

Avoiding Garbage in Sex Offender Re-Offense Risk Assessment: A Case Study

Gregory DeClue

ABSTRACT. This case study presents an extreme example of a problem that appears to occur frequently: misuse of risk assessment instruments in sex offender evaluations. This study does not critique the instruments themselves, but highlights problems in their application to specific cases. A particular challenge to evaluators is how to assess risk when some data about the subject are unknown or unreliable. For some currently used risk-assessment instruments, there are no established guidelines for how to accommodate missing data. Evaluators must not use actuarial instruments when there is insufficient reliable data to score the instruments. Instead, evaluators should list relevant risk factors and describe which factors are present, absent, or unknown. *[Article copies available for a fee from The Haworth Document Delivery Service: 1-800-HAWORTH. E-mail address: <getinfo@haworthpress.com> Website: <http://www.HaworthPress.com> © 2002 by The Haworth Press, Inc. All rights reserved.]*

Gregory DeClue, PhD, ABPP, is a Fellow of the American Academy of Forensic Psychology. He received his PhD in counseling psychology from the University of Missouri in 1983. He practices counseling and forensic psychology in Sarasota, Florida.

Address correspondence to: Gregory DeClue, 16443 Winburn Place, Sarasota, FL 34240-9228 (E-mail: gregdeclue@mailmt.com).

The author wishes to thank Randy Otto, Joel Dvoskin, Peter Bursten, and Donna DeClue, who commented on previous versions of this article.

A version of this article was presented at the 34th Annual Conference of the American Association of Sex Educators, Counselors, and Therapists, Miami Beach, Florida, May 3, 2002.

Journal of Threat Assessment, Vol. 2(2) 2002
<http://www.haworthpress.com/store/product.asp?sku=J177>
© 2002 by The Haworth Press, Inc. All rights reserved.

10.1300/J177v02n02_05

KEYWORDS. Sex offender, risk assessment, recidivism, actuarial, guided clinical, prediction, civil commitment

One approach to reducing sexual violence in society is to identify who is most likely to commit acts of sexual violence, and then take steps to reduce those people's opportunities and/or inclinations to commit sexually violent acts. Psychologists and psychiatrists are frequently asked to conduct risk assessments to assist in identifying which people are most likely to commit acts of sexual violence. Clinicians are often recommended, and in some states (e.g., Florida) required, to use forensic assessment instruments as they conduct these evaluations. This case study illustrates misuse of these assessment instruments. Examination of what went wrong in this case may help clinicians avoid similar errors in other cases.

LEGAL AND ETHICAL ISSUES

In *Kansas v. Hendricks* (1997), the United States Supreme Court ruled that it can be constitutional for states to civilly commit some sex offenders after they serve their criminal sentences. Several states have enacted laws and begun the civil commitment process, including a role for psychologists and psychiatrists to evaluate the respondent prior to a court hearing.

There is considerable scholarly debate about whether this type of civil commitment is fair or wise, and about whether and how scientist-practitioners can make accurate predictions (cf., Winick & LaFond, 1998). Careful consideration of issues in this debate is important, but the debate does not preclude practitioners from conducting risk assessments now. After considering how courts have ruled on the admissibility of predictions of dangerousness, Janus and Meehl (1997) write, ". . . it seems well established that there is no constitutional impediment to using predictions of dangerousness in legal proceedings, up to and including those that may result in loss of liberty or death. As a legal matter, prediction is not, in all of its forms and for all purposes, so inaccurate as to violate the due process clause" (p. 36).

The case study discussed in this article arose in Florida. The following summary of the commitment statute is provided for context. Chapter 394 of Florida Statutes (1999) includes a section relating to civil commitment of sexually violent predators. The law is concerned with "a

small but extremely dangerous number of sexually violent predators” who are “likely to engage in criminal, sexually violent behavior” (Florida Statutes §394.910).

“Sexually violent predator” is defined as “any person who (a) has been convicted of a sexually violent offense; and (b) suffers from a mental abnormality or personality disorder that makes the person likely to engage in acts of sexual violence if not confined in a secure facility for long-term control, care, and treatment” (Florida Statutes §394.912(10)). The population affected by the law includes people who would otherwise be released from a prison or jail, a psychiatric hospital, or a juvenile commitment. “Likely to engage in acts of sexual violence” is defined as “the person’s propensity to commit acts of sexual violence is of such a degree as to pose a menace to the health and safety of others” (Florida Statutes §394.912(4)).

Evaluators are given the task of gathering and analyzing data—including a direct examination of the person—to assist the trier of fact in determining whether the person meets criteria for civil commitment as a sexually violent predator. Either the petitioner (the State) or the respondent can elect to have the case decided by a jury rather than the trial judge. The burden of proof is on the State, which must show by clear and convincing evidence that the person meets these criteria for being a sexually violent predator. If the person meets criteria, he or she is confined indefinitely, with yearly reviews.

RATIONALE FOR USING ACTUARIAL INSTRUMENTS

A recent meta-analysis of sexual re-offense studies identifies some of the variables most predictive of re-offense, including sexual deviance and antisocial personality (Hanson & Bussiere, 1996, 1998). “Sexual offense recidivism was best predicted by measures of sexual deviancy (e.g., deviant sexual preferences, prior sexual offenses) and, to a lesser extent, by general criminological factors (e.g., age, total prior offenses)” (Hanson, 1998, p. 348).

As relevant risk factors are identified, evaluators no longer use the traditional clinical approach, which relies on an unstructured clinical interview and a battery of psychological tests. Indeed, it might be considered unethical for a practitioner to rely on the traditional clinical approach, ignoring relevant risk factors and basing the recommendation on factors that have no empirical relationship to sexual re-offense (Grisso, 2000).

Actuarial instruments can be seen as attempts to identify the base rate of re-offending for a specifiable subgroup of sex offenders. For example, we might imagine that the rate of re-offending by convicted rapists might be different than the rate of re-offending by convicted child molesters. A well-designed study suggests that this is true. Using survival analysis, Prentky et al. (1997) found a 25-year sexual recidivism rate of 39% for rapists and a 52% sexual recidivism rate for child molesters. A hypothetical actuarial instrument—call it *SEXTEST*—could be developed from this data alone. The *SEXTEST* would have one item, with two possibilities: child molester or rapist. We could say that a particular sex offender's "score" (child molester or rapist) on the *SEXTEST* is similar to the scores of offenders whose likelihood of being charged with a new sexual offense if at risk for 25 years is 52% (for a child molester) or 39% (for a rapist).

Actuarial instruments currently in use operate similarly to the hypothetical *SEXTEST*.¹ An evaluator should first determine that a target sex offender is comparable to those in the research samples for a particular actuarial instrument, then determine whether the data regarding the target offender are comparable to the data used to categorize the offenders in the development sample. Because our *SEXTEST* was developed from a study of males who had been committed to a special hospital after conviction for repetitive and/or aggressive sexual offenses, it would be inappropriate to use the instrument for predicting recidivism of a female sex offender, or for a male who had not been convicted of repetitive and/or aggressive sexual offenses.

Because the relevant study excluded offenders who had some victims who were both over and under 16 years of age with no primary target age, the *SEXTEST* would not be appropriate for use in such cases. It would be inappropriate to use the instrument when it was unknown whether the offender's victims were over or under age 16. An actuarial instrument is only useful when the target offender is similar to the offenders in research sample(s), and when the quantity and quality of the data regarding the target offender is comparable to that available for the offenders in research sample(s).

Consider two more points before we turn to commonly used actuarial instruments. First, the base rate for sexual recidivism depends on a number of factors other than the items that show up in the instruments. The outcome measure for recidivism in different studies might be re-offense, re-conviction, or re-incarceration. These different measures lead to different estimates of the rate of recidivism, and there is no outcome measure that tells exactly how many offenders in a research sample

have committed new sexual offenses (cf., Doren, 1998). Other factors that are likely to affect base rates include type of offense (e.g., extra-familial child molestation, intrafamilial child molestation, rape) and case disposition (e.g., probation versus incarceration). For a case involving an incest offender on probation, one should not use an actuarial tool normed on incarcerated rapists.

The second point is that actuarial instruments can be seen as offering a convenient—not a unique—way of organizing data. Doren (1998) [citing Hanson et al. (1993) and Hanson (1996)] reports a 77% likelihood of sexual recidivism (over a lengthy follow-up period) for extrafamilial child molesters who offended against one or more male victims, had been convicted of at least one prior sexual offense (in addition to the offense leading to current incarceration), and had never married. That provides a little more detailed information than the hypothetical SEXTEST, and that type of information could be used to design an actuarial test with a few more items. Indeed, that study was incorporated into the development of the first actuarial instrument we will consider, the RRASOR.

USING ACTUARIAL INSTRUMENTS

The following actuarial instruments were used by evaluators in the case study presented below: the Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR; Hanson, 1997), the Minnesota Sex Offender Screening Tool-Revised (MnSOST-R; Epperson et al., 1998), and the Sexual Offender Risk Appraisal Guide (SORAG; Quinsey et al., 1998). In this section, I address issues regarding the use of these instruments in conducting risk assessments, focusing on the research samples and, in particular, the type of data necessary for scoring the instruments. Some comments about the use of a specific instrument will apply to other actuarial instruments as well.

RRASOR

The Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR; Hanson, 1997) consists of four items: prior sex offenses, age at release, victim gender, and relationship to victim. Hanson (1997) describes research that supports his conclusion that sexual offense recidivism can

be usefully predicted through the consideration of a limited number of uncomplicated variables.

The development of the RRASOR was guided by the meta-analysis (Hanson & Bussiere, 1996; 1998). Seven development samples and one validation sample were utilized in the development of the instrument. The samples differed along a number of important variables, including recidivism criteria, length of follow-up, and participation in sex offender treatment. Because the recidivists and nonrecidivists were matched within each study, legitimate comparisons can be made to support the conclusion that higher scores on the RRASOR are associated with a higher rate of detected recidivism. The heterogeneity of the research samples suggests that the RRASOR is a robust instrument that can potentially be used across various sex offender populations, for males who have been confined (to prison or a secure, long-term treatment facility) for sex offenses.

In the development samples that were used to test the RRASOR, information was available on all the RRASOR predictor variables. In clinical practice, then, the RRASOR should only be utilized when sufficient reliable data are available to identify the number of prior sex offenses, age at release, victim gender, and relationship to victim(s).

MnSOST-R

The Minnesota Sex Offender Screening Tool-Revised (MnSOST-R; Epperson et al., 1998) is a 16-item scale that was developed from three samples of male sex offenders (totaling 387 people) released from a Minnesota correctional facility. Scoring the items on the MnSOST-R requires accurate, detailed information about the number of sex convictions, length of sex offending history, force used in committing the sex offense, age of victim, relationship of the offender to the victim, substance abuse history, employment history, and other information.

The MnSOST-R has been cross-validated on a sample of 95 sex offenders, also from the Minnesota prison system (Epperson, 2000a). Most intrafamilial child molesters were excluded from both the development and the cross-validation samples. The instrument might be useful for incarcerated rapists and extrafamilial child molesters outside Minnesota, but pending further research an evaluator who uses the instrument for non-Minnesotans must be prepared to respond to challenges to the assumption that a target offender is similar to offenders in Minnesota. Also, even if the released offenders in another area are simi-

lar to those in Minnesota, the meaning of MnSOST-R scores would be different if the environment in another area is different than it was in Minnesota in the 1990s.²

In cases where the offender's prior criminal history is unknown or partially unknown, and/or where the details of the offender's prior sexual offenses is unknown or partially unknown, use of the MnSOST-R is likely to be misleading. The MnSOST-R has a procedure for accommodating missing data, but this should only be used when most data are reliably known and a detail or two is missing. When substantial information is missing, or when it is unknown how much information is missing, the MnSOST-R should not be used.

SORAG

The developers of the Sexual Offender Risk Appraisal Guide (SORAG; Quinsey et al., 1998) recommend that evaluators not adjust predictions on the basis of factors not included in the prediction equation. They also recommend that the decision to confine a dangerous person should not require a finding of a mental disorder, but should be made on considerations of dangerousness alone. While these recommendations are at odds with existing laws in many areas, they do not preclude evaluators from using the SORAG in a manner consistent with local laws. But because Quinsey et al. (1998) have conducted the validation studies with the SORAG to predict the likelihood of any (sexual or nonsexual) violent re-offense, predictions that rely on the SORAG are vulnerable to arguments that what is predicted (any violent offense) is irrelevant to the legal question in many jurisdictions (likely to commit a new sexually violent act).

The SORAG consists of 14 items that, in the best of circumstances, can be scored from prison or treatment facility files. In practice, some information will rarely be available in the file, notably phallometric measures of deviant sexual arousal. Item 14, the Hare Psychopathy Checklist-Revised (PCL-R; Hare, 1991), can be scored from files if there is enough information, but in practice—at least in Florida—this is rarely the case. File information relevant to some SORAG items such as elementary school maladjustment, substance abuse history, and family and marital history may be unreliable due to the ultimate data source: unverified statements made by the offender. It is not unusual to find a description written as if it were the facts—such as, “no history of substance abuse”—without any indication that the offender's claim was ver-

ified by an independent, reliable source. As with the MnSOST-R, procedures are available for accounting for missing data, but again, substantial missing or unreliable data should preclude the use of the SORAG.

Static-99

Hanson has incorporated the RRASOR into the more recent Static-99 (Hanson & Thornton, 2000), and at the time of this writing, he recommends using the Static-99 instead of the RRASOR (Hanson, 2000). Hanson expects further developments, which will eventually lead to recommendations to use future instruments in place of the Static-99. The Static-99 was not used in the evaluations for this case study because it had yet to be developed. Because it is being used frequently at the time of this writing, and is currently required to be used in some areas, including Florida,³ some comments about the use of the Static-99 are in order.

Hanson and Thornton (2000) instruct: "Static-99 is intended for males aged at least 18 who are known to have committed at least one sex offense involving a child or a nonconsenting adult" (p. 134). Development samples substantially overlap with those used to develop the RRASOR; offenders were incarcerated or placed in secure, long-term treatment. The Static-99 is intended to be scored from the offender's official criminal record, but some information for some items can come from a clinical interview.

There are ten items on the Static-99: prior sexual offenses, prior sentencing dates, any convictions for non-contact sex offences, current convictions for non-sexual violence, prior convictions for non-sexual violence, unrelated victims, stranger victims, male victims, age at release, and marital status. The only procedure described for accommodating missing data is to score 0 if marital status is unknown (Phenix et al., 2000). Because there is no procedure for accommodating other missing data, use of the Static-99 when other data are missing or unreliable would constitute non-standard use of the instrument. Additional information (e.g., another sexual offense or establishing that a victim of a known offense was male) will typically lead to an increase or no change in Static-99 score. Therefore, if an evaluator were required (by contract or agency policy) to use the Static-99 in a case with some missing data, the Static-99 could at best be scored to indicate a range of possible scores (e.g., 4-11), based on known data and scores that would be possible if the missing data were known. While such an approach would still

constitute non-standard use of the instrument, it would be clearer and more accurate than scoring 0 for all the unknown items.

General Comments

Is it necessary for an evaluator in one state to be cautious about using actuarial instruments developed in another state? Yes it is. The goal of the MnSOST-R development team was to create a reliable and valid predictive instrument that could be easily scored by Minnesota's correctional case managers using only information routinely available to them from correctional records. While it is an open question whether Minnesota's sex offenders are similar to those elsewhere,⁴ it appears that Minnesota's sex offender files are different from those in some other states, including Florida. In reviewing files for an unpublished reliability study of the MnSOST-R in Florida, Epperson (2000b) noted that Florida's files are not as well documented as Minnesota's. This difference is so severe that Epperson (2000b) stated to a group of evaluators, "I think you're vulnerable to challenge about scoring any actuarial instrument, due to the incompleteness of the files in Florida."

I frequently communicate with classification officers in Florida prisons as I conduct evaluations of sex offenders relevant to civil commitment proceedings. Some acknowledge gaps in the database, in part due to chronic understaffing. Although some national databases are routinely searched for prior offenses, no information is sought regarding offenses in states that do not participate in the national databases. Thus, some prior offenses are never identified. For prior offenses that are identified, the details (including original arrest reports) are not always obtained, particularly for offenses that were committed out of state. The files often contain a Presentence or Postsentence Investigation (PSI) Report, but some of these are sparse. Some PSIs refer the reader to prior PSIs for details, but those prior PSIs are typically not available in the file. By the time an evaluator conducts a pre-release evaluation, some of the records that could have been obtained by a diligent, fully staffed classification department can no longer be obtained because the records have been expunged. This is particularly true for juvenile records, but frequently occurs for adult records as well. The result is that there is often not enough reliable information available about a given offender to accurately score assessment instruments.

In assessing risk for sexual recidivism, actuarial instruments are only useful when the target offender is similar to the offenders in research sample(s), and when the quantity and quality of the data regarding the

target offender is comparable to that available for the offenders in research sample(s). In some states, the quantity and quality of data regarding offenders is routinely not comparable to that in the research samples, and this should be considered in developing policies and contracts regarding which actuarial instruments, if any, should be required in risk assessments in a given state. In every case, the evaluator must determine whether or not a given instrument can be used appropriately. The following case illustrates obvious misuse of actuarial instruments.

CASE STUDY

Because the adversarial system is in full force and the stakes are so high—full freedom versus indefinite confinement—one might expect that sex offender re-offense risk assessments would be models of forensic psychology and psychiatry at their best. Although that should be the case, it is not hard to find examples of ill-prepared clinicians misusing unfamiliar instruments.⁵

A person was referred for assessment under Florida's Sexually Violent Predator Act. Three evaluators performed evaluations of him prior to a trial to determine whether he met criteria for civil commitment as a sexually violent predator. Because the relevant law includes a prong of "likely to [re-offend]," the evaluators included re-offense risk prediction as part of their evaluations. Two of the evaluators used actuarial instruments. That section of those reports is presented next.⁶

Evaluator 1's Risk Assessment

The first evaluator listed the following instruments: "The Hare PCL-R, Hanson's (1997) Rapid Risk Assessment for Sex Offense Recidivism (RRASOR), Minnesota Sex Offender Screening Tool (MnSost), and Risk Assessment Guide (RAG)."⁷

Evaluator 1's discussion of the results, *in toto*, is as follows: "Mr. X attained a RRASOR of 3, predicting a recidivism rate of 24.8% over five years and 36.9% over ten years against a similar population. The RAG indicated that the risk of recidivism for Mr. X, when compared to a similar population is 44% over seven years and 58% over ten years."⁸

Evaluator 2's Risk Assessment

Evaluator 2 listed and discussed risk assessment instruments in one paragraph: "Both static and dynamic risk factors were utilized in assess-

ing Mr. X's likelihood to engage in future violent sexual acts. The Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR), Minnesota Sex Offender Screening Tool-Revised (MnSOST-R), and the Sex Offender's Risk Appraisal Guide (SORAG) were all utilized in this assessment. Mr. X obtained a RRASOR score of 1, reflecting a recidivism rate below the base rate of the overall population of sex offenders. Mr. X's score on the SORAG reveals a recidivism rate of .07 over seven years and .09 over 10 years when compared to a similar population. The inmate score on the MnSOST-R does not reflect a significant risk for sexual recidivism."

Evaluator 1's Recommendation

Evaluator 1 relied, in part, on data from the actuarial instruments in offering an opinion: "It is predicted with a degree of certainty that he will be a repeat offender." The "degree of certainty" is not specified in the recommendations section of the report; presumably that is a reference to the percentages specified in the risk assessment section.

Evaluator 2's Recommendation

Evaluator 2 wrote, "In view of the limited static risk factors in this case, utilizing an actuarial assessment approach, it is this examiner's opinion that Mr. X does not meet the criteria for involuntary commitment."

DISCUSSION

Something is clearly wrong when one evaluator reports a 10-year recidivism risk of nearly 60% and another evaluator working on the same case at approximately the same time reports that the risk would be less than 10%! Why are their results so different? Consider the data that the evaluators entered into the actuarial prediction systems.⁹

At the time of these evaluations, reliable information about Mr. X consisted of the following. He was arrested in Florida in the early 1980s. At the time, he was an illegal alien. He was initially charged with sexual battery of a child and attempted murder with a deadly weapon. After plea bargaining, he was convicted of attempted sexual battery of a child and aggravated battery with a deadly weapon. Information about the offense was clear in the referral documents (including victim state-

ment, witness statements, and a confession), but will not be detailed here. No information was known about the man's identity or history. He had no identification documents. There was no confirmation of who he was, where he was from, or his history.

Hanson (1997) writes, "The predictive accuracy of the [RRASOR] was sufficient to justify its use as a screening instrument in settings that require routine assessments of sexual offender recidivism risk" (p. 1). But routine should not be confused with rote. In this case, there was inadequate information to score items 1 and 3. No one knew whether or not Mr. X had been arrested or convicted of a sexual offense in whatever country or countries from which he came. Therefore, the RRASOR could not be adequately scored, and should not have been used.

For the 16-item MnSOST-R, there was insufficient information to score items 1-3, 7, 8, and 10-12. For the 14-item SORAG, the following items could not be scored: 1-9, 11, 13, and 14. These instruments could not be adequately scored, and should not have been used.¹⁰

Both evaluators made misleading statements about the meaning of scores on the actuarial instruments. Even when there is sufficient data to properly score an instrument, the evaluator should not report that there is a recidivism rate of $x\%$ over 5 years and $y\%$ over 10 years. It would be less misleading for an evaluator to say that a given individual shows characteristics similar to a group of people in which, say, 3 out of 10 committed a new sexual offense in 15 years. The level of precision implied in saying that the recidivism rate is 36.9% is also misleading. And as independent studies of actuarial instruments are performed in new samples, it is likely that higher scores on the instruments will be associated with higher recidivism rates, but the actual recidivism rates for individual scores may vary from study to study (Dempster, 1998; Slovic et al., 2000).

These egregiously flawed risk assessments illustrate a problem that I have observed frequently, albeit usually to a less extreme degree.¹¹ Why does this occur? Boer et al. (1997) write, "[A]ctuarial risk assessment tends to disengage professionals from the evaluation process. By design, actuarial instruments are so structured that they require minimal professional judgment. Unless evaluators are sufficiently schooled in psychometric theory to have a healthy respect for the strengths and limitations of test data, professionals may tend to over- or under-utilize actuarial data when making decisions about individuals" (pp. 5-6). By design, when one enters inadequate or otherwise flawed data into an actuarial prediction system, the results will be flawed. Across a variety of fields, this process is known as "garbage in, garbage out." Often abbreviated as

GIGO, this is a famous computer axiom meaning that if invalid data are entered into a system, the resulting output will also be invalid. Although originally applied to computer software, the axiom holds true for all systems, including, for example, decision-making systems.

Should the test developers share the blame for this type of misuse of actuarial instruments? Perhaps. At the time of this writing, the instructions for the RRASOR, Static-99, and MnSOST-R are available on the Internet, but that information does not include the breadth of information one can find in the manual of more established instruments such as the PCL-R (Hare, 1991). These shortcomings are described in more detail elsewhere (Otto & Petrila, in press; Petrila & Otto, in press). A good deal of the information that would go into a manual for the SORAG is included Quinsey et al. (1998), but is not organized in a convenient way for new users. As developers of actuarial instruments declare them ready for use in clinical cases, the developers should produce a detailed manual that reports data on reliability, validity, and other information consistent with the current Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999). Whenever possible, evaluators should attend training seminars for each instrument prior to using it.

What is an evaluator to do when there is insufficient information to properly score any of the actuarial instruments? Such an evaluator is not unarmed.

GUIDED CLINICAL ASSESSMENT

Hanson (1998, citations omitted) writes:

Given the current state of knowledge, I believe that there are three plausible approaches to conducting risk assessments: guided clinical, pure actuarial, and adjusted actuarial. In the guided clinical approach, expert evaluators consider a wide range of empirically validated risk factors and then form an overall opinion concerning the offender's recidivism risk. In contrast, the pure actuarial approach evaluates the offender on a limited set of predictors and then combines these variables using a predetermined, numerical weighting system. The adjusted actuarial approach begins with an actuarial prediction, but expert evaluators can then adjust (or not) the actuarial predic-

tion after considering potentially important factors that were not included in the actuarial measure.

Petrila and Otto (in press, citations omitted) write: "Although actuarial methods have the potential to produce more reliable and more valid decisions than those made using clinical methods, there is little evidence that actuarial judgments of risk for violent or sexual reoffending are more reliable or more valid than other types of assessments for reoffending." Hanson (1998) expects that actuarial methods will eventually prove superior to clinical assessments in their ability to predict recidivism, but acknowledges, "[T]he research on actuarial measures for sexual offense recidivism has yet to demonstrate a clear superiority to the best clinical assessment methods."

Meanwhile, the one study that has directly compared a guided clinical approach to actuarial instruments "did not reveal a clear superiority for actuarial assessments over clinical judgments of risk, nor for actuarial scales that are empirically constructed over non-empirically constructed scales. In short, it appears that actuarial assessments are not always superior to clinical judgments, at least when those judgments are structured around empirically validated and theoretically derived variables. In addition, there is no evidence to support empirical construction as the preferred method of producing a risk assessment instrument" (Dempster, 1998, p. 68).

Thus, when the quantity and quality of information about an offender are abundant, an evaluator could choose from available actuarial or guided clinical approaches, or use both actuarial and guided clinical instruments (see Dempster, 1998, for a data-based recommendation for a mixed battery). In cases such as the present one, when too little information is available to properly score actuarial instruments, a guided clinical approach is clearly superior.

An evaluator using a guided clinical approach utilizes an a priori list of risk factors such as the Sexual Violence Risk-20 (SVR-20). The manual (Boer et al., 1997, p. 25) includes: "The SVR-20 is an assessment method or procedure rather than a test or scale. Although it is an attempt to systematize the assessment of individuals, it is not sufficiently structured or standardized to be a test and does not yield norm-referenced or criterion-referenced scores. Similar procedures have been developed for use in a wide range of psycholegal assessments, and have been labeled elsewhere "forensic assessment instruments" (Grisso, 1986). The SVR-20 includes eleven items regarding psychosocial adjustment (e.g., sexual deviation, psychopathy, substance abuse problems), seven items

regarding sexual offenses (e.g., high density, physical harm to victim(s), use of weapons or threats of death), and two items regarding future plans (including attitude toward intervention).

Evaluator 3's Risk Assessment

The third evaluator in this same case used the SVR-20. Because the SVR-20 presents relevant risk factors but does not entail mathematical combination to produce a summary score, this instrument could have been used to show what risk factors were clearly present (e.g., physical harm to victim, use of a weapon), what risk factors were clearly not present (none), and what risk factors were unknown (most of the 20 items included in the SVR-20). That would be one useful way for the evaluators to help the judge or jury consider the information—and lack of information—relevant to Mr. X's risk for future sexual violence. Unfortunately, that is not the way the instrument was used.

Evaluator 3's discussion includes: "When coded on the SVR-20 Coding Sheet, Mr. X clearly scores on nine of the twenty risk factors: substance abuse problems, . . . relationship problems, . . . employment problems, . . . physical harm to victim, . . . uses weapons or threats of death, . . . extreme minimization/denial of offenses, . . . lacks realistic plans, . . . negative attitude toward intervention."¹²

Evaluator 3 only listed the SVR-20 items that he believed were present for Mr. X; he did not list those that he did not believe to be present, or that he recognized he had insufficient information to rate. There was sufficient information to rate a few of the SVR-20 items, including item 14 (physical harm to victim) and item 15 (uses weapons or threats of death). Evaluator 3 apparently used Mr. X's description to rate item 7 (relationship problems; "He has had no enduring relationships.") and item 8 (employment problems; "He has maintained no steady employment."), despite the significant possibility that Mr. X gave no information about relationships or employment to avoid giving any identifying information about himself. In fact, Mr. X's relationship history and employment history were as unknown as were his histories of sexual offenses, nonsexual criminal offenses, and supervision failure (e.g., violations of probation).

It is generally advisable for a forensic clinician to prepare the same report regardless of which side in a legal dispute has retained her services. By reporting only those risk factors that he deemed to be present and ignoring those that are absent or unknown, Evaluator 3 gives the appearance of bias. A fairer and clearer use of the SVR-20 would be to list

all the items, along with data to show whether and to what degree each item is present, absent, or unknown.

CONCLUSIONS AND RECOMMENDATIONS

All three of the evaluations described above were seriously flawed. Yet it was largely on the basis of these assessments that a group of citizens had to make an important decision: whether to involuntarily and indefinitely commit a man who had served his time, or to give him his freedom and hope that he does not attack another child.

Janus and Meehl (1997) highlight the importance of these risk predictions:

There is a continuing debate about whether predictions of dangerousness are accurate enough to support deprivation of liberty. As in other civil commitment settings, the stakes in these determinations are high. The debate is especially important in the context of sex offender commitments, because the consequences of the predictions are so severe. If predictions about future violence are too optimistic, sexual violence may result. Unduly pessimistic predictions result in unnecessary, prolonged deprivations of liberty. In addition, sex offender commitments entail treatment that is expensive and intrusive, while sexual violence is exceedingly destructive. Thus, both types of prediction errors are costly in many ways. (p. 34)

This case study has implications at several levels. We begin at a broad, policy level.

National and International Policymakers

Assuming that sex offender re-offense risk predictions will continue to be relevant to some legal decisions (e.g., civil commitment of sexually violent offenders), it is likely that actuarial methods will prove useful when there is sufficient accurate data. Risk predictions will be enhanced if detailed information about sex offenders' crimes is gathered and archived, and if the files are not expunged, purged, or destroyed. Pre- and post-sentencing reports, psychological evaluations, and social histories contain important, relevant information, and should be maintained. Valuable information is lost if juvenile records are sequestered or destroyed. An accurate national, and even international,

database is increasingly valuable as people, including sex offenders, are more mobile. Careful thought and debate should address the merits of competing scientific and privacy interests.

State Agencies and Contract Organizations

At the level of state agencies and contracting organizations, it is necessary to do a detailed analysis of the type of offenders, the type of environment to which the offenders will be returned, and the type of records in the research sample(s) for each actuarial instrument, and compare this to the type of offenders, the type of environment, and the type of files typically available for the cases to be considered. It makes little sense to require that the evaluators in a given state use a particular instrument if the files in that state routinely lack information necessary to score that instrument.

Evaluators

Clinicians who accept referrals to do evaluations relevant to civil commitment of sexually violent offenders “have an obligation to maintain current knowledge of scientific, professional and legal developments within their area of claimed competence. They are obligated also to use that knowledge, consistent with accepted clinical and scientific standards, in selecting data collection methods and procedures for an evaluation” (Committee on Ethical Guidelines for Forensic Psychologists, 1991).

The demand for sex offender re-offense risk assessments has recently increased dramatically and is likely to continue to increase as more states enact sexually violent predator commitment laws. Psychologists and psychiatrists who accept referrals to perform such assessments must accept responsibility to undergo adequate training, carefully choose assessment instruments, utilize instruments properly, and present reports and testimony in a clear and careful manner. Evaluators must recognize and clearly communicate the limitations of the risk assessment instruments they use, and must not misuse actuarial instruments when the input data are unreliable or insufficient.

When available data are too sparse or unreliable to adequately score actuarial instruments, evaluators can appropriately present risk data via a guided clinical approach such as the SVR-20. All of the *a priori* risk factors should be listed, with indications of whether and to what degree each is present, absent, or unknown.¹³

As an evaluator prepares to offer testimony on a particular case, there will be pressures to formulate an opinion and advocate strongly for that opinion. Along with external pressure from the attorney arguing the case, the evaluator may feel internal pressure to present results forcefully and to “withstand” cross examination. Evaluators must resist the temptation to focus on apparent strengths of an assessment instrument in a way that distracts from real weaknesses in the data. When data about relevant risk factors are missing or unreliable, the evaluator can best assist the trier of fact by identifying relevant risk factors and describing which are present, absent, or unknown.

NOTES

1. At present, though, none of the actuarial instruments are constructed to predict different recidivism rates for child molesters and rapists.

2. For example, Area A might have a different recidivism rate than Area B, where Area A requires all released sex offenders to be on lifetime parole with close supervision and continuing outpatient sex offender treatment, and Area B requires no parole or outpatient treatment following release. If so, then offenders in Area A with a MnSOST-R score of 8 would be expected to recidivate at a different rate than offenders with the same MnSOST-R score who are released in Area B.

3. A private agency contracts with the state to conduct the evaluations, and sub-contracts with individual evaluators who actually do the evaluations. The private agency requires evaluators to use the Static-99 “unless contraindicated.” Evaluators responding to requests from the Respondent (usually represented by the Public Defender) or the petitioner (the State Attorney) are not bound by conditions of that contract, and can choose which instruments, if any, to use.

4. Another open question is the similarity of the environments (including level of supervision) to which offenders will be released.

5. *Kansas v. Hendricks* was decided in 1997. Florida’s civil commitment law went into effect in 1999, and these evaluations were done in 1999.

6. The evaluators included some caveats regarding use of the instruments, which are not included here.

7. Traditional clinical instruments, including the MMPI-2 and the Rorschach, were used by one or more evaluators, but are not listed or discussed in this article.

8. A number of criticisms could be made regarding the limited description of the instruments, including the facts that: (a) Evaluator 1 did not specify whether he was using the VRAG (which is used to assess risk for re-offense generally) or the SORAG (which is used to assess risk for re-offense by sex offenders), (b) no raw score was reported for the “RAG,” (c) no score of any kind was reported for the PCL-R, and (d) none of the evaluators gave sufficient details to show the basis for the scores (i.e., item scores).

9. Although this largely accounts for the unreliability of the predictions, the direction of the difference may have been influenced by the following. The evaluator who

reported a 60% recidivism rate was retained by the state; the evaluator who reported a 10% rate was retained by the respondent.

10. For comparison, consider the appropriateness of using an objective personality test when half of the responses are missing or illegible.

11. Examples include using actuarial instruments for cases in which it is known that the respondent had a juvenile record, but the record has been destroyed; and cases in which it is known that the offender has prior convictions for sex offenses, but details (including age of victim, whether force was used, etc.) are unknown.

12. Although Evaluator 2 said there are nine identified risk factors, I found only eight listed in his report.

13. If, due to agency or contract requirements, an evaluator must use an actuarial instrument in such a case, then the least harm can be done by, again, indicating whether and to what degree each item is present, absent, or unknown. Actuarial scores, if reported, should describe a range of what scores would be possible if missing data were known.

REFERENCES

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for educational and psychological testing (3rd ed.)*. Washington, DC: Authors.
- Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). *Manual for the sexual violence risk-20*. Vancouver, BC: British Columbia Institute Against Family Violence.
- Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior, 15*, 655-665.
- Dempster, R. J. (1998). *Prediction of sexually violent recidivism: A comparison of risk assessment instruments*. Unpublished master's thesis, Simon Fraser University.
- Doren, D. M. (1998). Recidivism base rates, predictions of sex offender recidivism, and the "sexual predator" commitment laws. *Behavioral Sciences and the Law, 16*, 97-114.
- Epperson, D. L., Kaul, J. D., & Hesselton, D. (1998, October). *Final report of the development of the Minnesota sex offender screening tool-revised (MnSOST-R)*. Presentation at the 17th Annual Research and Treatment Conference of the Association for the Treatment of Sexual Abusers, Vancouver, BC, Canada.
- Epperson, D. L. (2000a). Cross-validation update. (http://psych-server.iastate.edu/faculty/epperson/mnsost_download.htm; retrieved April 9, 2001).
- Epperson, D. L. (2000b, October). *The assessment of the sexually violent predator using the Minnesota sex offender screening tool-revised*. Presentation for the Florida Department of Children and Families, Orlando, FL.
- Grisso, T. (1986). *Evaluating competencies: Forensic assessments and instruments*. New York: Plenum.
- Grisso, T. (2000, March). *Ethical issues in sex offender re-offense risk prediction*. Presentation at the Sex Offender Re-Offense Risk Prediction Symposium, Madison, WI.
- Hanson, R. K. (1996). Child molester recidivism. *Research Summary: Corrections Research and Development, 1*, 1-2.

- Hanson, R. K. (1997). *The development of a brief actuarial scale for sexual offense recidivism (User Report No. 1997-04)*. Ottawa, Ontario, Canada: Department of the Solicitor General.
- Hanson, R. K. (1998). What do we know about sex offender risk assessment? *Psychology, Public Policy, and Law*, 4, 50-72.
- Hanson, R. K. (2000, March). *Meta-analysis of sex offender recidivism risk studies: Rapid risk assessment of sexual offense recidivism (RRASOR), Static-99, Sex offender needs assessment rating (SONAR)*. Presentation at Sex Offender Re-Offense Risk Prediction Symposium, Madison, WI.
- Hanson, R. K. & Bussiere, M. T. (1996). *Predictors of sexual offender recidivism: A meta-analysis (User Report 96-04)*. Ottawa, Ontario, Canada: Department of the Solicitor General.
- Hanson, R. K. & Bussiere, 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., 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. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. *Law and Human Behavior*, 24, 119-136.
- Hare, R. D. (1991). *Manual for the psychopathy checklist-revised*. Toronto: Multi-Health Systems.
- Janus, E. S. & Meehl, P. E. (1997). Assessing the legal standard for prediction of dangerousness in sex offender commitment hearings. *Psychology, Public Policy, and Law*, 3, 33-64.
- Kansas v. Hendricks*, 521 U.S. 346 (1997).
- Otto, R. K & Petrila, J. (in press). Admissibility of testimony based on actuarial scales in sex offender commitments: A reply to Doren. *Sex Offender Law Report*.
- Petrila, J. & Otto, R. K. (in press). Admissibility of expert testimony in sexually violent predator proceedings. In A. Schlank & F. Cohen (Eds.). *The sexually violent offender (Volume II: Legal issues, clinical issues, special populations)*. Kingston, NJ: Civic Research Press.
- Phenix, A. & Hanson, R. K., & Thornton, D. (2000). Coding rules for the Static-99. (<http://www.sgc.gc.ca/epub/corr/ecodingrules/ecodingrules.htm>; retrieved April 9, 2001).
- Prentky, R. A., Lee, A. F. S., Knight, R. A., & Cerce, D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and Human Behavior*, 21, 635-659.
- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (1998). *Violent offenders: Appraising and managing risk*. Washington, DC: American Psychological Association.
- Slovic, P., Monahan, J., & MacGregor, D. G. (2000). Violence risk assessment and risk communication: The effects of using actual cases, providing instruction, and employing probability versus frequency formats. *Law and Human Behavior*, 24, 271-296.
- Winick, B. J. & LaFond, J. O. (Eds.). (1998). Sex offenders: Scientific, legal, and public policy perspectives [special issue]. *Psychology, Public Policy, and Law*, 4(1/2).

RECEIVED: May 30, 2002
REVISED: August 5, 2002
ACCEPTED: August 6, 2002