). A multi-site comparison of actuarial risk instruments for sex offenders. *Psychological Assessment, 15(3),* 413–425.

Helmus, L. (2007). A multi-site comparison of the validity and utility of the Static-99 and Static-2002 for risk assessment with sexual offenders. Unpublished B.A. thesis, Carleton University, Ottawa, ON, Canada.

Helmus, L. M. D., & Hanson, R. K. (2007). Predictive validity of the Static-99 and Static-2002 for sex offenders on community supervision. *Sexual Offender Treatment, 2,* 1–14.

HM Prison Service. (2005). *Structured Assessment of Risk and Need (Sexual Offenders): Manual, v. 2.* London: Offending Behaviour Programmes Unit, HM Prison Service.

Interstate Commission for Adult Offender Supervision. (2007). Sex offender assessment information survey. ICAOS Documents 4-2007. Lexington, KY.

Jackson, R. L., & Hess, D. T. (2007). Evaluation for civil commitment of sex offenders: A survey of experts. *Sexual Abuse: A Journal of Research and Treatment, 19,* 409–448.

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.

Looman, J. A. (2006). Comparison of two risk assessment instruments for sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 18,* 193–206.

McGrath, R. J., Cumming, G. F., & Burchard, B. L. (2003). *Current practices and trends in sexual abuser management: The Safer Society 2002 Nationwide Survey.* Brandon, VT: Safer Society Foundation, Inc.

Mossman, D. (2006). Another look at interpreting risk categories. *Abuse: A Journal of Research and Treatment, 18,* 41–63.

Nicholaichuk, T. (2001, November). *The comparison of two standardized risk assessment instruments in a sample of Canadian Aboriginal sexual offenders.* Paper presented at the annual conference of the Association for the Treatment of Sexual Abusers, San Antonio, TX.

Olver, M., Wong, S. C. P., & Nicholaichuk, T. P., & Gordon, A. (2007). The validity and reliability of the Violence Risk Scale–Sexual Offender Version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment, 19,* 318–329.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. (2005). *Violent offenders: Appraising and managing risk,* (2nd ed.). Washington, DC: American Psychological Association.

Rettenberger, M., & Eher, R. (2006). Actuarial assessment of sex offender recidivism risk: A validation of the German version of the Static-99. *Sexual Offender Treatment, 1*(3), 1–11.

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC area, Cohen's *d,* and *r. Law and Human Behavior, 29,* 615–620.

Roberts, C. F., Doren, D. M., & Thornton, D. (2002). Dimensions associated with assessments of sex offender recidivism risk. *Criminal Justice and Behavior, 29,* 569–589.

Russell, D. E. H. (1983a). The incidence and prevalence of intrafamilial and extrafamilial sexual abuse of female children. *Child Abuse & Neglect, 7,* 133–146.

Russell, D. E. H. (1983b). The prevalence and incidence of forcible rape and attempted rape of females. *Victimology: An International Journal, 7,* 81–93.

Sjöstedt, G., & Långström, N. (2001). Actuarial assessment of sex offender recidivism risk: A cross-validation of the RRASOR and the Static-99 in Sweden. *Law and Human Behavior, 25,* 629–645.

Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment, 14,* 139–153.

Tough, S. E. (2001). *Validation of two standardized risk assessments (RRASOR, 1997; Static-99, 1999) on a sample of adult males who are developmentally disabled with significant cognitive deficits.* Unpublished master's thesis, University of Toronto, Toronto, ON, Canada.

Watanabe, K., Yokota, K., Yoshimoto, K., Ihara, N., & Fujita, G. (2007). *Recidivism in child rapists: Identifying high risk factors.* Unpublished manuscript, obtained from National Research Institute of Police Science, Tokyo, Japan.

Weinrott, M. R., & Saylor, M. (1991). Self-report of crimes committed by sex offenders. *Journal of Interpersonal Violence, 6,* 286–300.