

CONTEXTUALIZED RISK ASSESSMENT IN CLINICAL PRACTICE: UTILITY OF  
ACTUARIAL, CLINICAL AND STRUCTURED CLINICAL APPROACHES TO  
PREDICTIONS OF VIOLENCE

Rebecca L. Jackson, M.S.

Dissertation Prepared for the Degree of  
DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2004

APPROVED:

Richard Rogers, Major Professor  
Kenneth W. Sewell, Committee Member  
Craig Neumann, Committee Member  
Daniel Shuman, Committee Member  
Linda Marshall, Interim Chair of the Department  
of Psychology  
Sandra L. Terrell, Dean of the Robert B. Toulouse  
School of Graduate Studies

Jackson, Rebecca L., *Contextualized risk assessment in clinical practice: Utility of actuarial, clinical, and structured clinical approaches to predictions of violence*. Doctor of Philosophy (Clinical Psychology), August 2004, 140 pp., 25 tables, references, 143 titles.

Assessing offenders' risk of future violent behavior continues to be an important yet controversial role of forensic psychologists. A key debate is the relative effectiveness of assessment methods. Specifically, actuarial methods (see Quinsey et al., 1998 for a review) have been compared and contrasted to clinical and structured clinical methods (see e.g. Hart, 1998; Webster et al., 1997). Proponents of each approach argue for its superiority, yet validity studies have made few formal comparisons. In advancing the available research, the present study examines systematically the type of forensic case (i.e., sexual violence versus nonsexual violence) and type of assessment method (i.e., actuarial, structured clinical, and unstructured clinical). As observed by Borum, Otto, and Golding (1993), forensic decision making can also be influenced by the presence of certain extraneous clinical data. To address these issues, psychologists and doctoral students attending the American Psychology Law Society conference were asked to make several ratings regarding the likelihood of future sexual and nonsexual violence based on data derived from actual defendants with known outcomes. Using a mixed factorial design, each of these assessment methods were investigated for its influence on decision-makers regarding likelihood of future violence and sexually violent predator commitments. Finally, the potentially biasing effects of victim impact statements on resultant decisions were also explored.

## TABLE OF CONTENTS

LIST OF TABLES .....	iv
Chapter	
1. INTRODUCTION.....	1
History of Violence Prediction in Psychology and Psychiatry	
Clinical Constructs of Dangerousness and Risk Assessment	
Reviving Risk Assessment Research	
Actuarial Prediction	
Clinical Judgment	
Structured Clinical Judgment	
Applying Risk Assessment Techniques to Clinical Practice	
Sexually Violent Predator Statutes	
Victim Impact Statements	
Current Study and Research Questions	
2. METHOD.....	43
3. RESULTS .....	53
4. DISCUSSION .....	81
Implications for Sexually Violent Predator Evaluations	
Recommendations for Conducting SVP Evaluations	
Next Steps Toward Validating an SVP Measure	
Limitations of the Current Study and Directions for Future Research	
Summary and Conclusions	
APPENDIX A .....	110
APPENDIX B .....	113
APPENDIX C .....	115
APPENDIX D .....	117
APPENDIX E .....	119
APPENDIX F .....	121
APPENDIX G .....	123

REFERENCES .....126

## LIST OF TABLES

### Table

1. Areas Under the Curve for Actuarial Risk Assessment Instruments in Predicting Serious Offenses and Sexual Offenses .....	17
2. Schematic Depiction of 6 X 3 Factorial Research Design.....	44
3. Training and Specialization of Study Participants .....	45
4. Education/Occupation of Study Participants.....	53
5. Risk Assessment Experience and Percentage of Forensic Practice by Specialty Training in Psychology.....	54
6. Risk Assessments and Forensic Practice of Doctoral-Level Clinicians (i.e., Clinical and Counseling Psychologists) Versus Other Psychologists and Graduate Students .....	55
7. Violence and Sexual Violence Likelihood Ratings and Indicated Support for SVP Commitment by Violent and Nonviolent Predictions .....	56
8. Utility Estimates Calculated for Clinicians' Ability to Predict Sexual Violence.....	58
9. Categorical Prediction Accuracy (Violent versus Nonviolent) Utilizing a Single Decision Making Method.....	59
10. Hit Rate and $X^2$ Statistics for Clinical Decision Making Regarding Predictions of Nonsexual and Sexual Violence.....	60
11. Likelihood Estimates (Expressed in Percentages) of Nonsexual and Sexual Violence for Nonsexually Violent Offenders by Clinical Decision Making Method.....	61
12. Likelihood Estimates (Expressed as Percentages) of Nonsexual and Sexual Violence for Sexually Violent Offenders by Clinical Decision Making Method.....	63
13. Estimates of Sexual Violence: Effects of One Versus Two Assessment Methods .....	65
14. Testing Order Effects on Sexual Versus Nonsexual Prediction Accuracy (i.e., Hit Rate) by Assessment Method .....	66

15. Mean (SD) Ratings of Potential Violence Based on Actuarial and Actuarial Plus Clinical Data .....	67
16. Clinicians' Confidence in Their Ratings of Future Violence (Means/Standard Deviations) .....	69
17. Confidence of Sexual and Nonsexual Violence Predictions when Judgments Are Rendered Following One Decision Making Method Versus a Combined Assessment Approach.....	69
18. Confidence of Accurate Versus Inaccurate Clinicians' Estimates of Sexual and Nonsexual Violence .....	70
19. Means (Standard Deviations) for Amenability to Treatment Estimates Based on Single and Combined Clinical Decision Making Methods .....	71
20. Professionals' Support of Sexually Violent Predator (SVP) Commitment by Predicted Violence Type.....	73
21. Overall Hit Rates and $X^2$ for Sexual and Nonsexually Violent Outcomes by Level of Training .....	74
22. Stepwise Multiple Regression: Predictors of Support for SVP Commitment .....	76
23. Biasing Effects of Victim Impact Statement on Clinicians' Composite Predictions of Sexual and Nonsexual Violence (Means/Standard Deviations).....	77
24. Magnitude of Violence Prediction Bias Following Review of a Victim Impact Statement by Primary Assessment Method.....	78
25. Separate Discriminant Function Analyses: Predicting Over- and Underpredictors of Sexual Violence Based on Clinician Characteristics .....	80

## CHAPTER 1

### INTRODUCTION

Psychologists are often called upon to perform forensic assessments on specific psycholegal issues. In the clinical domain, prime examples of these assessments include competency to stand trial (CST) and insanity evaluations. In the past two decades, psychologists have become increasingly involved in risk assessments for the purpose of predicting violent and dangerous behavior (see, e.g., Monahan et al., 2001), particularly with the resurgence of sexually violent predator laws (SVP; see e.g., Community Protection Act [Revised Code of WA], 1990; Kan. Stat. Ann. 59-29[a], 1994) allowing for the involuntary commitment of “sexually violent predators.”

Societal restrictions of persons with mental disabilities is accomplished using two sources of control: *parens patriae* and *police power*. Prior to the 1960s, involuntary civil commitment was a common vehicle utilized to treat mentally ill individuals (Norko, 2000). Commitment was used to provide treatment until the individual was again able to care for himself or herself. This *parens patriae* approach to patient management emphasizes the psychologist’s role as helper and protector, extending the 18<sup>th</sup> century tradition of familial care of the mentally ill. With the population growth in America, institutions were created to house the mentally ill and provide treatment (Falk, 1999). In this manner, society adopted the role previously held by the family. *Parens patriae* considerations remained the primary avenue for committing mentally ill individuals until the 1960s (Falk, 1999; Monahan et al., 2001; Norko, 2000). Legal and societal forces converged to change the practice and purpose of civil commitment statutes beginning in the 1960s and 1970s. *Police power* justifications for committing the mentally ill were becoming prominent (Janus, 2000; *Lessard v. Schmidt*, 1976; Norko, 2000). For example, a federal district

court found in *Lessard v Schmidt* (1976) that a pure *parens patriae* justification for involuntary civil commitment was insufficient. The judges ruled that involuntary commitment was only justifiable when a clear likelihood of harm to self or others was evident. Several states followed suit, revising their commitment laws to limit the practice of *parens patriae* commitments. *Police power* considerations became predominant in commitment laws, depriving individuals of their freedom only when they present an extreme danger to themselves or others. The “dangerousness” requirement had two large effects on the practice of civil commitment: (a) fewer non-violent mentally ill individuals were provided with inpatient treatment, and (b) the purpose of commitment became less about protecting the patient (*parens patriae*) and more about protecting society from that patient (*police power*). In addition to the effects of the “dangerousness” requirement, the deinstitutionalization movement was also limiting the number and type of individuals being hospitalized. The new social climate viewed civil commitment not as a benevolent force, but as a restrictive power that deprived individuals of their civil liberties. Changes in the process of civil commitment led Stone (1975) to voice concern that legitimately mentally ill individuals were being denied treatment and that civil commitment was increasingly becoming an avenue for preventive detention.

The adoption of police power and focus on dangerousness required psychologists and other mental health professionals to become involved in predicting dangerous behavior (Monahan, 1981; Steadman & Coccoza, 1974; Thornberry & Jacoby, 1979). Clearly, commitments predicated on dangerousness necessarily require accurate predictions of dangerous behavior. The study of risk assessment emerged from this fundamental paradigm shift in the 1970s and evolved into a field of research that is freshly energized and forensically relevant. In the present review, the development of “risk assessment” as a program of research is addressed



in order to provide a framework for the current investigation. As such, the review of risk assessment identifies deficiencies in the knowledge base and reveals important avenues for future research.

### History of Violence Prediction in Psychology and Psychiatry

Psychology and psychiatry's role in determinations of dangerousness is rooted in American constitutional and case law (see e.g., *Barefoot v. Estelle*, 1983; *Baxstrom v. Herold*, 1966; *Kansas v. Hendricks*, 1997; *Schall v. Martin*, 1984). This section briefly chronicles mental health professionals' involvement in legally sanctioned predictions of future violence.

The U.S. Supreme Court's 1966 decision in *Baxstrom v. Herold* provided a rare opportunity to empirically investigate mental health professionals' ability to accurately predict future violence. Briefly, Baxstrom, a prisoner in New York's correctional system, was diagnosed as mentally disordered and transferred to a hospital for the criminally insane shortly before the expiration of his prison term. Despite New York's commitment law requiring a judicial finding of dangerousness, Baxstrom was detained at the hospital past the expiration of his original prison sentence without a hearing to determine his present dangerousness. Consequently, Baxstrom argued his right to equal protection had been violated. The Court agreed, finding that Baxstrom's rights had been violated by committing him without benefit of current finding of dangerousness as was required under New York law for all other civil commitments.

As a result of *Baxstrom*, 967 offenders being held in hospitals without judicial determinations of dangerousness were released to lower security civil hospitals throughout New York State. This decision presented the unique opportunity to assess the accuracy of clinicians' predictions of dangerousness as all offenders had been deemed too dangerousness to be released from maximum security hospitals. Raising doubts about clinicians' expertise, these offenders had

low rates of reoffending. Overall, only about 20% of the patients were subsequently violent according to hospital and police reports. In stark contrast to predictions of dangerousness, many patients did surprisingly well. For example, 176 patients (18%) were discharged to the community within 1 year of transfer, and only 7 (less than 1%) were subsequently returned to secure hospitals. Over a 4.5 year period, more than half of the patients were discharged to the community and fewer than 3% returned to secure hospitals (Monahan & Steadman, 2001; Steadman & Cocozza, 1974).

Proving the Baxstrom cohort was not unique, Thornberry and Jacoby (1979) published a parallel study of 586 patients released from Pennsylvania's Farview Institution (*Dixon v. Pennsylvania*, 1971). The 3-year general recidivism rate was 23.7%. Only 14% of the total sample were rearrested or returned for a violent offense during the four-year follow-up period.

Both the Baxstrom and Dixon studies served to highlight two very important findings: (a) the base rate of violence among presumably "dangerous" mentally disordered offenders was relatively low, and (b) clinicians greatly overestimated the likelihood of future violence.

Early research (Goldberg, 1968; Oskamp, 1965; Quinsey & Ambtman, 1979; Steadman & Cocozza, 1974; Thornberry & Jacoby, 1979) on predictive accuracy suggested that clinicians had no special abilities to predict violent behavior. In particular, these studies suggested that training, experience, and confidence of clinicians had very little effect on their accuracy. For example, Quinsey and Ambtman (1979) demonstrated that experienced psychiatrists' predictions of violence were no better than laypersons' (i.e., school teachers) judgments on the same patients. Goldberg's (1968) review of the prediction literature cited several studies echoing this result: the amount of professional training and experience does not relate to predictive accuracy. In addition, he noted that the addition of information was not related to the accuracy of

participants' resulting inferences. While not specific to predictions of dangerousness, Oskamp (1965) reported that the confidence of decision makers increases with the amount of information available; however, the accuracy of those predictions does not improve.

The combined influence of the above findings led many mental health professionals to conclude that accurate predictions of dangerousness simply cannot be accomplished. The American Psychiatric Association (1974) declared in its task force report, *Clinical Aspects of the Violent Individual*, "Psychiatric expertise in the prediction of dangerousness is not established and clinicians should avoid conclusory judgements in this regard" (p. 30).

Monahan's (1981) *The Clinical Prediction of Violent Behavior* discredited the tradition of clinical predictions of dangerousness. This monograph articulated the many inaccuracies of clinical predictions of violent behavior and concluded that "psychiatrists and psychologists are accurate in no more than one out of three predictions of violent behavior" (p. 47). This work had such a profound influence on attitudes towards predictions of violence that Otto (2001) ironically referred to it as "too good a book" on the fallibility of clinical judgment, noting that it essentially eliminated research interest in the topic for many years.

Pessimism about dangerousness predictions deterred researchers from the topic. Practically speaking, prediction research was nonexistent in the 1980s. Paradoxically, this same period saw American courts deciding on many cases that directly expanded the clinician's role in assessing dangerous behavior.

For example, the Supreme Court's decision in *Barefoot v. Estelle* (1983) served to make clinical predictions of dangerousness virtually unavoidable. In *Barefoot*, the petitioner objected to the use of psychiatrists' testimony regarding his risk for future dangerousness based on the assertion that psychiatrists were not competent to predict future violence. In conjunction with

this case, the American Psychiatric Association (APA) submitted an *amicus* brief detailing their view of expert psychological testimony regarding dangerousness. In its brief, the APA cited the available research, including Monahan's (1981) analysis, and concluded that psychiatrists have no expertise in predicting future violence and asserted that laypersons could do as well as "experts." The brief informed the Court that psychiatrists were wrong more often than they were right and had a tendency to over-predict violence (Monahan, 1981; Steadman & Coccoza, 1974; Thornberry & Jacoby, 1979). The Court's response to the brief and its resultant opinion were remarkable.

Essentially, the Court ignored the APA's brief, stating that "we are [not] convinced that the view of the APA should be converted into a constitutional rule barring an entire category of expert testimony" (*Barefoot v. Estelle*, 1983, p. 3387) and that they were "not persuaded that such testimony is almost entirely unreliable" (p.3398). Despite the "experts" themselves providing evidence of their inability to predict accurately, the Court found that "neither petitioner nor the Association suggests that psychiatrists are always wrong with respect to future dangerousness, only most of the time" (p. 3398). Bolstering their argument, the Court also relied on an earlier decision (*Addington v. Texas*, 1979) which reinforced the mental health professional's role in legal proceedings. *Addington*, in referring to commitment hearings, stated, "whether the individual is mentally ill and dangerous to either himself or others . . . turns on the meaning of the facts *which must be interpreted by expert psychiatrists and psychologists*" (*Addington v. Texas*, 1979, p. 1811; emphasis added). Ironically, the prediction of dangerous and violent behavior was becoming an integral and possibly unavoidable part of psychologists' and psychiatrists' duties (Monahan et al., 2001).

The legal system has continued to evidence greater confidence than many mental health professionals in the accuracy of clinical predictions. The American Bar Association's (1998) position on psychologists' role in risk assessments is highlighted in its *National Benchbook on Psychiatric and Psychological Evidence and Testimony* which stated that "it [clinical evaluation] is the best information available. The alternative is to deprive fact finders, judges and jurors of the guidance and understanding that psychiatrists and psychologists can provide" (p. 49). Further, Monahan et al. (2001) asserted that the general expectation is that mental health professionals can distinguish with a reasonable degree of accuracy between dangerous and nondangerous persons with mental disorders (see also Monahan, 2000; Mossman, 2000). Clearly, mental health professionals have been drafted into the field of violence risk assessment, some more readily than others.

The professionals who either choose or are mandated to perform risk assessments are expected to have the requisite knowledge and expertise (Monahan et al., 2001), despite their lack of formal training in risk assessment methods. As a result of this expectation, the question has shifted from (a) "Should mental health professionals be making violence predictions" to (b) "How do mental health professionals improve their predictions of violence?" (see Grisso & Appelbaum, 1992).

#### Clinical Constructs of Dangerousness and Risk Assessment

The "dangerousness" standard for civil commitment presented many difficulties for forensic psychologists. The apparent purpose of this standard is to ensure that only the most dangerous individuals are deprived of their civil liberties and freedom. Steadman (2000) argued that this requirement increased the difficulty of the psychologist's task rather than clarifying it because the term "dangerous" does not have a clinically meaningful definition. Instead, he

argued, the clinical parallel to the legal concept of dangerousness is “risk of future harm.” The difference between “dangerous” and “risk of future harm” can be construed as a categorical versus dimensional judgment. Determinations of dangerousness imply a binary (yes or no) outcome. In contrast, risk is inherently a probabilistic judgment. Employing the term “risk assessment” reflects the belief that the psychologist’s task is one of judging probabilities, not arriving at a dichotomous decision.

From one perspective, the distinction between binary predictions of dangerousness and probabilities of future violence is useful. Construing violence as a future probability allows us to focus on those variables that increase or decrease risk. It allows us to be flexible in our assessments and tailor treatment to those variables likely to decrease risk. From another perspective, this distinction is likely to be spurious in certain contexts. Regardless of the probabilistic nature of judgements, the resulting outcome remains binary. In the case of civil commitment, most decisions are comprised of yes/no decisions about involuntary hospitalization. The fact remains that legal decisions (i. e., commit or not to commit) and psycholegal constructs (i. e., gradations of risk) appear at odds.

Hart (1998; see also Steadman, 2000) argued for the replacement of “dangerousness” with “risk assessment” by pointing out that dangerousness implies a stable attribute about an individual. He contended that research does not support a single, global trait that operates across all individuals and situations to produce violent behavior. Instead, he defined risk assessment as “the process of evaluating individuals to (1) characterize the likelihood they will commit acts of violence and (2) develop interventions to manage or reduce that likelihood” (p. 122). Further, Hart argues against the term “predicting violence” on similar grounds. Predicting violence implies that the clinician is a detached observer who makes predictions and subsequently takes

no action based on the prediction. In stark contrast, the express purpose for making those predictions, according to Hart, is to facilitate management of the risk. In this manner, clinicians are active participants in the evaluation process and therefore, the term “risk assessment” most accurately captures current practice.

The term “risk assessment” will be employed throughout this review because it is the most comprehensive term used to describe psychologists’ activities. As noted above, the first prong of this term relates to characterizing an individual’s likelihood of future violence. In other words, the term risk assessment subsumes both predictions of dangerousness and predictions of violence. Likewise, a clinician’s responsibility may involve estimates of risk that change over time, situations, and legal contexts. The use of the term risk assessment also addresses important situational variables. The bulk of the present investigation focuses on the first prong of risk assessment, namely, assessment of the likelihood for future violence.

### Reviving Risk Assessment Research

Grisso and Appelbaum (1992) advocated a paradigm shift in the way psychologists viewed the risk assessment process. Instead of arguing the merits of psychologists’ involvement, they encouraged programmatic research aimed at increasing the validity of the assessments. At least two potential models have been proposed to improve risk assessments. The first model is offered by those individuals advocating the use of actuarial scales to predict violence (Quinsey, Harris, Rice & Cormier, 1998). This model articulates the position that clinicians are inaccurate predictors of violence, and therefore, should be eliminated from the equation. The second model involves efforts to increase clinicians’ accuracy in the risk assessment process via standardization of their clinical evaluation (e.g., structured clinical judgment).

Methodological advances have also been partially responsible for the revitalization of risk assessment research. Several researchers (Monahan, 1981; Mossman, 1994; Rice & Harris, 1995; Douglas, Ogloff, Nicholls, & Grant, 1999) have noted problems with the traditional utility estimates of prediction accuracy, such as positive predictive power, negative predictive power, sensitivity and specificity. These indices are influenced by base rates of behavior and clinicians' preferred errors, i.e., whether they are more or less conservative in predicting violence. Douglas et al. (1999) noted that traditional 2 X 2 contingency tables perform maximally at base rates of the criterion at 50%. In most research of violent recidivism, base rates are often less than 50%. Mossman (1994; see also Rice & Harris, 1995; Douglas et al., 1999; Swets, 1996) advocated for the adoption of Receiver Operating Characteristic (ROC) analysis for risk prediction research, noting that it is less influenced by base rates than traditional indices. Since Mossman's (1994) review and reanalysis of risk prediction research using ROC analysis, it has become a frequently-used index of accuracy among violence researchers.

The Receiver (or Relative; Quinsey et al., 1998) Operating Characteristic (ROC) was originally developed in communications and signal detection theory (Swets, 1996). With false alarms plotted on the *x axis* and hit rate plotted on the *y axis*, the ROC curve represents a ratio of false positives to accurate "hits." ROCs produce a statistical index called the Area Under the Curve (AUC). AUCs can range from 0 (perfectly wrong predictions) to 1.0 (perfect accuracy). AUCs of .5 indicate accuracy at chance levels. The value of the AUC of a test is equal to the probability that a randomly chosen, actually violent person, will score higher on the measure than a randomly chosen nonviolent person (Douglas et al., 1999; Rice & Harris, 1995; Mossman, 1994). In terms of clinical decision making, the value of AUC equals the likelihood that a clinician would rate a randomly selected violent person as more likely to be violent than a



randomly selected nonviolent person. Douglas et al. (1999) reported that AUCs of .70 and above are considered large.

Importantly, both utility estimates and ROC analysis are limited in their applicability to individual decisions. As outlined previously, utility estimates are vulnerable to fluctuations in base rate. ROC analysis partially controls for this phenomenon and instead yields an “overall” AUC. Individual cut points are absent, eliminating the ability to examine a clinician’s or a test’s performance at different cut scores. In other words, Positive Predictive Power and Negative Predictive Power are not discerned from ROC analysis. In making decisions regarding future violence, and reporting the associated error rates of these predictions, PPP and NPP are perhaps of paramount importance.

The following sections review the research in actuarial and clinical predictions of violence. In most cases, ROC analysis was utilized and AUCs reported. When appropriate, this review also reports traditional analyses of predictive accuracy such as correlations and 2 X 2 contingency tables.

#### Actuarial Prediction

As an early and vocal advocate of statistical methods of predictions, Meehl (1954/1996) argued for the superiority of actuarial over clinical predictions of psychological and behavioral phenomena. Monahan (1981) applied this idea specifically to predictions of violent behavior, demonstrating that clinical predictions of future violence were wrong more often than not. His 1981 monograph influenced many professionals and researchers who began to investigate the feasibility of actuarial (i.e., statistical) predictions of violence.

Early attempts by Quinsey, Pruesse, and Fernley (1975; also Pruesse & Quinsey, 1977) to apply actuarial techniques to clinical predictions of violence found mixed results. In a sample of

60 involuntary psychiatric patients who were discharged to the community, 18 (30.0%) were rearrested during a 39-month period. Quinsey and colleagues constructed a scale in which one point was assigned for each of the variables that discriminated between successes and failure. The variables included: (a) diagnosis of a personality disorder, (b) age < 31 at discharge, (c) < 5 years spent in psychiatric hospitals, (d) admission offense not against persons, and (e) separation from a parent before age 16. This scale classified the patients with 78% accuracy. In an attempt to cross-validate this actuarial scale, Pruesse and Quinsey (1977) followed-up 206 patients released from a maximum security hospital. Accuracy of prediction dropped from 78% to 65%.

Harris, Rice, Quinsey and colleagues have been instrumental in the development of statistical techniques aimed at the prediction of violent behavior. As the actuarial method's most vocal proponents (Quinsey et al., 1998), they have developed the most popular actuarial instrument used to predict violent behavior, the Violence Risk Appraisal Guide (VRAG; Harris, Rice, & Quinsey, 1993; © American Psychological Association, Washington, DC.).

#### *Construction of the VRAG*

The purpose of the VRAG is to predict future incidents of violence in individuals who have violence in their histories. The VRAG was constructed based on samples of mentally disordered offenders who had been previously followed by the researchers (Harris et al., 1993). The combined validation sample consisted of 685 men of whom 618 had an opportunity to recidivate. The sample was derived from previous studies of Oak Ridge psychiatric patients (Rice, Harris, Lang, & Bell, 1990; Rice, Harris, & Cormier, 1992) and included violent and sexually violent offenders. Offenders in the Rice et al. (1990) study consisted of insanity acquittees and matched comparisons who had spent at least one day in Oak Ridge between the years 1975 and 1981. The offenders in the Rice et al. (1992) study included patients who spent at

least 2 years in the therapeutic community program during the period between 1968 and 1978. While acknowledging that the construction sample consisted of individuals who were mental patients, the authors ensured “that the instrument would work as well with mentally disordered offenders as with offenders free of serious mental disorder” (Quinsey et al., 1998, p. 145).

For the purposes of scale development, the authors defined the outcome variable as any new criminal charge for a violent offense or incident that could have resulted in a criminal charge if committed in the community. Violent offenses that were included ranged in severity from assault to homicide<sup>1</sup>. Official records and institutional files were examined for incidents of violent behavior that did or would have resulted in criminal charges if committed in the community. The occurrence of violent offenses was then dichotomized for purposes of VRAG validation; the resultant outcome variable was at least one violent offense.

The authors collected approximately 50 predictor variables on each of the individuals. The variables were selected on the basis of their empirical relationship to violence or the authors’ curiosity about their relationship. These variables included sociodemographic variables, childhood and adult adjustment variables, characteristics of the index offense, and psychological assessment variables, including IQ. Using discriminant function analysis, twelve variables were chosen based on their independent contribution to predicting violence (Harris et al., 1993). Each item is scored according to a weighting procedure based on how different the individual is from the base rate for that item in the validation sample. Therefore, each item has its own range of scores, with total scores ranging from -26 to +38.

---

<sup>1</sup> Interestingly, the authors noted that in their jurisdiction (assumed to mean Canada), assault can be charged in the absence of bodily contact. Because such assaults are apparently rare, they included all assaults as violent incidents (Quinsey et al., 1998).

Harris et al. (1993) reported a correlation between total VRAG scores and violent recidivism of .44 ( $p < .01$ ). The AUC in the construction sample was .76, indicating that randomly drawing a violent and nonviolent patient will result in the violent patient achieving the higher VRAG score 76% of the time. Harris et al. reported no optimal cut score in applying the VRAG, but instead assigned probabilities of risk to each VRAG category. Quinsey et al. (1998) also investigated whether individuals scoring in the highest category were also those individuals likely to commit the most violent offenses. The seriousness of each offender's act was weakly related to VRAG score ( $r = .18, p < .01$ ).

In cross-validating the VRAG with a sample of sex offenders, Rice and Harris (1997) utilized a sample of 159 sex offenders who were not part of the initial construction sample. Follow-up data included incidents of violence as well as sexual violence. With sex offenders, the VRAG predicts general violence as well as it did in the construction sample. Total scores on the VRAG correlated .47 (compared with .44 in the construction sample) with violent recidivism. The relationship between VRAG scores and sexual recidivism was weak ( $r = .20$ ). The use of total scores does not allow for inspection of the VRAG's performance at various cut points.

### *Specialized Actuarial Instruments*

Beyond measures of general violence, a host of other actuarial instruments have been constructed to predict specific violent behaviors. These instruments tend to be focused on predicting specific types of violent behavior (e.g., sex offenses) or situation-dependent behavior (e.g., inpatient violence) and consist of variables that are easy and inexpensive to collect (e.g., demographic information). The following section outlines a few of the more commonly used instruments and highlights their utility in the process of risk assessment.

*Sex offense prediction.* Predicting recidivism among sex offenders has become a growth industry, particularly since the enactment of sexually violent predator laws (discussed in detail in the section titled *Sexually Violent Predator Statutes*). Efforts to develop techniques to accurately predict the future behavior of this subgroup of offenders has been confounded by statistical as well as conceptual difficulties. Statistically, sex offenses are sometimes difficult to detect because of the low reporting rates of sexually motivated crimes and also due to the tendency of some jurisdictions to enter in plea bargains with defendants for lesser, nonsexual crimes. Conceptually, investigators must decide upon the behavior of interest. The methods outlined in this section were developed to predict sexual recidivism exclusively, with less attention given to general violence.

Hanson developed the Rapid Risk Assessment of Sexual Offense Recidivism (RRASOR; Hanson, 1997) solely to predict sexual reoffending. Using data later published in a meta-analysis of over 20,000 sex offenders (Hanson & Bussiere, 1998), Hanson (1997) identified four variables with a minimum correlation of .10 with sexual recidivism. He developed a brief actuarial scale consisting of: (a) prior sex offenses, (b) any unrelated victims, (c) any male victims and (d) age < 25 at the time of the risk assessment. Across seven developmental samples, comprising a total of 2,592 sex offenders, Hanson found that RRASOR scores had an average correlation of .27 with sexual recidivism. In a large cross-validation sample of 1,400 sex offenders in Sweden, Sjostedt and Langstrom (2000) found that the RRASOR had a correlation of .22 with sexual recidivism. Ideally, correlations with sexual recidivism for the RRASOR would be greater, however, given that the RRASOR consists of four easily scored, easily accessible variables, it may be a useful tool as part of a larger risk assessment.

The Static-99 (Hanson & Thornton, 1999) augmented the RRASOR with items from the Structured Anchored Clinical Judgement Scale (SACJ; Grubin, 1998). Hanson (1999) noted that the RRASOR and the Structured Anchored Clinical Judgment scale were assessing related, but not identical constructs. The SACJ measures domains of risk (e.g., previous convictions) and aggravating risk factors (e.g., deviant sexual interests and psychopathy). Hanson combined the scales to test whether this would lead to better prediction ability than either scale by itself. The resulting scale was called the Static-99 because all the variables were static in nature<sup>2</sup> and the most recent version was developed in 1999.

Hanson and Thornton (1999) tested the Static-99 against the RRASOR and the SACJ in four samples of offenders. For the prediction of sex offense recidivism, the Static-99 (average AUC = .71) was marginally more accurate than the RRASOR (average AUC = .68,  $p < .05$ ) or the SACJ (average AUC = .67,  $p < .01$ ). Although beyond its original purpose, the Static-99 (average AUC = .69) also outperformed the RRASOR (average AUC = .64) and the SACJ (average AUC = .64,  $ps < .001$ ) in predicting any violent recidivism.

Barbaree, Seto, Langton, and Peacock (2001) tested the relative effectiveness of several actuarial techniques in predicting violent behavior. These authors employed the use of the VRAG, RRASOR, and Static-99 to predict violence and sexual violence among a sample of 215 sex offenders released from prison. Additionally, the authors tested the Sex Offense Risk Appraisal Guide (SORAG; Rice & Harris, 1997), an extension of the VRAG. Results indicated that the VRAG and the SORAG performed equally well in predicting any reoffense (AUCs = .77 and .76, respectively, see Table 1). However, the RRASOR outperformed the other actuarial

---

<sup>2</sup> It is arguable whether all the variables are static. The Static -99 includes age < 25 at the time of the risk assessment. Naturally, this variable will change over time.

instruments in the prediction of sexual reoffending. This finding is particularly impressive considering the RRASOR consists of only four items (Hanson & Thornton, 1999). However, the inconsistency of the RRASOR across samples is troubling. Hanson and Thornton (1999) and Barbaree et al. (2001) reported virtually identical AUCs for the Static-99 in predicting sexual recidivism (.71 versus .70). In contrast, the RRASOR showed much more variability across samples, with an average AUC of only .68 in Hanson and Thornton and a large .77 in Barbaree et al. These findings suggest the Static-99 may be the more consistent measure in predicting sexual violence.

Table 1

*Areas Under the Curve for Actuarial Risk Assessment Instruments in Predicting Serious Offenses and Sexual Offenses (from Barbaree et al., 2001)*

Outcome	VRAG	SORAG	RRASOR	Static-99
Any reoffense	.77	.76	.60	.71
Serious reoffense	.69	.73	.65	.70
Sexual reoffense	.61	.70	.77	.70

*Summary and Critique of Actuarial Research*

This brief review highlights the advances in actuarial predictions of violence. Actuarial methods have been shown to have better than chance predictive accuracy for any type of recidivistic behavior. When a clinician is asked to estimate the likelihood that a particular individual will commit any type of offense at some point in the future, actuarial methods such as the VRAG are likely to be helpful. Quinsey et al. (1998) concluded that the “VRAG performed exactly as one would have wished” (p. 150). Depending upon your view, this conclusion could

be considered a marked overstatement. High scores suggest an increased probability of violence; they do not address the level of violence. Predicted violence can range from assault without bodily contact to murder. It is unknown how helpful VRAG predictions are to the clinician faced with making decisions regarding commitment to or release from a secure hospital. Litwack (2001) cogently argued this point in declaring that even a 100% chance of committing a simple assault does not necessarily justify secure commitment.

In addition, the generalizability of the VRAG to other forensic settings is questionable. During the 1960s and 1970s, Oak Ridge was operating its Social Therapy Unit which has been the subject of much debate. By today's standards, the therapeutic practices employed on the unit would be considered unethical. For example, published reports from Oak Ridge detail the use of LSD (Barker & Buck, 1977) and a contraption called the "Total Encounter Capsule" (Barker & McLaughlin, 1977). Some of the patients included in the VRAG construction and cross validation samples were among the patients treated with these and other unacceptable practices (see also Barker & Mason, 1968). Given the unusual nature of the validation sample, it was initially questioned whether the authors' early findings would generalize to other samples, even other mentally ill samples. More recent research with the VRAG (see e.g., Barbaree et al., 2001) has found results comparable to initial validation samples, suggesting this may not be as great of a concern as it once was.

An additional consideration with the VRAG and all actuarial instruments is their sole focus on violence prediction with no consideration of treatment or management. Proponents of comprehensive risk assessment (Boer, Hart, Kropp, & Webster, 1997; Hanson & Harris, 2000; Hart, 1998; Webster, Douglas, Eaves, & Hart, 1997) argue that dynamic variables must be included to facilitate treatment and provide for changes (positive and negative) over time.



Actuarial instruments focus on static variables that, by definition, do not change. The exclusive use of static variables is inconsistent with comprehensive risk management. As such, actuarial instruments can only address the first prong of risk assessment: predicting likelihood of future violent behavior. While problematic in certain contexts, their narrow focus is not necessarily a limitation when the sole purpose of the risk assessment is the prediction of future violence.

Supporters of the actuarial approach to risk assessment vary in their enthusiasm. Webster, Harris, Rice, Cormier and Quinsey (1994) advocated that actuarial techniques be used as the principal, but not exclusive, method of rendering clinical decisions. They recommended that clinicians may want to adjust actuarial estimates of risk within a narrow range when there were compelling circumstances to do so. These authors have recently modified their recommendation and currently are advocating for the exclusive use of actuarial methods. They justify this position stating that “actuarial methods are too good and clinical judgment too poor to risk contaminating the former with the latter” (Quinsey, et al., 1998, p. 171). In contrast, Hanson (2000; see also Hanson & Harris, 2000) argued for a more balanced approach suggesting that evaluators use both clinical data (e.g., dynamic variables) and actuarial findings.

#### Clinical Judgment

Monahan (1981), as previously cited, concluded that clinical predictions of violence are correct only one out of three times. The meaningfulness of this estimate has been questioned. Mossman (1994) argued that traditional analyses such as contingency tables are too reliant upon base rates of behavior and clinicians’ preferred errors (e.g., tendency to maximize positive or negative predictive power). Using detailed clinical examples, Mossman demonstrated how accuracy can be obscured by base rates and preferred errors. In a hypothetical example, he demonstrated that one clinician achieved accuracy 54% of the time, but achieved sensitivity of

.93, meaning she missed only 7% of violent patients. A second clinician achieved overall accuracy of 89%, but missed 50% of the violent patients. Mossman demonstrated that the ratio between true positive and false positives, one indicator of predictive accuracy, was identical for both clinicians (1:2) as was the AUC (.856). Rice and Harris (1995) reinforced this point by explaining that true positive, true negatives, sensitivity and specificity will vary at different decision points (i.e., cut scores or thresholds). These estimates are necessarily confounded by the conservatism of clinicians' judgments. The authors argue that the constancy of the AUC is its strength.

Mossman (2000) reanalyzed data from 44 studies of violence prediction, including the data that Monahan (1981) used to make his conclusion. He found that estimates of accuracy using ROC analysis were quite higher than those originally reported. Mossman reported that the Areas Under the Curve (AUCs) for these 44 studies ranged from .48 to .98, with an average AUC of .78, well within, if not above, the range of AUCs often found using actuarial techniques. For example, Steadman and Cocozza's (1974) very influential original findings were reanalyzed and yielded an AUC of .76. Mossman's findings suggest that clinical predictions of violence may not be as poor as once assumed.

Other investigators have questioned whether accuracy is even the appropriate question. Gardner, Lidz, Mulvey and Shaw (1996; see also Mossman, 2000; Mulvey & Lidz, 1995) have framed the question as one of conditional probabilities of violence. They argue that clinical judgment may have increased accuracy when specified in terms of conditions. A goal is to discover *which* clinicians under *what* circumstances can predict certain *types* of behavior (e.g., Gardner, Lidz, Mulvey & Shaw, 1996; Skeem, Mulvey, & Lidz, 2000).

### *Emergency Commitment*

Monahan (1981) was the first to consider differences in accuracy of clinical predictions of violence between institutionalized patients and psychiatric emergency room (ER) patients. In psychiatric emergencies, the potential for violent behavior is usually temporally and situationally close to the prediction. The threat of violence likely happened directly before presentation to the ER. Monahan's writing stresses the importance of situational variables in clinical prediction.

In the first violence prediction study to utilize psychiatric ER patients, Lidz, Mulvey, and Gardner (1993) tested clinical predictions of violence in two samples of psychiatric patients from a large urban psychiatric emergency department. After a 6-month follow-up period, the overall base rate of violence was 44.6%. Clinicians' positive predictive power was significantly better than chance, with 53.2% of the patients predicted to be violent later becoming violent. Of those predicted to be nonviolent, more than one-third (36%) became violent (Negative Predictive Power = .64). Overall predictive accuracy (i.e., hit rate) was 58.5%. Mossman reanalyzed Lidz et al.'s data and found an AUC of .66.

### *Inpatient Violence*

Inpatient facilities may provide a particularly appropriate setting for studying risk assessment partly because of their high base-rate of violent behavior. According to Otto (1992), approximately 15-28% of hospitalized individuals engage in some type of physically assaultive behavior, and as many as 40-50% engage in some type of dangerous behavior, including threats and other physical acts indicative of violence.

A particularly positive aspect of studying risk assessment in inpatient settings is the typically short time between prediction and outcome. The short follow-up time allows research to be completed in a timely fashion. A potentially negative aspect of this population (from a

research perspective) is that the provided treatment may act as a mediating variable which prevents violent behavior from occurring. The research programs of McNiel and Binder (1991) and Werner and colleagues (1983, 1984) with inpatient samples have highlighted common pitfalls of violence prediction that affect predictive accuracy.

McNiel and Binder (1991) asked clinicians to estimate the probability that newly admitted psychiatric patients would become violent within the first week of hospitalization. Twenty-two nurses and 41 physicians provided estimates on the likely dangerousness of 149 patients. Both nurses and physicians rated patients' violence in proportion to their actual violent acts. However, both disciplines tended to overpredict violence, placing more patients in the moderate and high violence categories than actually committed violent acts. In reanalyzing their results to calculate traditional utility estimates, I dichotomized the physicians' predictions into binary categories reclassifying low risk as a nonviolent prediction. Moderate and high estimates were reclassified as violent predictions. The results suggest physicians achieved a true positive rate of 31.5% and a true negative rate of 87.4%. Sensitivity and specificity were 85.7% and 78.9%, respectively. Physicians overpredicted violence, but achieved an overall hit rate of 73.4%. Using Mossman's (1994) FP:TP ratio, the physicians achieved a 2:1 ratio. This is analogous to Monahan's (1981) assertion that predictions are correct only one out of three times. However, Mossman reanalyzed this data using ROC analysis and showed an AUC of .71, an AUC significantly better than chance. This illustration demonstrates that "accuracy" is at least partially dependent upon statistics, with traditional utility estimates possibly underestimating predictive accuracy.

Recognizing the need for more structure in violence risk assessments, Werner, Rose, Yesavage and Seeman (1984) investigated the variables that clinicians used in order to make

predictions of violence. Fifteen psychiatrists predicted potential violence for 40 male psychiatric inpatients using the following data: (a) Brief Psychiatric Rating Scale (BPRS; Overall & Klett, 1972) scale scores and (b) physical assaults precipitating hospitalization. The correlation between actual violence and predicted violence was nonsignificant ( $r = .14$ ). Although psychiatrists weighed five of the 18 BPRS factors positively in predicting violence, none were significantly related to violence at all. This finding suggests clinicians' misuse of available data may contribute to poor predictive accuracy. This research demonstrates a need to educate clinicians about factors empirically related to violent behavior because predictions based on irrelevant variables (i.e., those unrelated to violence) are likely to be inaccurate.

Quinsey and Ambtman (1979) first recognized poor interrater reliability as a pitfall of clinical predictions of violence. The validity of predictions is necessarily limited by the poor reliability. To address this issue, Werner, Rose and Yesavage (1983) investigated both interrater reliability as well as accuracy of prediction. Using the same 40 patients in Werner et al. (1984), they asked an additional 15 psychologists to rate these patients. Interrater reliability was only moderate for psychologists ( $ICC = .47, p < .001$ ) and modest for psychiatrists ( $ICC = .37, p < .001$ ). In terms of prediction accuracy, the mean hit rate for all judges was .39, with only two out of every five violent patients being correctly identified. Werner et al.'s results suggest several hypotheses regarding clinicians' accuracy. As previously noted, the low interrater reliability necessarily constrained validity. In addition, the use of the BPRS may have limited value in making predictions. Clearly, this information was not helpful to clinicians making judgments. In fact, other researchers (e.g., Lidz et al., 1993, McNeil & Binder, 1991) demonstrated that clinicians using completely unstructured judgment can achieve better predictive accuracy than clinicians using irrelevant information.

### *Summary and Critique of Clinical Judgment*

Recent research utilizing clinical predictions of violence has been more positive than Monahan's (1981) initial review. Clinicians are often able to achieve predictive accuracy at better than chance levels, even when these predictions are essentially unstructured (Lidz et al., 1993; McNeil & Binder, 1991). Advances in the literature have also pointed to the common pitfalls of clinical predictions and suggested avenues to increase predictive accuracy. For example, the work of Werner and colleagues (1983, 1984) suggest that clinical variables may be important, but only when the correct relationship between these clinical variables and violent behavior is known. Like Quinsey and Ambtman (1979), Werner et al. (1983) recognized the importance of interrater reliability in maximizing clinicians' predictive accuracy. Observations from the available research suggest the need for a structured method that allows clinicians to (a) achieve high levels of interrater reliability, (b) weigh variables accordingly to risk, and (c) use their clinical skills in assessing those variables. Advances in the clinical prediction of violent behavior, through the use of structured clinical judgments, directly address these goals.

#### Structured Clinical Judgment

The term "clinical judgment" in the risk assessment literature refers most often to *unstructured* clinical judgment. In reviewing the process of clinical predictions of risk, it becomes apparent that this lack of structure hinders the accuracy of judgments. Numerous authors (Monahan, 1981; see also Hart, 1998; Monahan & Steadman, 1994; Webster, Harris, Rice, Quinsey & Cormier, 1994) have criticized the use of the unstructured clinical judgment approach to risk assessment. More recently, structured approaches to risk assessments have been proposed to overcome the difficulties encountered with unstructured methods.

Hart (1998) argued for the adoption of structured (i.e., guided) clinical judgment for risk assessment. Structured approaches to risk assessment improve over unstructured approaches by standardizing how (a) evaluations are conducted and (b) variables are weighed. Structured clinical approaches also differ from actuarial methods in several important ways. Unlike actuarial approaches, the flexibility of clinical approaches allow for decisions to be rendered considering the “totality of the circumstances,” not simply a set of static variables. The flexibility of structured approaches allow professionals, in certain circumstances, to decide that the test results are not meaningful in a given case and should be ignored. As Hart pointedly asks, “Does it matter at all what an offender’s total score is on the VRAG, and how many risk factors are present or whether he scores above a specific cut-off, if he also expresses genuine homicidal intent?” (p. 126).

#### *Advances in Structured Clinical Judgment*

Structured clinical approaches to risk assessment have two main objectives. First, structured approaches strive to provide consistency. Secondly, they balance consistency with flexibility. The importance of these considerations led to the development of structured clinical guides to risk assessment. Currently, the Historical-Clinical-Risk 20 (HCR-20; Webster, Douglas, Eaves, & Hart, 1997; © Mental Health Policy and Law Institute, Burnaby, Canada) and the Sexual Violence Risk –20 (SVR-20; Boer, Hart, Kropp & Webster, 1997; © Mental Health Policy and Law Institute, Burnaby, Canada) comprise the available guides for structured risk assessments. Each of these measures is described below.

The HCR-20 identifies 20 variables that are organized into three scales: *Historical*, *Clinical*, and *Risk management*. The HCR-20 manual (Webster et al., 1997) provides definitions and scoring guidelines for each of the variables. In addition to static variables, the HCR-20 also

considers situational and environmental variables, such as exposure to destabilizers in the environment, lack of personal support, and stress. Webster et al. underscore this final aspect of the HCR-20 as particularly important for allowing flexibility in assessments. Identifying dynamic variables suggests targets for intervention, thus making structured approaches more consistent with the second prong of risk assessment, namely risk management. Borum (1996) wrote of the HCR-20 that “The promise of this instrument lies in its foundation on a conceptual model or scheme for assessing dangerousness; its basis in the empirical literature; [and] its operationally defined coding system” (p. 950).

Monahan (1981) identified four common “blind spots” in the clinical prediction of violent behavior, namely, (a) the lack of specificity in defining the criterion; (b) a reliance on illusory correlations; (c) a failure to incorporate situational or environmental information; and (d) the neglect of base rates. As described above, the HCR-20 addresses three of Monahan’s chief concerns. First, it specifically defines the items and provides scoring criteria. Second, it uses clinical and risk variables that are empirically related to violence. Third, its ratings take into consideration situational and environmental variables. In summary, the HCR-20 has a sound conceptual basis that addresses these relevant domains. The next subsection summarizes the available research on its validation and predictive accuracy.

*Research Findings with the HCR-20.* Several researchers (Douglas, Ogloff, Nicholls, & Grant, 1999; Douglas & Webster, 1999; Grann, Belfrage, & Tengstrom; 2000; Kroner & Mills, 2001) have examined the HCR-20’s validity in correctional, forensic psychiatric, and civil psychiatric settings.

Douglas and Webster (1999) investigated the validity of the H (Historical) and C (Clinical) scales as well as combined scores in a sample of Canadian offenders through



comparisons to past violence. The authors noted that the R (Risk) scale was not completed due to insufficient information for coding. Utilizing chart reviews, they demonstrated that the H and C scales were significantly ( $p < .01$ ) related to past violent offenses ( $r$ s .30 to .50). Also utilizing a sample of Canadian offenders, Kroner and Mills (2001) followed 87 offenders an average of 790 days post-release. They reported the HCR-20 correlated modestly ( $r = .16$ ) with reconvictions for violent offenders and achieved a modestly greater than chance AUC of .62.

Several studies have investigated the HCR-20 with mentally disordered offenders. Investigating the Historical scale's predictive validity in a sample of mentally disordered offenders in Sweden, Grann et al. (2000) followed released patients for two years. Using ROC analysis, the AUC in predicting violent reconviction was .71. Additionally, they calculated the scale's sensitivity and specificity at a cut score of 12 (.71 and .61, respectively). Finally, Douglas et al. (1999) evaluated predictive validity in civil psychiatric patients. In this sample of 193 patients, the HCR-20 and its subscales' AUCs ranged from .63 to .76.

In summary, the above findings suggest that the HCR-20 has demonstrated predictive and postdictive accuracy that parallels that found with actuarial schemes, such as the VRAG and the Static-99, and does so in both correctional and civil samples. Demonstrating its generalizability to civil settings may provide extra support for the HCR-20. The VRAG and the Static-99 were developed on and validated with criminal samples. Therefore, a benefit of the HCR-20 may be its ability to predict violence in criminal and noncriminal settings. Preliminary research suggests strengths of the HCR-20 include reliability and generalizability.

*The Sexual Violence Risk-20 (SVR-20).* The Sexual Violence Recidivism-20 (SVR-20; Boer et al., 1997), a version of the HCR-20, was developed specifically for use with sex offenders. In constructing the SVR-20, the authors attempted to identify risk factors that are

empirically related to future sexual violence, clinically useful, yet parsimonious. For the purposes of the SVR-20, the definition of sexual violence is “actual, attempted, or threatened sexual contact with a person who is nonconsenting or unable to give consent” (Boer et al., 1997, p. 328). This broad definition of sexual violence includes rape, sexual touching, exhibitionism, obscene letters or phone calls, distribution of pornography, voyeurism, and theft of fetish objects. Clinicians are likely to be divided regarding whether this wide array of acts should be defined as sexual violence. This overly broad definition of sexual violence may have implications for its predictive validity, particularly the number of false positives, especially as it relates to its usefulness in making determinations about sexually violent predator commitments.

The items chosen for inclusion for the SVR-20 were derived through a review of the literature on sex offenders. Special attention was given to factors that discriminate between sexual and nonsexual offenders and on those associated with recidivistic violence or sexual violence in sex offenders. However, the SVR-20 remains to be tested empirically.

Many of the items included in the first and third sections of the SVR-20 are similar to instruments that assess for risk of general violence, such as the VRAG. The second section, Sexual Offending, includes items believed to be associated specifically with risk for sexual violence (Boer et al., 1997).

### *Summary and Critique of Structured Clinical Judgment*

The authors of the HCR-20 and the SVR-20 avoid offering a decision-making algorithm with associated cut scores and categories of risk. Instead, they invite clinicians to use these items as a guideline to help them assess probabilities of risk and make recommendations regarding treatment and management of sex offenders. They provide no guidelines to achieve this. The absence of cut scores increases the clinician’s flexibility and allows him or her to consider other

important information in decisions regarding risk. An inherent limitation to this approach is that clinicians are left without guidance at the most crucial step of the assessment process: making an ultimate determination.

### Applying Risk Assessment Techniques to Clinical Practice

Forensic clinicians have been legitimized in their risk assessment practice by the courts (see e.g., Janus, 2000; Monahan et al., 2001). Despite the fact that research has shown that no risk assessment method is highly accurate, courts have continued to place this responsibility on clinicians, arguing that there is nothing “inherently unattainable about a prediction of future criminal conduct” (*Schall v. Martin*, 1984, p. 2417). One possible conclusion is that in the absence of a proven method, selection of a risk assessment method may depend on the clinician’s personal preferences. Another possible conclusion maintains that no method has demonstrated superiority because all studies of accuracy have obscured the prediction task by asking clinicians to make overarching predictions regarding risk that are independent of situation and context (see e.g., Heilbrun, 1997; Mulvey & Lidz, 1998; Skeem, Mulvey, & Lidz, 2000).

### *Contextualized Risk Assessment*

Twenty years ago, Monahan (1981) stated that what was needed for moral, legal, and empirical “progress in the area of prediction is a dramatic increase in the degree to which mental health professionals articulate what it is they are predicting and how they went about predicting it” (p. 17). Despite this plea, prediction research has continued to focus on the accuracy of clinicians’ judgments and not on process of decision-making itself.

Most research in the area of violence prediction has been grounded in the cue-utilization model of human judgment. Grisso (1991) explained that this model frames the task of predicting dangerousness as a “clinical exercise” in applying a context-free algorithm for combining risk

factors. This context-free model of risk assessment does not parallel real-life decision making in many situations. Clinicians are likely to be asked “under what conditions is this person likely to be violent?” Importantly, the legally relevant question is: “Is this person’s risk of future violence severe enough to warrant detaining him or her involuntarily?” This kind of conditional risk assessment more closely resembles the type of questions and issues that clinicians actually face. It is naïve to suggest that (a) predictions of violence occur in a context-free framework, and (b) such predictions are accurate regardless of environmental or situational events. Similarly, Skeem, Mulvey, and Lidz (2000) criticized the cue-utilization model because it assumes that behavior is largely independent of context. As they noted, “a substantial body of research on mental health professionals’ (MHPs’) predictive accuracy appears to be based on a model of clinical decision making that inadequately represents the actual nature and goals of MHPs risk assessments” (p. 609).

Contextual variables that have the potential to influence the likelihood of risk include (a) situational or environmental variables and (b) legal decision-making context. Situational or environmental variables include a myriad of issues, such as inpatient or community violence, substance use, domestic disputes, and medication compliance. Legally relevant contextual variables are often intertwined with situational variables. Examples include legal determinations, such as release from a maximum security hospital or certification as a sexually violent predator.

*Situational/Environmental Context.* Past empirical evidence has suggested that clinicians tend to overpredict violence in their patients (Lidz et al., 1993; McNiel & Binder, 1991; Monahan, 1981; Steadman & Cocozza, 1974; see also previous sections titled *History of Violence Predictions* and *Clinical Judgment*). Skeem et al. (2000) proposed that the dichotomous nature of their decisions (dangerous/not dangerous) may hide the more subtle context-specific

decisions being rendered. For instance, a clinician may have predicted that a male patient will be violent toward his mother due to his alcohol intake, noncompliance with medication, and residence with his mother. However, if the patient subsequently complies with his medication and abstains from alcohol, his risk of violence is presumably low. These contextual factors are often not made explicit when factored into binary decisions. As observed in this example, the clinician appears to have overpredicted violence because the salient situational variables are not explicated.

A fundamental problem with the notion of actuarial risk assessment techniques is that they do not take into account these contextual variables. Their use may be incompatible with the flexibility needed to incorporate environmental and situational variables into a prediction. In contrast, their use may maximize predictive ability in certain contexts. More research is needed into the contextual variables that influence decision making and the process clinicians use in formulating their judgments.

*Legal Decision Making Context.* Environmental/situational contexts are important for clinicians to consider in making responsible risk assessments. Heilbrun's (1997) review illustrates the importance of legal context in the decision making process. Heilbrun puts forth two different models (i.e., prediction and management) for decision making in risk assessments. The goal of the prediction model is to determine the occurrence or nonoccurrence of some event in a specified time period. In contrast, the primary goal of the management model is to reduce the risk of this event's occurrence. Different goals of the risk assessment process may dictate the need for different methods of risk assessment. For example, actuarial techniques may have the most utility when binary decisions are needed, such as a decision to commit an individual under a sexually violent predator act. On the other hand, a clinically guided risk assessment may be

more useful during decisions regarding graduated release. The use of these clinical variables allows for the identification, and hence, changes in, dynamic variables amenable to treatment.

Heilbrun's (1997) conceptualization of prediction versus management models provides a framework for shifting the way we think about risk assessment and the decisions resulting from that process. The enactment of sexually violent predator laws affords a rich context in which to discuss this issue.

### Sexually Violent Predator Statutes

Since 1990, the United States has seen a resurgence in the enactment of Sexually Violent Predator (SVP) laws. Similar to sexual psychopath laws enacted in the 1930s and 1940s, these laws allow for the indefinite civil commitment of sexually violent individuals.<sup>3</sup> The early sexual psychopath laws were enacted to help society deal with sexual offenders who were "too sick to deserve punishment" (Janus, 2000). In contrast to these original laws, this "second generation" of commitment laws were enacted as a public safety measure in extending the incapacitation of offenders who had already served their criminal sentences. The state of Washington was the first to pass a second-generation commitment law, followed by 16 additional states (Lieb & Matson, 1998). Although these laws have come under the scrutiny of the scientific and legal (see e.g., Janus, 2000; Schopp, 1998) communities, SVP laws continue to enjoy constitutional and popular support.

Washington's Community Protection Act (Revised Code of Washington, 1990) provided for the civil commitment of individuals who were deemed "sexually violent predators."

According to this statute, a sexually violent predator is "any person who has been convicted of,

---

<sup>3</sup> Minnesota actually has the longest-standing sex offender commitment statute. Its 1939 law allowing for the commitment of "sexual psychopaths" was never repealed. Instead a 1994, "sexually dangerous persons" statute was enacted that is more in line with modern commitment statutes.

or charged with, at least one crime of sexual violence and suffers from a mental abnormality or personality disorder that makes the person likely to engage in future predatory acts of sexual violence if not confined in a secure facility” (Lieb & Matson, 1998, p. 27). Although state-to-state variation exists in the exact language of these laws, each shares four common elements: (a) a past act of sexually harmful conduct, (b) a current mental disorder or abnormality, (c) a finding of risk of future sexually harmful conduct, and (d) some connection between the mental abnormality and the danger. Most states require that the person be convicted or charged with a sexually violent offense; other states (see e.g., North Dakota House Bill 1047, 1997) simply require that the individual “be shown” to have engaged in sexually predatory conduct.

Kansas has the most well-known of the second-generation commitment laws, due to the Supreme Court’s (1997) *Hendricks* decision. In *Hendricks*, the Court upheld the state’s police power rights and legitimized the constitutionality of SVP commitment laws. Briefly, *Hendricks* was a convicted child molester serving a sentence in a Kansas state penal institution. Shortly before his release in 1994, the state petitioned to have *Hendricks* civilly committed under a newly-enacted sexually violent predator law (see Kan. Stat. Ann. 59-29[a], 1994). *Hendricks* argued against the constitutionality of this law and was eventually denied relief by the US Supreme Court. The Kansas statute allows for the commitment of “any person who has been convicted of or charged with a sexually violent offense and who suffers from a mental abnormality which predisposes the person to commit sexually violent offenses in a degree constituting such person a menace to the health and safety of others” (Lieb & Matson, 1998, p. 21).

Second-generation commitment statutes require that the individual undergo an evaluation. The nature of these evaluations is not well articulated but each calls for the

“certification of the person as a sexually violent predator” (e.g., New Jersey, Laws of 1998) or determination that the “person will engage in sexually violent behavior” (e.g., California Welfare and Justice Code, 1996). Other states (i.e., Wisconsin Statutes, 1994) simply call for an “evaluation.” While these statutes provide for the final determination to be made by the court, the role of the mental health professional is integral. These legal determinations are informed by mental health professionals’ findings. Without clear parameters regarding their responsibilities, mental health professionals operating within the legal arena face challenges in answering the psycholegal issues posed to them.

#### *Sex Offender Recidivism as a “Special” Case of Risk Assessment*

Within the context of risk assessment, SVP laws present a unique challenge to mental health professionals conducting evaluations. Importantly, individuals who commit sexual offenses do not represent a specialized criminal class. Sex offenders frequently commit nonsexual crimes (Hanson & Bussiere, 1998). Hanson (1999) noted that antisocial sex offenders are more likely to recidivate with nonsexual crimes than with sexual crimes. In addition, not all sex offenders have deviant sexual interests or preferences. Hanson and Bussiere reported that 10-25% of male community samples admit to some form of sexual offending. It is unlikely that all of these individuals warrant the label “sexually violent predator.” The challenge facing clinicians is to determine which of these individuals with sexual violence in their histories is likely to both (a) commit another act of sexual violence (as opposed to another criminal act) and (b) commit acts of sexual violence resulting from a mental abnormality as required by sexually violent predator laws.

Hanson and Bussiere’s (1998) meta-analysis of sex offenders identified factors predicting general recidivism different from those predicting sexual recidivism. Specifically, sex offenders



who recidivated with nonsexual crimes tended to be young, unmarried, and have a history of juvenile and adult antisocial behavior. These factors are also risk factors in samples of non sex offenders. In contrast, the strongest predictors of sexual recidivism were factors related to sexual deviance, such as deviant sexual interest, prior sexual offenses, and victim choices such as boys and strangers. Thus, it is the combination of sexual deviance and an antisocial lifestyle that places offenders at high risk for sexual offense recidivism (Hanson, 1999). The authors concluded that the assessment of sexual recidivism needs to consider factors uniquely related to sexual offending, such as sexual deviance and victim type.

#### Victim Impact Statements

A final consideration in the review of risk assessment literature concerns potentially biasing information. Much empirical attention has been given to those variables that improve clinicians' abilities to predict risk. Equally important to consider is information that may decrease or bias those judgments. The development of actuarial methods and structured clinical guides for aiding in the process of risk assessment highlights the need to consider relevant variables when making assessments regarding risk. Ideally, the use of well-validated instruments will help decision-makers avoid reliance upon emotionally-laden, yet irrelevant information.

Victim impact statements have been considered in the past as a potentially biasing force in attorneys' as well as psychiatrists' (Lynett & Rogers, 2000) decisions regarding dangerousness. Forensic examiners are often encouraged to gather multiple sources of information when conducting risk assessments (Boer, et al., 1997; Rogers & Shuman, 2000; Webster et al., 1997). In the first systematic study to investigate the effects of victim impact statements on decision-makers, Lynett and Rogers (2000) found that the presence of an emotionally evocative victim statement exerted a moderate effect ( $d = .52$ ) on psychiatrists'

determinations of dangerousness. The authors noted a limitation of their study may be the use of psychiatrists who did not necessarily have specialized forensic training or practice.

### Current Study and Research Questions

Forensic psychologists who become involved in risk evaluations undertake an enormous responsibility that has social and legal implications that are far from trivial. Monahan (1981) recognized

The consequences of erroneous predictions of violence include the injury or death of the victim of the person wrongly predicted to be safe and the extended institutionalization in a prison or mental hospital of the person wrongly predicted to be violent, or even, as we have noted, his or her execution. While the prediction of violent behavior shares many features with the prediction of other forms of human conduct, the potential consequences of its misapplication give it a priority in professional and ethical concern. (p. 14).

The importance of accurate predictions of violence and the legal and ethical responsibility of clinicians to perform them is well established. The most appropriate avenue for achieving this accuracy is still in debate. This review of the literature highlights the continuing need for well-designed research. The present research contributes to the literature by comparing different models of risk assessment and issues related to clinical decision making.

### *Effectiveness of Techniques in Assisting Decision-Makers*

Borum (1996) noted that predictive accuracy has improved from first generation to second generation of violence research due to advances in research methodology. Predictive accuracy is an important component of responsible risk assessment. As was detailed previously,

both actuarial and clinical judgments have produced better than chance accuracy rates (Barbaree et al., 2001; Hanson & Thornton, 1999; Lidz et al., 1993; Mossman, 1994; Quinsey et al., 1998). These studies examine a clinician or a test's ability to predict a wide variety of violence, ranging from simple assaultive acts to murder. However, general recidivism is rarely required by the referral question. Instead, clinicians are asked to evaluate the likelihood of particular types of recidivism (e.g., sexual violence) to the degree that warrants involuntary detainment. Risk assessments must balance a person's right to be free from unnecessary confinement against the public's right to be protected from dangerous individuals. Litwack (2001) eloquently argued that actuarial methods, such as the VRAG, do not predict the absolute risk an offender poses of committing a sufficiently serious offense to justify that offender's involuntary confinement. He cautioned clinicians that estimates of risk made with these instruments may not justify indefinite commitment by any reasonable cost-benefit analysis.

Despite achieved accuracy of any actuarial or clinical method designed to predict violent or dangerous behavior, the responsibility falls upon the decision maker (i.e., clinician) to assimilate relevant data into an *assessment of risk*. Furthermore, investigators (e.g., Hart, 1998; Sreevinsan et al., 2000) have stressed that tests do not make decisions. Clinicians are ultimately responsible for their own decisions. Reliable and valid tests inform clinicians' judgments, but they do not replace them. Studies to date have not investigated the utility of the different assessment approaches (i.e., actuarial, unstructured clinical judgment, and structured clinical judgment) in helping clinicians make informed, valid, and relevant decisions.

*Research Question #1.* Do clinicians achieve greater levels of predictive accuracy when their decisions are informed by actuarial, clinical, or structured clinical methods?

### *Discriminant Validity*

Referral questions naturally guide any responsible assessment process. The clinician structures his or her assessment (choice of tests or methods as well as presentation of results) around the chief referral question. This basic tenet of assessment applies to the area of risk assessment as well. An all-encompassing statement that either actuarial methods or clinical judgment is better is akin to stating that the WAIS-III is “the best test.” Issues to be addressed include for what purpose, under which conditions, and about whom judgments are being made. These contextual variables are pivotal in assessing any test or technique’s utility.

Quinsey et al. (1998) contend that discriminating between future sexual violence and nonsexual violence is not an important question. In acknowledging the existence of specific sexual predators laws, they recognize that actuarial methods cannot identify all dangerous sex offenders (and discriminate them from nonsex offenders) without an unacceptably high false positive rate. They contend, however, that the real problematic issue is that the laws are missing the important point. Society should be protected from violent offenders of all kinds, not simply sex offenders. Therefore, they argue that the important prediction is that of violent behavior in general. This stance is much more consistent with Canada’s Dangerous Offender laws than American sexually violent predator acts. Quinsey et al. do not appear to recognize this limitation when recommending actuarial predictions. Regarding sexually violent predator commitments, clinicians are asked to distinguish between general violence and sexual violence. Hart (1999) articulated this argument directly, noting that when Quinsey et al. (1998) “advise readers to ignore the law . . . and to disregard the laws of the jurisdiction within which they work,” they are “guilty of counseling professional suicide” (p. 487). Furthermore, Hanson and Bussiere’s (1998)

meta-analysis of sex offender recidivism provided an empirical basis for considering sexual recidivism separately from nonsexual recidivism.

In the present study, Sexually Violent Predator laws will be employed to provide the decision making context needed to structure a relevant risk assessment. Questions inherent in SVP laws include the likelihood that an individual will commit *sexual* violence in the future and that the risk posed is sufficient to justify involuntary detainment of the individual. The various risk assessment methods will be examined to assess their utility for decision makers.

*Research Question #2.* Do assessment methods have differential utility in predicting general versus sexual violence?

#### *Actuarial versus Structured Clinical Predictions of Violence*

Strong support for actuarial techniques has caused Quinsey and his colleagues to modify their previous recommendation that clinicians may want to adjust actuarial estimates of risk when there were compelling circumstances to do so (Harris, Rice, & Quinsey, 1993; Webster, Harris, Rice, Cormier, & Quinsey, 1994). Currently, Quinsey et al. (1998) are advocating for the exclusive use of actuarial approaches on the grounds that clinical judgment contaminates the accuracy of actuarial judgments. Other research with structured clinical judgement (see e.g., Douglas & Webster, 1999; Litwack, 2001; Strand, Belfrage, Fransson, & Levander, 1999) suggests that this complete replacement may be premature. The assertion that clinical judgment contaminates the accuracy of actuarial predictions is an empirical question that remains untested.

*Research Question #3.* Does the addition of clinical information reduce the accuracy of actuarial judgements?

### *Confidence of Judgments*

The relationship between confidence of clinicians' judgments and accuracy has been periodically investigated in the literature with mixed results. Oskamp (1965) found that confidence of raters' judgments regarding personality traits had no effect on their accuracy. On the other hand, McNiel, Sandberg, and Binder (1998) found that confidence of clinicians' predictions of violence was associated with the outcome of violence. Litwack (2001) argues that if confident clinical judgments were found to be more accurate than actuarial predictions, then confident clinical judgments should be relied upon above actuarial assessments. This recommendation seems a bit premature (and difficult to follow practically), yet it highlights the lack of information we have regarding clinicians' confidence. The current study will address the relationship between confidence and relative accuracy in clinicians' judgments as well as the relationship between assessment method and level of confidence.

*Research Question #4.* Do clinicians' confidence levels vary as a function of amount of data and decision making method?

*Research Question #5.* Is there a relationship between confidence and accuracy of judgments?

### *Treatment Amenability*

Hart (1998; see also Boer et al., 1997; Hanson & Thornton, 1999; Webster et al., 1997) suggested that actuarial instruments are constrained by their use of static variables that do not allow for interventions, such as treatment and other remediations. An implied hypothesis is that the use of actuarial methods serve to make clinicians pessimistic about a patient's amenability to treatment.

*Research Question #6.* Do predictions of treatment success vary as a function of assessment method?

#### *Support of Sexually Violent Predator Commitment*

Research Question #7 investigates differences in indicated support for Sexually Violent Predator commitment. Specifically, the research question inspects differences between violent and sexually violent offenders as well as the effects that the varying decision making methods may have on clinicians' recommendations.

*Research Question #7.* What factors influence a clinicians' indicated support for committing an offender as a Sexually Violent Predator?

#### *Professional Training Experience*

Research Question #8 investigates the effects of clinicians' level of training on various predictions. Previous research (Lidz, Mulvey, Apperson, Evanczuk, & Shea, 1992) suggested that more experienced clinicians tend to make more conservative judgments. In the present study, it is expected that experienced forensic psychologists will rate defendants as less likely to engage in both violent and sexually violent behavior in the future. It is also hypothesized that experienced forensic psychologists will indicate less support of a defendant's commitment under sexually violent predators laws.

*Research Question #8.* Do judgments regarding future dangerousness and support of SVP commitments vary as a function of professional experience?

## Ancillary Research Question

### *Biasing Effects*

One potentially biasing source of information is the inclusion of victim impact statements in comprehensive risk assessments. Limited available research (Rogers & Lynett, 2000) suggested that the availability of emotionally provocative victim statements may substantially influence psychiatrists' judgments.

*Ancillary Research Question #1.* Does the inclusion of Victim Impact Statements influence clinicians' decisions regarding potential violence and sexually violent predator determinations?



## CHAPTER 2

### METHOD

#### Design

A 2 x 6 x 3 mixed factorial design was employed to test the various hypotheses.

Independent variables included: outcome type, assessment method, and a repeated measures variable. The design allowed for between-subjects comparisons of assessment methods as well as within-subjects comparisons of incremental validity and/or biasing effects of different kinds of information (see Table 2 for an illustration). An additional strength of the design is the inclusion of the within-subjects factor, rendering the analysis statistically more powerful than reliance solely on between-subjects comparisons. The independent variables are operationalized:

1. The first independent variable, outcome type, consisted of two categories: sexual violence and nonsexual violence.
2. The second independent variable, assessment method, consisted of three primary categories: actuarial, structured clinical judgement (SCJ), and unstructured clinical judgement (UCJ).

Each of these primary methods was then coupled with a second method, totaling six different categories of assessment method: actuarial/SCJ, actuarial/UCJ, SCJ/actuarial, SCJ/UCJ, UCJ/actuarial and UCJ/SCJ.

3. The third independent variable was a repeated-measures factor consisting of successive ratings made by the participants. Time 1 ratings were made following exposure to the primary assessment method. Time 2 ratings were made following exposure to the second method. Time 3 ratings were made following exposure to the victim impact statement.

Table 2

*Schematic Depiction of 6 X 3 Factorial Research Design*

Group	Time 1	Time 2	Time 3
Group 1	Actuarial	Structured	Victim Impact Statement
Group 2	Actuarial	Unstructured	Victim Impact Statement
Group 3	Structured	Actuarial	Victim Impact Statement
Group 4	Structured	Unstructured	Victim Impact Statement
Group 5	Unstructured	Actuarial	Victim Impact Statement
Group 6	Unstructured	Structured	Victim Impact Statement

*Note:* Columns present repeated measures factor, rows depict between-subjects categories.

Dependent variables included both categorical and dimensional judgments including binary predictions of violence (i. e., yes or no), estimates of the likelihood of general violence and sexual violence, confidence of those judgements, and estimates of treatment amenability.

#### Participants

A total of 392 individuals attending the Biannual American Psychology and Law Society (APLS) conference participated in the current research. The sample consisted of 220 (56.1%) females and 166 (42.3%) males. Six (1.5%) individuals failed to provide information regarding

their gender. The two largest groups (see Table 3) were Ph.D. psychologists ( $n = 204$  or 52.0%) and psychology graduate students ( $n = 134$  or 34.2%). Participants had an average of 9.64 ( $SD = 10.09$ ) years of experience with a range of 0 to 40 years. Nearly one-third of the sample (32.1%) reported having no experience conducting risk assessments, with a smaller proportion (20.2%) having very extensive experience (i.e.,  $\geq 50$  risk assessments conducted).

Table 3

*Training and Specialization of Study Participants*

Training	Frequency
Ph.D.	204
Clinical	81
Counseling	10
Experimental	8
Social	22
Other Psychology	22
Not Reported	61
Psy.D.	16
Graduate Student	134
Attorney/Judge	6
Undergraduate	16
Other/Not Reported	16
Total	392

## Measures and Materials

### *Case Materials*

An important facet of risk assessment research is external validity. In order to increase the ecological validity of the current study, case materials from six actual criminal defendants were obtained. Two types of cases (sexually violent recidivators and nonsexually violent recidivators) were chosen to be representative of individuals who are likely candidates for SVP commitment. Each case was selected based on a history of acts of both sexual violence and nonsexual violence. Each case was also chosen by the availability of sufficient detail to rate each of the assessment methods (actuarial and structured clinical guides) as well as provide an interview and offense history for the unstructured clinical judgment condition.

Cases were collected from a researcher (S. D. Hart , personal communication, January, 2002) with extensive clinical and research experience in risk assessments. In addition to comprehensive case materials suitable for the current study, he had recidivism data for each defendant that was collected as part of an unpublished research project. Follow-up data were available for five years for each defendant, allowing proper categorization of the offender as a future violent offender or sexually violent offender (three cases of each type).

### *Actuarial Instruments*

*Violence Risk Appraisal Guide (VRAG)*. The VRAG (Harris et al., 1993; © American Psychological Association, Washington, DC) is the best-validated actuarial instrument for the prediction of violent behavior (Quinsey et al., 1998). The VRAG consists of 12 variables each scored according to a weighting procedure based on the degree to which the individual varies from the base rate for that item in the initial sample. Based on this weighting procedure, the 12 VRAG variables are summed to a total score ranging from -26 to +38. The average VRAG score

in the original sample was close to zero (.91) with a standard error of measurement of 4.1 (Quinsey et al., 1998). The total scores are converted into nine categories of risk with probabilities of future violence at follow-up periods of 7 and 10 years assigned to each category.

Several investigators have evaluated the psychometric properties of the VRAG. Harris et al. (1993) calculated the interrater reliability ( $r = .90$ ) of the VRAG total score using a small sample of 20 randomly selected offenders. Other researchers also obtained excellent interrater reliability for the VRAG total score in terms of Pearson's  $r$  ( $r = .90$ ; Barbaree et al., 2001) and ICCs (ICC = .95; Kroner & Mills, 2001). For predictive validity, ROCs were moderate, with AUCs ranging from .61 for sexual recidivism (Barbaree et al., 2001) to .77 for all violent recidivism (Rice & Harris, 1997). Further, Barbaree et al. reported that the VRAG correlated moderately ( $r = .45, p < .001$ ) with general recidivism and modestly ( $r = .24, p < .001$ ) with violent recidivism. The latter correlation with violent recidivism was considerably lower than originally reported by Harris et al. (1993;  $r = .44$ ).

*Static-99.* The Static-99 (Hanson & Thornton, 1999), described previously, was chosen for the present investigation based on its validity in predicting future sex offenses. This instrument is a conglomeration of both the Rapid Risk Assessment for Sex Offence Recidivism (RRASOR; Hanson, 1997) and the Structured Anchored Clinical Judgment (SACJ; Grubin, 1998). Hanson and Thornton (1999) demonstrated its incremental validity over both individual measures. The total score of the Static-99 ranges from 0-12 and is categorized as four different levels of risk (i.e., 0-1 for low risk, 2-3 for medium-low risk, 4-5 for medium-high, and  $\geq 6$  for high risk).

Regarding its psychometric properties, Barbaree et al. (2001) reported excellent interrater reliability ( $r = .90$ ) for the Static-99 total score based on 30 cases. Predictive validity for sexual

recidivism in terms of AUC is reported to be .70 (Barbaree et al., 2001; Hanson & Thornton, 1999). However, total scores correlated very modestly with sexual recidivism ( $r = .18, p < .05$ ).

### *Structured Clinical Judgment*

*Historical Clinical Risk- 20 (HCR-20) Version 2.* Webster et al. (1997; © Mental Health Policy and Law Institute, Burnaby, Canada) developed the HCR-20 Version 2 as a revision of the earlier HCR-20 (Webster, Eaves, Douglas, & Wintrup, 1995). The HCR-20 is a 20-item clinical guide that consists of three separate scales: *Historical*, *Clinical*, and *Risk*.

Few investigators have examined the HCR-20's psychometric properties. Reliability analyses have demonstrated adequate internal consistency and interrater reliability. Specifically, Cronbach's alpha for the total score varies from .78 (Belfrage, 1998) to .95 (Douglas et al., 1999). In terms of interrater reliability, Douglas et al (1999) reported an Intraclass Correlation Coefficient (ICC) for the total score of .80 with subscales ranging from .70 (Clinical subscale) to .87 (Historical subscale).

Validity studies have demonstrated that the HCR-20 also has criterion-related validity. For example, Kroner and Mills (2001) found a significant but modest ( $r = .16$ ) relationship between the HCR-20 and reconvictions for violence in a sample of male offenders. Investigating concurrent validity, Douglas and Webster (1999) reported significant correlations between the HCR-20 scales and violence (Historical  $r = .50, p < .001$ ; Clinical  $r = .30, p < .01$ ; and Total  $r = .44, p < .001$ ). The considerable differences found by the researchers is likely due to the differences in methodology: predictive follow-up (Kroner & Mills, 2001) versus retrospective chart review (Douglas & Webster, 1999).

*The Sexual Violence Risk-20 (SVR-20).* Boer et al. (1997; © Mental Health Policy and Law Institute, Burnaby, Canada) developed the SVR-20 specifically for use with sex offenders.

Data on this instrument's validation have not been published. Conceptually, the measure has promise for risk assessment given its special attention to factors related to sexual violence. The present investigation is the first known study to utilize the SVR-20 and report its psychometric properties.

Interrater reliability was calculated based on three randomly chosen cases. The principal investigator and an advanced doctoral student independently rated each of the actuarial and structured clinical assessment methods from the available case materials. Intraclass Coefficients (ICC) were calculated for each measure's total score. ICCs for the actuarial measures were moderate (VRAG = .74 and Static-99 = .77)<sup>4</sup>. ICCs for the structured clinical judgment measures were outstanding (HCR-20 = .99, SVR-20 = .95).

#### *Unstructured Clinical Judgement*

Weiner (1999) has articulated the need to focus forensic reports on the referral question and recommends that risk assessment reports include past recidivism, treatment response, and community resources (p. 514). Weiner's (1999) recommendations were followed in developing unstructured clinical reports for the current study. Information given to participants in unstructured clinical judgment groups included a summary of clinical interviews conducted with the defendant that includes important variables such as background information, substance abuse history, course of treatment while incarcerated, and plans for release (see Appendix A).

---

<sup>4</sup> Interestingly, one item in particular attenuated the reliability estimates of the VRAG and Static-99. One offender had conflicting information regarding his marital status, resulting in conflicting ratings.

### *Victim Impact Statements*

The Committee on Ethical Guidelines for Forensic Psychologists (1991; see also Rogers & Shuman, 2000) recommend reviewing multiple sources of data when conducting a forensic assessment. Commentator such as Ziskin (1995) have suggested that emotionally evocative information may substantially influence a forensic expert's decisions. Few empirical studies (Lynett & Rogers, 2000) have investigated the potential effects of emotionally-laden information on experts' conclusions.

An important source of emotionally evocative information is accounts given by victims of crime. Victim impact statements (see Appendix B), provide a description of the offense from the perspective of the victims/survivors and includes their immediate physical and psychological reactions. In this investigation, a victim impact statement was adopted from prior research (Lynett, 1990; Lynett & Rogers, 2000) on dangerousness that demonstrated a potentially biasing effect.

## Procedure

### *Consent and Recruitment*

Institutional Review Board (IRB) approval was obtained from the University of North Texas. Following IRB approval, APLS representatives organizing the convention were contacted. The purpose and relevance of the project was explained and approval to conduct data collection at the conference was obtained.

Professionals and graduate students attending the American Psychology Law Society (APLS) convention and the American Academy of Forensic Psychology (AAFP) workshop were invited to participate in the current study by one of three researchers. As part of registration, each APLS attendee was required to check-in at the conference registration table. Following



registration, a researcher explained the general purpose of the study and invited participation. The primary researcher also invited AAFP workshop attendees to participate by addressing the group during the opening remarks of each workshop session and distributing data collection packets. All conference and workshop attendees were composed of forensic psychologists, psychiatrists, and doctoral students in psychology; all were invited to participate.

### *Experimental Assignment*

Thirty-six unique experimental packets were assembled (six offenders by six unique orderings of information). Prior to the convention, the researcher assembled batches of data collection packets, with each batch consisting of one each of the 36 unique packets. Prior to assembly, a randomized order was established and repeated for each batch. Individuals agreeing to participate were given the next available packet in the batch. After a batch was completely distributed, the next batch was opened. This quasi-random distribution assured relatively equal representation in each group. Data on six different offenders were utilized in order to minimize stimulus restriction concerns. No analyses by specific offender were hypothesized. Therefore, the number of participants in the most detailed level of analysis (i. e., order effects) ranged from 61 to 71.

### *Completion of Experimental Protocols*

Participants received a packet of information introduced by a cover letter (see Appendix C). The cover letter provided a description of the study, the confidential nature of responses, and a statement that IRB approval had been received. In accordance with IRB approval, the cover letter also stated that completion of the questionnaire served as informed consent. Following the cover letter, the packet included an information sheet consisting of the referral question and offense history of the offender (see Appendix D) and the experimental materials.

Participants completed the study in three consecutive phases. In Phase I, the participants reviewed the referral question and offense history of the offender and received the first component of risk assessment data: actuarial, structured clinical, or unstructured clinical. Participants made initial ratings based upon this Phase I data (see Appendix E).

During Phase II, participants reviewed the second type of assessment data and rated the offender's risk a second time. Victim impact statements were introduced during Phase III. Each participant read an identical victim impact statement and completed ratings a third time based on the addition of this information. Finally, participants provided details regarding the nature of their clinical practice and experience conducting risk assessments (see Appendix F).

Upon completion and return of the data packet, the participant received a thank-you gift for his or her involvement. The primary researcher attempted to answer any general questions regarding the experiment. A full debriefing was deemed unfeasible in the convention setting. The researcher was not completely confident that specific hypotheses and details of the study would remain confidential throughout the convention. Therefore, specific hypotheses were not disclosed in an effort to limit possible demand characteristics associated with such knowledge.

## CHAPTER 3

### RESULTS

#### Demographic and Descriptive Characteristics

The American Psychology-Law Society (APLS; Division 41 of the American Psychological Association) attracts a multidisciplinary membership including psychologists and attorneys as well as judges and sociologists. The society also enjoys a large student membership consisting mostly of graduate students. APLS conference attendance is equally diverse and as a result, participants in the current study reflect a diversity of backgrounds, training, education, and experience (see Table 4).

Table 4

*Education/Occupation of Study Participants*

Degree Held	Frequency ( <i>n</i> )
Ph.D. Psychologists	204
Psy.D. Psychologists	16
Graduate Student	134
Attorney/Judge	6
Undergraduate Students	16
Other <sup>a</sup>	12
Not Reported	4
Total	392

*Note:* <sup>a</sup> The “Other” category includes individuals with diverse backgrounds such as sociologists, parole officers, and law enforcement/corrections officials.

An interesting question is whether different specialties within psychology differ in their professional experience, particularly forensic experience. Within the doctoral-level psychologists, four main specialties were reported including clinical, counseling, experimental, and social. Doctoral-level psychologists reported an equal number of years practicing at the doctoral level, but specialized experience varied according to discipline.

Table 5 presents data reported for years practicing at the doctoral level, number of risk assessments conducted, and percentage of practice devoted to forensic issues for each of the doctoral-level participants. As expected, Clinical Ph.D., Clinical Psy.D., and counseling Ph.D. psychologists each reported a significantly greater percentage of forensic practice ( $F [6,190] = 13.84, p < .001$ ) and significantly more experience conducting risk assessments ( $F [6, 210] = 8.09, p < .001$ ) than social or experimental psychologists.

Table 5

*Risk Assessment Experience and Percentage of Forensic Practice by Specialty Training in Psychology*

Specialty ( <i>n</i> )	Years Practicing <i>M (SD)</i>	Risk Assessments Conducted <i>M (SD)<sup>a</sup></i>	Forensic Practice (%) <i>M (SD)</i>
Clinical Ph.D. (81)	11.26 (9.59)	68.96 (87.16)	67.70 (38.46)
Clinical Psy.D. (16)	7.44 (4.30)	83.53 (80.57)	76.00 (36.52)
Counseling (10)	8.50 (8.20)	38.33 (34.27)	58.00 (42.11)
Experimental (8)	9.57 (11.24)	0.00 (0.00)	10.00 (17.32)
Social (22)	8.31 (8.72)	.28 (1.18)	1.94 (6.28)

<sup>a</sup> Due to the presence of outliers ( $z$  score  $> 3.3$ ; Tabachnick & Fidell, 2001) in the number of risk assessments conducted, 7 cases were deleted from this analysis.

*Note.* Tukey’s post hoc analysis indicates no difference ( $p > .05$ ) among clinical Ph.D., clinical Psy.D., and counseling Ph.D. psychologists.

For purposes of the current analyses, the most relevant subgroup of participants consists of clinical and counseling psychologists. As Table 6 demonstrates, this combined sample ( $n = 107$ ) has significantly more forensically-relevant experience than other Ph.D. level psychologists. As expected, the refined sample also reported significantly more forensic experience than graduate students (see Table 6).

Table 6

*Risk Assessments and Forensic Practice of Doctoral-Level Clinicians (i.e., Clinical and Counseling Psychologists) Versus Other Psychologists and Graduate Students*

	Doctoral-Level Clinicians ( $n = 107$ )	Other Psychologists ( $n = 30$ )	Graduate Students ( $n = 134$ )	$d1^b$	$d2^c$	$F$
Risk Assessments Conducted <sup>a</sup>	68.96 (87.16)	.19 (.98)	16.84 (37.96)	.89	.81	19.64***
Forensic Practice (%)	68.04 (38.40)	4.39 (11.09)	39.22 (41.72)	1.85	.72	31.73***

*Note:* According to Tukey’s post hoc analysis, doctoral-level clinicians have conducted significantly ( $p < .05$ ) more risk assessments and devote more of their practice to forensic issues than other psychologists and graduate students.

<sup>a</sup> Outliers eliminated from analysis

<sup>b</sup>  $d1$  = Cohen’s  $d$  calculated between doctoral-level clinicians and other psychologists.

<sup>c</sup>  $d2$  = Cohen’s  $d$  calculated between doctoral-level clinicians and graduate students.

\*\*\*  $p < .001$

Primary research questions are predominately investigated utilizing this refined group of doctoral-level clinicians ( $n = 107$ ). The refined sample's experience is most likely to parallel individuals conducting risk assessments for the courts, thereby bolstering the ecological validity of the results of the present study. Particular research questions investigating the effects of training and experience employ the additional sample of 134 psychology graduate students.

#### Accuracy of Violence Assessments

Eighty-five (79.4%) of the doctoral-level clinicians accurately predicted future violence during Phase I. Following review of the second assessment method, the number remained comparable with 88 (82.2%) correctly predicting violence. Interestingly, very few participants (i.e., 5 or 4.7%) altered their categorical prediction of violence on the basis of additional assessment information.

As expected, clinicians predicting future violence rated offenders as significantly more likely to commit acts of nonsexual violence and sexual violence. They also supported Sexually Violent Predator (SVP) commitment to a greater degree than did clinicians predicting nonviolence (see Table 7).

Table 7

*Violence and Sexual Violence Likelihood Ratings and Indicated Support for SVP Commitment by Violent and Nonviolent Predictions*

Estimate	Categorical Prediction		<i>F</i>	<i>d</i>
	Violent <i>M (SD)</i>	Nonviolent <i>M (SD)</i>		
Likelihood of nonsexual violence	57.36	23.05	38.28***	1.57

	(22.82)	(16.68)		
Likelihood of sexual violence	54.13	22.00	45.49***	1.71
	(19.72)	(13.56)		
Support of SVP Commitment	43.28	14.74	13.56***	.93
	(32.79)	(16.37)		

---

*Note:* Estimates are rated as continuous variables ranging from 0-100%.

\*\*\*  $p < .001$

### *Clinical Decision Making in Predicting Future Sexual or Nonsexual Violence*

The proper application of sexually violent predator laws require a finding that the offender is likely to commit future acts of *sexual* violence. To most adequately address the standard, experts must identify which offenders are likely to be sexually violent in the future. Doctoral-level clinicians' ability to predict which offenders are likely to commit future acts of sexual violence versus those likely to commit nonsexual acts of violence was inspected both dimensionally and categorically.

A within-subjects ANOVA was utilized to inspect dimensional ratings of violence. A significant interaction was expected between the continuous dependent variables of likelihood of future sexual and nonsexual violence and the categorical variable of actual outcome, with sex offenders receiving higher ratings of potential for sexual violence and nonsexual offenders receiving higher ratings for general violence potential. However, a within-subjects ANOVA indicated no significant interaction ( $F [1, 104] = .30, p > .05$ ), indicating that clinicians' estimates of type of violence did not vary systematically with actual outcome.

Chi-square analysis was used to inspect categorical estimates of sexual and nonsexual violence. No significant relationship was found between type of predicted violence and actual

outcome ( $X^2 [1] = .90, p > .05$ )<sup>5</sup>. Thirty-seven (34.6%) of the 107 doctoral-level clinicians accurately discriminated type of violence. Utility estimates were calculated to inspect participants' ability to accurately predict sexual violence (see Table 8). Of the participants who made inaccurate predictions of sexual violence an equal number overpredicted or underpredicted sexual violence ( $ns = 18$  and  $19$ , respectively).

Table 8

*Utility Estimates Calculated for Clinicians' Ability to Predict Sexual Violence*

PPP	NPP	Sensitivity	Specificity	Hit Rate
.44	.55	.42	.56	.50

*Note:* PPP = Positive Predictive Power; NPP = Negative Predictive Power

*Research Questions #1 and #2*

Overall, clinicians were unable to discriminate between offenders likely to commit future acts of violence versus sexual violence, making correct predictions only half the time. Research Questions #1 and #2 both begin to explore systematic differences in accuracy based on specific assessment methods, testing the relative utility of actuarial, clinical, and structured clinical judgements in making predictions regarding future risk. Research Question #1 explores predictive accuracy in general while Research Question #2 investigates predictions of sexual versus nonsexual violence specifically.

<sup>5</sup> In order to categorically inspect participants' ability to discriminate between offenders' likelihood of future sexual or nonsexual violence, a dichotomous prediction variable was calculated based on participants' ratings of likelihood of violence or sexual violence. Participants who rated an offender's likelihood of nonsexual violence higher than his likelihood of sexual violence were considered to be predicting nonsexual violence and vice versa. All ratings were used, regardless of absolute magnitude, to detect all instances of discrimination between types of violence.



*Category Accuracy by Clinical Decision Making Method*

First, doctoral-level clinicians’ ability to predict violence based on a single assessment method was inspected. As presented earlier, doctoral-level clinicians predicted future violence with over 80% accuracy. Chi-square analysis indicates that categorical predictions do not differ based on type of assessment method ( $X^2 [2] = 1.93, p > .05$ ). Table 9 presents clinicians’ binary prediction based on the primary assessment method.

Table 9

*Categorical Prediction Accuracy (Violent versus Nonviolent) Utilizing a Single Decision Making Method*

Clinical Decision Making Method	(Nonviolent) Inaccurate Prediction <i>n</i> (%)	(Violent) Accurate Prediction <i>n</i> (%)	$X^2 / p$
Actuarial	6 (17.6)	28 (82.4)	1.93/ <i>ns</i>
Structured Clinical Judgment	7 (16.7)	35 (83.3)	
Unstructured Clinical Judgment	9 (29.0)	22 (71.0)	

*Note:* As all offenders in the current study were violent recidivators, nonviolent predictions are *inaccurate* predictions. Nonsignificant  $X^2$  indicates no relationship between accuracy and decision making method.

Doctoral-level clinicians’ use of the assessment methods did not establish any specific method as superior in its accuracy at distinguishing between future sexual or nonsexual violence. As noted in Table 10, none of the three methods differed significantly from the base rate (50.0%) for sexual violence.

Table 10

*Hit Rate and  $X^2$  Statistics for Clinical Decision Making Regarding Predictions of Nonsexual and Sexual Violence*

Decision Making Method	Overall Hit Rate	$X^2$ / <i>p</i>
Actuarial	.48	.00/ <i>ns</i>
Structured Clinical Judgment	.56	.57/ <i>ns</i>
Unstructured Clinical Judgement	.52	.05/ <i>ns</i>

*Note:*  $X^2$  value indicates accuracy within each category of decision making method.

*Dimensional Accuracy by Decision Making Method*

Clinicians' ability to distinguish sexual from nonsexual violence was also tested via dimensional ratings. Specifically, clinicians' likelihood ratings were examined via repeated measures analysis. In the current analysis, decision-making method was added to the within subjects ANOVA to detect potential interactions with offender type and violence estimates. A nonsignificant three-way interaction between clinical decision making method, actual outcome, and likelihood ratings of sexual and nonsexual violence indicated no difference in ability to discriminate ( $F [2, 100] = 1.54, p > .05$ ; see Tables 11 and 12).

Three-way interactions are particularly difficult to interpret. Despite the nonsignificance of this interaction, closer inspection is warranted to more fully address the research questions. Specifically, the research questions explore the utility of different decision making methods in discriminating between sexual and nonsexual violence. Therefore, Tables 11 and 12 were designed to facilitate inspection of that interaction, with each table representing the results for each condition of the third independent variable, namely, offender outcome. Table 11 presents

the interaction between decision making method and estimates of sexual and nonsexual violence for offenders who were nonsexually violent. Table 12 presents information regarding those offenders who were sexually violent.

Table 11

*Likelihood Estimates (Expressed in Percentages) of Nonsexual and Sexual Violence for Nonsexually Violent Offenders by Clinical Decision Making Method*

Decision Making Method ( <i>n</i> )	Nonsexual Violence Estimate <i>M (SD)</i>	Sexual Violence Estimate <i>M (SD)</i>	<i>F*</i>	<i>d</i>
Actuarial (21)	48.62 (21.34)	50.57 (18.43)		.10
Structured Clinical (22)	59.14 (26.38)	48.50 (21.08)		.45
Unstructured Clinical (15)	42.07 (22.61)	32.07 (15.65)		.51
			1.54 <i>ns</i>	

*Note:* F value represents omnibus F obtained for the nonsignificant three-way interaction.

The results presented in Table 11 can be interpreted in two different ways. First, the use of both structured and unstructured clinical judgment resulted in the largest effect sizes between violent and sexual violence estimates, indicating that clinicians utilizing the clinical judgment methods may have the best *discriminability*. They tend to rate violent offenders as more likely to commit violent (rather than sexually violent) acts. Despite the lack of statistical significance, the Cohen's *d* values (*ds* = .45 and .51) suggest a modest effect. Importantly, the mean ratings offered by clinicians utilizing unstructured clinical judgment fall below 50%. It is difficult to

assert that any ratings indicating less than 50% likelihood of future violence do indeed constitute a prediction of violence at all.

A second interpretation of Table 11 suggests that clinicians are better able to predict violence of any kind when they utilize a standardized method of assessment (i.e., actuarial or structured clinical judgment). Inspection of mean ratings (i.e.,  $M$  ratings of nonsexual and sexual violence) demonstrates that clinicians utilizing structured clinical judgment ( $M = 53.82$ ) and actuarial methods ( $M = 49.60$ ) produced higher estimates of potential violence than clinicians utilizing unstructured clinical judgment ( $M = 37.07$ ). The difference between mean ratings made by clinicians utilizing structured clinical methods versus those utilizing unstructured methods is moderate ( $d = .75$ ). Furthermore, clinicians utilizing structured clinical judgment indicated a greater than 50% likelihood that the offender will commit some act of violence, at least meeting a logical threshold to be considered a prediction of violence.

An additional hypothesis generated by inspected of Table 11 is that clinicians utilizing unstructured clinical judgment recognize the limitations of the unstructured approach and lower their estimates accordingly. The data allow for inspection of this hypothesis via inspection of confidence estimates. If clinicians utilizing unstructured clinical judgment offered low estimates of violence based on the limitations of the unstructured approach, confidence ratings of their estimates will also be modest. This hypothesis will be addressed in the discussion of confidence estimates results (i.e., Research Question #4).

Table 12 presents the interaction between clinical decision making method and violence estimates for *sexually* violent offenders. Effect sizes between ratings of sexual and nonsexual violence were even more modest for sexually violent offenders (range of  $ds = .14$  to  $.28$ ).

Clinician ratings indicate that they did not perceive sexual offenders to be any more likely to commit sexual rather than nonsexual violence.

A potential hypothesis is that sex offenders are simply perceived as more likely to be violent (both sexually and nonsexually) than nonsexual offenders, attenuating any difference between ratings for a specific type of violence. However, a nonsignificant main effect for offender outcome ( $F [1, 100] = .03, p > .05$ ) demonstrated that violent and sexually violent offenders were seen as equally likely to become dangerous. Table 12 indicates that clinicians were less able to discriminate between potential for sexual and nonsexual violence in future sex offenders than they were for nonsexually violent offenders. In addition, an interesting trend emerges from inspection of Tables 11 and 12 that suggests the use of unstructured clinical judgment in isolation appears to *underestimate* the risk of both sexual and nonsexual violence potential.

Table 12

*Likelihood Estimates (Expressed in Percentages) of Nonsexual and Sexual Violence for Sexually Violent Offenders by Clinical Decision Making Method*

Decision Making Method ( <i>n</i> )	Nonsexual Violence Estimate <i>M (SD)</i>	Sexual Violence Estimate <i>M (SD)</i>	<i>F*</i>	<i>d</i>
Actuarial (13)	50.15 (21.16)	43.77 (24.28)		.28
Structured Clinical (20)	51.25 (28.32)	47.50 (23.81)		.14
Unstructured Clinical (15)	39.60 (26.21)	44.53 (26.34)		.19
			1.54 <sub>ns</sub>	

*Note:* F value represents omnibus F obtained for the nonsignificant three-way interaction.

### *Combined Assessment Approach*

*Incremental Validity of the Combined Assessment Approach.* The second step in analyzing the effects of assessment method was to examine predictive ability when combining methods. To inspect incremental validity, judgments rendered during Phase II were based on information combining two different methods. The design allows for the inspection of incremental validity achieved with additional information in general as well as improvements in prediction based on the unique combination of decision making methods. As outlined in the Methods chapter, six unique orderings of assessment method were utilized: actuarial/structured, actuarial/unstructured, clinical/actuarial, clinical/unstructured, unstructured/actuarial, and unstructured/clinical.<sup>6</sup>

Clinicians rated all offenders as significantly more likely to be sexually violent when they inspected two sources of assessment data. A repeated-measures ANOVA inspected changes in perceptions regarding offender violence. A main effect for likelihood of sexual violence indicated that clinicians increased the magnitude of their estimates of sexual violence when additional information of any kind was reviewed ( $F [1,94] = 4.43, p < .05$ ). The effect of this difference was modest ( $M_{\text{time1}} = 45.25$  versus  $M_{\text{time2}} = 48.42; d = 14$ ; see Table 13). However, there was no interaction between order of assessment method and changes in magnitude of prediction ( $F [5,94] = 1.01, p > .05$ ). Clinicians' estimates of nonsexual violence did not increase significantly with the addition of data ( $F [1,94] = 3.57, p > .05$ ).

---

<sup>6</sup> Clinical = Structured Clinical Judgment; Unstructured = Unstructured Clinical Judgment

Table 13

*Estimates of Sexual Violence: Effects of One Versus Two Assessment Methods*

Decision Making Method	Sexual Violence Estimate		<i>d</i>
	One Assessment Method	Combined Assessments	
( <i>n</i> )	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	
Actuarial/Structured (21)	50.48 (21.19)	56.10 (21.70)	.26
Actuarial/Unstructured (13) <sup>a</sup>	43.92 (20.25)	49.31 (23.08)	.25
Clinical/Actuarial (19)	43.00 (22.60)	46.84 (24.16)	.16
Clinical/Unstructured (23)	52.17 (21.36)	49.13 (21.51)	.14
Unstructured/Actuarial (10) <sup>a</sup>	37.00 (22.01)	39.80 (20.99)	.13
Unstructured/Clinical (20)	38.95 (22.85)	44.80 (22.69)	.26
Total	45.25 (21.97)	48.42 (22.38)	.14

*Note:* <sup>a</sup> Cell sizes (< 15) were less than optimal.

Estimates of potential sexual violence increased nonsignificantly within each ordering condition (see Table 13). Surprisingly, increases in estimates of sexual violence were unrelated to the actual outcome of the offender ( $F [1,94] = .10, p > .05$ ). Following review of additional data (i.e, Phase II ratings) nonsexual offenders were still seen as equally likely to be sexually violent as were sexual offenders ( $M_{\text{violent}} = 48.64$  versus  $M_{\text{sexual}} = 48.17$ ;  $F [1,94] = .00, p > .05$ ).

*Categorical Accuracy of Phase II Violence Predictions.* Accurate discrimination of sexual and nonsexual violence was unrelated to order of assessment ( $X^2 [5] = 2.01, p > .05$ ). As presented in Table 14, clinicians were unable to distinguish significantly sexual from nonsexual offenders based on the combined assessment approach. Clinicians utilizing a combination of clinical and unstructured clinical judgement achieved the highest hit rate with 65% accurate

predictions. As the associated chi-square value indicates, however, the relationship between predictions and accuracy is still nonsignificant.<sup>7</sup>

Table 14

*Testing Order Effects on Sexual Versus Nonsexual Prediction Accuracy (i.e., Hit Rate) by Assessment Method*

Order of Assessment Method Presentation	Hit Rate*	$X^2/p^a$
Actuarial/Clinical	.50	.00/ <i>ns</i>
Actuarial/Unstructured	.46	.02/ <i>ns</i>
Clinical/Actuarial	.43	.12/ <i>ns</i>
Clinical/Unstructured	.65	1.63/ <i>ns</i>
Unstructured/Actuarial	.57	.06/ <i>ns</i>
Unstructured/Clinical	.50	.00/ <i>ns</i>

*Note:* Hit rate indicates percentage of correct predictions of either nonsexual or sexual violence.

<sup>a</sup>  $X^2$  values calculated separately for each condition.

*Research Question #3*

Research Question #3 asked specifically whether the addition of clinical data reduced the accuracy of actuarial data. Results presented above suggest this is not the case. However, to further inspect this question, data were analyzed separately for clinicians receiving actuarial data first, followed by either structured or unstructured clinical data. Dichotomous accuracy was

<sup>7</sup> Due to suboptimal cell sizes (< 15), between and within-subjects ANOVA analyses of the combined assessment approach are highly speculative. Results and discussion of these analyses are found in Appendix G.



unaffected by the addition of clinical data ( $X^2_{\text{actuarial}} [1] = .95, p > .05$  vs.  $X^2_{\text{actuarial + clinical}} [1] = .91, p > .05$ ).

Previous repeated measures analyses demonstrated that changes in estimates of sexual or nonsexual violence following the addition of clinical data were not significant depending on order assessment method ( $F_{\text{sex} \times \text{order}} [1,5] = 1.01, p > .05$ ;  $F_{\text{nonsexual} \times \text{order}} [1,5] = 1.34, p > .05$ ). A nonsignificant interaction indicates no particular assessment method appreciably affected appropriate ratings significantly. To present the issue more clearly, Table 15 displays clinicians' mean ratings of violence and sexual violence when based solely on actuarial data as well as ratings made following review of either structured or unstructured clinical data.

Table 15

*Mean (SD) Ratings of Potential Violence and Sexual Violence Based on Actuarial and Actuarial Plus Clinical Data*

<u>Violent Offenders' Estimated Risk of Future Violence</u>				
	<u>Phase I Estimates</u>		<u>Phase II Estimates</u>	<u>d</u>
Actuarial Data	48.62 (21.34)	Actuarial + Structured	55.50 (16.27)	.35
		Actuarial + Unstructured	48.89 (24.85)	.01
<u>Sex Offenders' Estimated Risk of Future Sexual Violence</u>				
Actuarial Data	43.77 (24.28)	Actuarial + Structured	54.11 (26.80)	.41
		Actuarial + Unstructured	38.75 (25.94)	.20

Contrary to Quinsey et al.'s (1998) assertion that clinical data reduced the accuracy of actuarial data, no significant decreases in continuous ratings of likelihood of sexual or nonsexual

violence were detected. Clinicians decreased, although not significantly, their estimates of sexual violence when they reviewed additional data in the form of unstructured clinical judgment. Despite this modest decrease ( $d = .20$ ), their ability to discriminate between sexual and nonsexual offenders actually increased (see also Table 1 in Appendix G). Of interest, clinicians who paired structured clinical data with actuarial data produced the only predictions of violence greater than 50% in the appropriate outcome category (i.e., violent offenders vs. sex offender).

### Confidence

#### *Research Question #4*

Research Question #4 investigated clinicians' confidence levels in predicting future violence. Clinicians rated their level of confidence in estimates of both nonsexual and sexual violence on a scale ranging from 0% confident to 100% confident. Clinicians were equally confident in predicting both nonsexual violence ( $M = 56.45%$ ) and sexual violence ( $M = 53.87%$ ,  $t [105] = 1.66, p > .05$ ).

Clinicians appeared uninfluenced by assessment method as no significant differences in confidence of either nonsexual or sexual violence were found depending on assessment method. In addressing the earlier hypothesis generated by Table 11, it appears that clinicians utilizing unstructured clinical judgment did not offer predictions with lower levels of confidence. Despite their modest estimates of potential violence, these clinicians were equally confident as clinicians utilizing other assessments methods.

Surprisingly, clinicians who predicted future violence were significantly more confident in their ratings than were clinicians predicting nonviolence (see Table 16). The current study lacked a control group of offenders who were actually not violent. Therefore, it is not possible to discern from these data whether this difference reflects either (a) a tendency for individuals who

are accurate in their predictions to be more confident or (b) whether individuals who predict future violence are more confident. However, this question is addressed further in supplementary analyses.

Table 16

*Clinicians' Confidence in Their Ratings of Future Violence (Means [Standard Deviations])*

	Categorical Prediction		<i>F</i>	<i>d</i>
	Violent ( <i>n</i> = 87)	Nonviolent ( <i>n</i> = 19)		
Confidence of nonsexual violence	62.42 (21.99)	48.95 (26.80)	5.39*	.59
Confidence of sexual violence	60.47 (21.58)	45.79 (25.35)	6.75*	.66

*Note:* Confidence ratings are based on a continuous scale ranging from 0 to 100%

\*  $p < .05$

Confidence of predictions of both nonsexual and sexual violence increased with the addition of data (see Table 17). No interaction between increase in confidence and assessment method was found, indicating simply the addition of any data, rather than a specific assessment method, was responsible for the increase in confidence.

Table 17

*Confidence of Sexual and Nonsexual Violence Predictions when Judgments Are Rendered*

*Following One Decision Making Method Versus A Combined Assessment Approach*

	One Assessment Method	Combined Methods	<i>F</i>	<i>d</i>
Nonsexual violence	56.45 (23.18)	59.90 (23.29)	15.57***	.15

Sexual violence	53.87 (23.14)	57.83 (22.79)	14.49***	.17
-----------------	---------------	---------------	----------	-----

\*\*\*  $p < .001$

### Research Question #5

Research Question #5 explored the relationship between accuracy and confidence. Are accurate judges likely to be more confident in their ratings? Confidence ratings of sexual and nonsexual violence were highly correlated with each other ( $r = .76, p < .01$ ), but were unrelated to clinicians' accurate predictions of sexual and nonsexual violence ( $r_s = .005$  and  $-.107$  respectively,  $p_s > .05$ ). As a second step, we investigated significant differences in confidence levels via ANOVAs using the dichotomous variable of accurate discrimination as an independent variable. As Table 18 demonstrates, no significant differences were found on either confidence rating. A slight, but nonsignificant trend emerged indicating that inaccurate clinicians were more confident in their predictions, although the effect sizes were modest ( $d_{\text{nonsexual}} = .26; d_{\text{sexual}} = .18$ ).

Table 18

### *Confidence of Accurate versus Inaccurate Clinicians' Estimates of Sexual and Nonsexual Violence*

	Inaccurate Prediction	Accurate Prediction	<i>t</i>	<i>d</i>
Confidence of Prediction				
Nonsexual Violence	67.68 (18.38)	61.74 (26.05)	.84 $ns$	.26
Sexual Violence	64.44 (21.82)	60.71 (18.49)	.51 $ns$	.18

## Treatment Amenability

### *Research Question #6*

Clinicians were asked to rate offenders' amenability to treatment on a Likert-type scale ranging from 1 (not at all treatable) to 7 (very treatable). Research Question #6 inspected differences in treatment amenability ratings based on the varying assessment methods. After reviewing one assessment method, clinicians perceived the offenders to be marginally treatable ( $M = 3.31, SD = 1.28$ ). Individuals reviewing unstructured clinical judgement rated offenders as significantly more amenable to treatment ( $M = 3.81, SD = 1.33$ ) than individuals reviewing structured clinical judgment ( $M = 3.00, SD = 1.17; F [2, 100] = 3.79, p < .05; d = .65$ ).

Following review of additional data, a significant interaction between order of assessment method and change in treatment amenability ratings emerged ( $F [5, 96] = 4.84, p < .001$ ). Table 19 presents clinicians' estimates of treatability based on a single assessment method as well as the combined assessment method approach. Individuals who reviewed structured clinical judgement first ( $M = 3.00, SD = 1.17$ ) did not change their ratings significantly based on the addition of actuarial data ( $M = 2.69, SD = .79, p > .05$ ). However, individuals who received structured clinical data followed by unstructured clinical data increased their estimates of treatability by nearly a point ( $M = 3.95, SD = 1.68; t [21] = 3.05, p < .01$ ; see Table 23), resulting in a significant difference in amenability to treatment ratings during Phase II ( $t [36] = 2.80, p < .01$ ; see Table 19).

Table 19

*Means (Standard Deviations) for Amenability to Treatment Estimates Based on Single and Combined Clinical Decision Making Methods*

	Single Method		Combined Methods	<i>d</i>
Actuarial	3.21 (1.24)	Actuarial/Structured	2.86 (1.11)	.29
Actuarial	3.21 (1.24)	Actuarial/Unstructured	3.77 (1.64)	.41
Structured	3.00 (1.17)	Structured/Actuarial	2.69 (.79) <sub>a</sub>	.29
Structured	3.00 (1.17)	Structured/Unstructured	3.95 (1.68) <sub>a</sub>	.69
Unstructured	3.81 (1.33)	Unstructured/Actuarial	3.20 (.92)	.49
Unstructured	3.81 (1.33)	Unstructured/Structured	3.90 (1.26)	.07

<sup>a</sup> Indicates significant difference in treatment amenability ratings via Tukey's post hoc test ( $p < .05$ ).

Violent and sexually violent offenders were seen as equally amenable to treatment ( $F [1,101] = 2.91, p > .05$ ) as were offenders who were predicted to be either violent or sexually violent ( $F [1,71] = 2.59, p > .05$ ). A significant relationship between likelihood of violence ratings and treatment amenability suggested that offenders perceived as less treatable were estimated to be more likely to commit acts of nonsexual violence ( $r = -.29, p < .01$ ) and sexual violence ( $r = -.24, p < .05$ ). Moreover, offenders perceived as less amenable to treatment were viewed as more likely to meet a sexually violent predator standard ( $r = -.24, p < .05$ ).

#### Support of SVP Commitments

##### *Research Question #7*

Research Question #7 inspected the degree to which professionals supported sexually violent predator commitments. For the current analysis, clinicians who rated an offender equally likely to commit acts of violence and sexual violence were included. Clinicians tended to support SVP commitment equally for offenders they perceived as likely to be sexually versus

nonsexually violent ( $F [2,102] = .48, p > .05$ ). Table 20 presents ratings of support for sexually violent commitment by clinicians' prediction. Paradoxically, offenders who were *predicted* to be nonsexually violent received slightly, although not significantly, greater support for SVP commitment ( $M_{\text{sexual}} = 35.52$  versus  $M_{\text{nonsexual}} = 41.90; d = .20$ ).

Table 20

*Professionals' Support of Sexually Violent Predator (SVP) Commitment\* by Predicted Violence*

*Type*

	Type of Violence Predicted			<i>F</i>
	Non Sexual ( <i>n</i> = 42)	Sexual ( <i>n</i> = 31)	Equal Likelihood ( <i>n</i> = 32)	
SVP Support	41.90% (28.86)	35.52% (37.10)	35.63% (32.34)	.48 $ns$

\* Support for Sexually Violent Predator (SVP) commitment was indicated on a continuous 0-100% scale.

Professionals' support for SVP commitment for offenders who were *actually* violent versus sexually violent was also nonsignificant ( $M_{\text{sexual}} = 37.02, SD 32.25$  vs  $M_{\text{nonsexual}} = 38.72, SD = 32.09, F [1,104] = .07, p > .05$ ). Indicated support of SVP commitment did not vary based on assessment method ( $F [1,5] = 1.45, p > .05$ ).

Professional Training Experience

*Research Question #8*

A final research question investigated the effect of level of training on participants' ability to predict violence. For the current analysis, doctoral-level clinicians ( $n = 107$ ) were compared to psychology graduate students ( $n = 134$ ).

Doctoral-level clinicians were no more able to accurately discriminate violent from sexually violent offenders than were graduate students ( $X^2 [1] = .19, p > .05$ ). Table 21 presents hit rates for doctoral-level clinicians and graduate students.

Table 21

*Overall Hit Rates and  $X^2$  for Sexual and Nonsexually Violent Outcomes by Level of Training*

	Hit Rate	$X^{2a}$
Graduate Students	.46	.74ns
Doctoral-Level Clinicians	.50	.02ns

<sup>a</sup>  $X^2$  values represent relationship between predicted and actual type of violence within each level of training category.

Graduate students and doctoral-level clinicians rated offenders as equally likely to commit acts of violence ( $M_{\text{students}} = 51.53, SD = 26.28$  vs.  $M_{\text{clinicians}} = 51.05, SD = 5.40; t [238] = .14, p > .05$ ) and sexual violence ( $M_{\text{students}} = 54.43, SD = 24.25$  vs.  $M_{\text{clinicians}} = 48.42, SD = 22.38, t [238] = 1.97, p = .05$ ).

Graduate students and doctoral-level clinicians did not differ in their indicated confidence levels or support of sexually violent predator commitment. Graduate students, however, rated offenders as significantly more amenable to treatment than did doctoral-level clinicians ( $M_{\text{clinicians}} = 3.43, SD = 1.38$  vs.  $M_{\text{students}} = 3.81, SD = 1.20; t [229] = 2.28, p < .05$ ).



## Supplemental Analyses: Professional Practice Issues

Research suggests that testimony offered with a high degree of confidence is more influential with the fact-finder than testimony offered with low levels of confidence (Fox & Walters, 1986; Penrod & Cutler, 1995; Wells, Lindsay, & Tousignant, 1980). This knowledge is likely to have two significant effects on the process of expert testimony in relation to future dangerousness: (a) fact-finders are likely to believe confident rather than non-confident witnesses, and (b) attorneys are likely to present expert testimony only when it can be offered with a high degree of confidence. To most closely approximate judgements in this study that would be most likely to be offered as expert testimony, a series of supplemental analyses were conducted utilizing only those clinicians who indicated a 50% or greater level of confidence in their judgements.

Experts' opinions presented with greater confidence are likely to be more persuasive. Therefore, a composite judgment consisting of estimated likelihood and adjusted for confidence level must be considered. By multiplying participants' likelihood estimates by their level of confidence, a composite estimate was calculated. The following analyses utilize the composite ratings as indicators of experts' final judgments.

### *Recommending SVP Commitment*

An exploratory analysis was conducted to investigate whether professionals use relevant information in making their commitment recommendations. Stepwise multiple regression found offenders who are most likely to receive support for an SVP commitment are seen as highly likely to be sexually violent, unable to control their behavior and as having a mental abnormality or personality disorder (see Table 22).

Table 22

*Stepwise Multiple Regression: Predictors of Support for SVP Commitment*

	R <sup>2</sup>	R <sup>2</sup> Change	Beta	p
Ability to Control Behavior	.19		.38	.001
Sexual Violence Composite	.32	.13	.39	.001
Presence of MA or PD	.37	.05	.24	.05

*Note.* Sexual Violence Composite was calculated by multiplying predictions of sexual violence by the confidence of that estimate; MA = mental abnormality; PD = personality disorder.

*Effects of Victim Impact Statements*

An important consideration when evaluating predictive accuracy is the quality of information that contributes to the judgment. The victim impact statement used in the present study was a statement used in previous research (Lynett & Rogers, 2000) to investigate the biasing effects of emotionally-laden material. The statement describes the process of a sexual assault from the victim's point of view.

Composite ratings were evaluated to assess the effects of the victim impact statement on the professionals' overall judgements regarding future risk. Following review of the victim impact statement, between-subjects analysis indicated that clinicians rated offenders as equally likely to commit acts of sexual and nonsexual violence ( $M_{\text{nonsexual}} = 42.38$ ,  $SD = 23.28$  vs  $M_{\text{sexual}} = 42.46$ ,  $SD = 21.62$ ,  $F [1, 79] = .00$ ,  $p > .05$ ). Estimates of sexual and nonsexual violence remained unrelated to accuracy ( $F [1, 79] = .05$ ,  $p > .05$ ).

*Magnitude of Change.* Repeated-measures analysis indicated that clinicians rated offenders as both more likely to commit acts of sexual violence and nonsexual violence

following review of the victim impact statement (see Table 23). Because the magnitude of change was unrelated to accuracy ( $X^2_{\text{change}} = .58, p > .05$ ), the effects produced by review of the victim impact statement can be construed as a biasing effect. The emotionally laden information increased predictions of both sexual and nonsexual violence, although the effect was slightly larger in biasing predictions of sexual violence (see Table 23).

Table 23

*Biasing Effects of Victim Impact Statement on Clinicians' Composite Predictions of Sexual and Nonsexual Violence (Means/Standard Deviations)*

	Pre-Victim Impact	Post Victim Impact	<i>F</i>	<i>d</i>
Nonsexual	39.32 (21.58)	42.38 (23.28)	5.83*	.14
Sexual	37.25 (18.05)	42.46 (21.62)	12.81***	.26

\*  $p < .05$

\*\*\*  $p < .001$

Results have already demonstrated that clinical decision-making methods are unrelated to accuracy. The biasing effect found with the victim impact statement provides the opportunity to inspect each decision making method's susceptibility to bias. A testable hypothesis is that certain assessment methods make the professional more susceptible to bias. Alternatively, the use of a specific assessment method may render the professional immune to the biasing effects of victim impact statements.

Order of assessment method yielded no significant difference in magnitude of change in predictions of sexual or nonsexual violence. However, a significant interaction between change in estimates of violence and primary assessment method suggested that utilizing unstructured

methods first, followed by either actuarial or structured clinical judgment resulted in more susceptibility to bias than did either of the other assessment methods (see Table 24). In terms of nonsexual violence predictions, the use of unstructured clinical judgment resulted in greater bias than either structured clinical judgment or actuarial ( $M_{\text{difference}} = 9.24, d = .83$ ). For sexual violence, unstructured clinical judgment also resulted in more bias than the use of actuarial or structured clinical data ( $M_{\text{difference}} = 8.99, d = .72$ ). Only individuals beginning their assessments with unstructured clinical data changed their risk estimates significantly following review of the victim impact statement, indicating that assessments anchored with either actuarial data or structured clinical data are less vulnerable to the biasing effects.

Table 24

*Magnitude of Violence Prediction Bias Following Review of a Victim Impact Statement by Primary Assessment Method*

	Unstructured M (SD)	Actuarial M (SD)	Structured M (SD)	<i>F</i>	d1	d2
<b>Magnitude of Change</b>						
Sexual Violence	11.65 (16.87)	2.42 (8.94)	2.85 (11.25)	4.27*	.70	.64
Nonsexual Violence	9.68 (16.80)	.80 (9.31)	.16 (6.45)	5.70**	.67	.80

*Note:* d1 = Cohen's *d* measures the magnitude of change for unstructured clinical judgement compared to actuarial; d2 = Cohen's *d* measures the magnitude of change for unstructured clinical judgement compared to structured clinical judgment.

\*  $p < .05$ ; \*\*  $p < .01$

Clinicians did not alter their support for SVP commitment following review of the victim impact statement, nor did victim impact statements produce changes in perceptions of the offender to control his behavior or perceived treatability of the offender.

The available clinician characteristics were not related to their susceptibility to bias. A stepwise multiple regression was performed to test the hypothesis that clinician characteristics, rather than decision making method, was responsible for the resulting bias,. No available demographic data entered the equation. Years of professional practice, number of risk assessments conducted, gender, and percentage of practice devoted to forensics all failed to predict significantly the extent of emotionally-based bias.

*Over/Under Prediction of Sexual Violence.* Within the framework of sexually violent predator evaluations, inaccurate judgements may fall into one of two categories: overprediction or underprediction of sexual violence. A final supplementary analysis inspected factors contributing to inaccurate predictions. Two separate discriminant functions were performed to investigate the possibility that different clinician characteristics influence either the tendency to over- or underpredict sexual violence.

Both discriminant analyses suggested that only the variable “percentage of forensic practice” discriminated between accurate and inaccurate clinicians. Number of risk assessments conducted, years practicing at the doctoral level, gender, and current practice in a state with an SVP commitment statute all failed to enter. Interestingly, both under and overpredictions of sexual violence indicated that they devoted a significantly greater percentage of their practice to forensic issues than did accurate clinicians (see Table 25).

Table 25

*Separate Discriminant Function Analyses: Predicting Over- and Underpredictors of Sexual Violence Based on Clinician Characteristics*

	% of forensic practice	Classification Rate
Accurate	63.38 (40.11)	
Overpredictors	88.67 (27.80)	61.0%
Underpredictors	90.53 (19.29)	61.0%

## CHAPTER 4

### DISCUSSION

A critical issue for public policy and professional practice is whether Sexually Violent Predator (SVP) commitment is an empirically-validated determination. Morse (1998) and Schopp (1998) have questioned its validity as a legal construct, noting that it blurs the line between criminal and civil commitment. Other psychologists (Melton, Petrila, Poythress, & Slobogin, 1997) caution against offering forensic opinions, particularly about dangerousness (p. 17), stating these “ultimate” opinions regarding legal constructs are moral or legal judgments not to be made by a psychologist. In contrast, Rogers and Ewing (1989) cogently argued that forensic opinions are not moral or legal in nature. Clinicians involved in forensically relevant evaluations attempt to apply the legal standard as it is written (Heilbrun, Rogers, & Otto, in press; Rogers & Ewing, 1989). Importantly, they do not evaluate the morality or utility of the law. The central issue for the current discussion is the validity of current knowledge and methodology in making the forensically-relevant determinations.

To equitably and ethically apply sexually violent predator (SVP) statutes, it is imperative that forensic clinicians identify offenders likely to commit future acts of *sexual* violence, rather than any form of violence. Specifically, this study examined professionals’ and paraprofessionals’ abilities to discriminate between different types of violent behavior. With the increase in sexually violent predator statutes, the implications of being labeled a sexually violent offender, rather than simply a dangerous offender are considerable. Importantly, the constitution forbids ex-post facto legislation, meaning that an individual cannot be committed as an SVP simply because he has committed a sexually violent crime in his past. Past behavior can,

however, be admitted as evidence in predicting future violent behavior. Therefore, the onus is on the courts, aided by mental health professionals, to determine *which* offenders with sexually violent crimes in their history are likely to be sexually violent again. As Janus and Nudell (2000) argue, “The testimony of clinical psychologists is given great weight in sex offender commitment proceedings and the accuracy and reliability of their testimony is of paramount importance to the correct resolution of the case” (p. 13). Rendering accurate and reliable testimony is a complex process with a myriad of nuances, some of which were tested here in order to examine their effects.

The discussion that follows reviews implications for sexually violent predator evaluations, including both policy and professional practice issues. The discussion then turns to future directions in practice and research and provides preliminary recommendations regarding the process of assessment.

### Implications for Sexually Violent Predator Evaluations

#### *Policy Issues*

Morse (1978) and Grisso (1986; see also Heilbrun, 2001) have proposed models for use in psycholegal mental health assessments. Morse’s (1978) assessment model for civil commitment was based on his observation that civil mental health laws have three primary components: (a) the existence of a mental disorder, (b) the functional abilities related to the legal task (e.g., ability to care for oneself), and (c) the strength of the causal connection between the mental disorder and the functional disabilities (Bank, 2002; Morse, 1978). The application of Morse’s model to sexually violent predator evaluations is clear. Sexually violent predator statutes share four common elements: (a) a past act of sexually harmful conduct, (b) a finding of risk of future sexually harmful conduct, (c) a current mental disorder or abnormality, and (d)



some connection between the mental abnormality and the danger, namely, the abnormality causes some impairment in the offender's ability to control his behavior. Applying Morse's model, the clinician must address (a) the presence of a mental disorder, (b) the functional act of committing sexual violence, and (c) the mental disorder *causes* the commission of the sexual violence.

Grisso's (1986) model was developed specifically for forensic evaluation of legal competencies. It consists of six aspects Grisso believed are shared by legal competencies: functional, contextual, causal, interactive, judgmental, and dispositional (see also Heilbrun, 2001, p. 84). Although more complicated than Morse's (1978), Grisso's (1986) model shares at least two common elements with the model proposed by Morse: functional and causal. Both models account for the importance of functional abilities related to the legal construct as well as the causal element (e. g., a mental disorder) responsible for the functional impairment. As such, each model provides guidance to the forensic clinician in collecting relevant data, answering the legally-relevant questions, and organizing and presenting conclusions.

Clinicians face a formidable task when they agree to conduct an assessment and offer testimony in a sexually violent predator evaluation. Similar to other referral questions, such as competency to stand trial and criminal responsibility, the expert is expected to understand the legal standard and apply it appropriately in his or her psychological evaluation and eventual testimony. The Specialty Guidelines for Forensic Psychologists (1991) clearly state that clinicians must be "prepared to explain the relationship between their expert testimony and the legal issues" (p. 665). The current data provide an opportunity to inspect the degree to which clinicians understood and applied the relevant legal standard in making recommendations regarding sexually violent predator commitments.

Significant predictors of SVP commitment recommendations were the following: (a) likelihood of future sexual violence, (b) presence of a mental abnormality or personality disorder, and (c) offender's ability to control his behavior. As expected, predictions of future sexual violence were most highly related ( $r = .54, p < .001$ ) to strength of SVP recommendation. Although 75 (70.8%) of the doctoral-level clinicians believed evidence of a mental disorder or abnormality was present, it was not significantly related ( $r = .12, p > .05$ ) to their ratings of the offender's ability to control his behavior. In other words, forensic experts do not appear to recognize the importance of volitionality in the legal definition of mental disorder.

Proper application of the SVP statutes requires an understanding of mental abnormality and ability to control behavior within a *legal* definition. Legal scholars (e.g., Janus, 1998; Schopp, 1998) articulated the distinction between *legal* and *clinical* mental disorders. Legally, a disorder requires an impairment in volitional control (Schopp, 1998). The simple presence of a clinical disorder and lack of control is not sufficient; the key element is that the disorder *causes* the lack of control. Clinically, mental disorders require no such connection. The DSM-IV (APA, 2000) clearly states that a clinical diagnosis "is not sufficient to establish the existence for legal purposes of a mental disorder, mental disability, mental disease or mental defect" (p. xxiii). The DSM-IV continues by reminding clinicians that "having the diagnosis in itself does not demonstrate that a particular individual is (or was) unable to control his or her behavior at a particular time" (p. xxiii).

*Hendricks* (1997) and *Crane* (2002) provided guidelines for the required impairment in volitionality. However, the Court purposefully avoided operationalizing the construct by suggesting that "safeguards of human liberty in the area of mental illness and the law are not always best enforced through precise bright-line rules" (*Crane*, 2002, p. 868). *Crane* clarified

that the statutes do not require a total lack of control, only a “serious difficulty” in controlling behavior (p. 868). In the absence of bright-line criteria, clinicians are faced with the challenging task of not only defining volitionality for themselves, but assessing it in dimensional terms. Investigators (e.g., Rachlin, Halpern, & Portnow, 1984) have noted the complexity the volitionality requirement adds to psycholegal assessments, including the complexity involved in assessing volitionality in its dimensional form. The clinician must struggle with the definition of “serious difficulty,” distinguishing it from “some” difficulty or “moderate” difficulty.

Volitional impairment is not novel to SVP statutes. As an integral component of a legal mental disorder, forensic mental health experts are required to assess volitionality as part of several legal standards (e.g., criminal responsibility). Drawing on the clinical assessment of criminal responsibility provides a framework for discussing clinicians’ experience with assessing volitional impairment in a legal context.

The lack of volitional control is a pivotal component of legal standards for insanity. For example, the American Law Institute (ALI, 1962) standard for legal insanity requires that, as the result of a mental disease or defect, the defendant either (a) lacks an appreciation of his criminality or (b) is unable to conform his conduct to the law. Shades of the ALI standard are undoubtedly seen in the modern SVP laws that require impaired volitionality *as a direct result of a mental disease or defect*. At least two approaches to the assessment of a defendant’s volitionality have been advanced (*Hendricks*, 1997; Rogers & Shuman, 2000). For the ease of discussion these models will be referred to as the “self-report” model and the “behavioral inference” model.

*Self-Reported Volitionality.* In *Hendricks* (1997), the Supreme Court appeared to rely heavily on Leroy Hendricks’ own testimony that he was “unable to control” his molestation of

children. His admission effectively allowed the Court to avoid questions regarding *the degree* of impairment required by the law. The Hendricks case provides the first avenue for gauging volitional impairment, namely, the offender's own admission. Clinicians often gather information from patients, clients, and defendants via self-report methods to assess symptoms of psychopathology and level of functioning, whether the methods used are unstructured clinical interviews, multiscale inventories, or the more sophisticated structured interviews. Despite the clinical tradition, two fundamental problems in self-reported impairment in volitionality cause this approach to be untenable in SVP evaluations in that it requires the defendant to: (a) possess a rather sophisticated degree of insight and (b) negotiate the "right" answer depending on the legal question.

Research suggests that sex offenders are unreliable in their self-reports (Langevin, 1988; Rogers & Dickey, 1991; Sewell & Salekin, 1997). Most notably, sex offenders are characterized by their denial, minimization, or externalization of blame (Kennedy & Grubin, 1992; Langevin, 1988). Kennedy and Grubin (1992) identified a number of different denial patterns common among sex offenders. Importantly, two of the four patterns identified were characterized by externalization of responsibility. These offenders acknowledged the offense, but attributed the cause of their behavior to an external force out of their immediate control (see also Sewell & Salekin, 1997). These studies suggest that in terms of sex offender evaluation, a defendant's "admission" that he had no control over his<sup>8</sup> behavior is a relatively common defense, questioning the diagnostic and legal relevance of this admission. A reasonable conclusion is that a defendant's externalizing the cause of behavior to forces beyond his control is so common as not to provide specific information.

The second flaw in asking a defendant to judge his level of volitional impairment is that it puts the offender in a very untenable position. Rogers and Dickey (1991) reviewed three competing models of denial and minimization in sex offenders: the pathogenic, criminogenic and adaptational models. The adaptational model posits a kind of cost-benefit analysis in a defendant's chosen presentation. Perceiving himself in an adversarial situation (e.g., a criminal trial or civil commitment hearing), an offender will adopt a response style that he deems most likely to facilitate a positive outcome for himself (Rogers, 1990; Rogers & Dickey, 1991). For example, a criminal defendant may perceive an insanity acquittal as more palatable than an extended prison term, and therefore, attempt to feign psychosis. Applying the adaptational model to sex offenders clarifies the difficulty with which the offender is faced, as the following example illustrates. A defendant may present his "inability to control his behavior" as a mitigating circumstance at trial because it argues against criminal intent. The resulting trial strategy may involve pleas of mental disorder and volitional impairment in an effort to avoid or minimize punishment. The convicted offender is sentenced to a prison term and eventually identified as a potential candidate for sexually violent predator commitment. The offender is then faced with the dilemma of changing strategies for this civil trial, and attempt to show that he does not suffer from mental illness or volitional impairments. Ironically, in the civil trial, in order to secure freedom from detainment, he must demonstrate that he is in full control of his mental faculties and made conscious choices to repeatedly and heinously offend against each of his victims. Furthermore, the inconsistencies in his account may be used to impeach his credibility in reference to volitional impairment. In addition to pointing out the paradoxical implications of

---

<sup>8</sup> Research suggests that most sex offenders are male (see e.g., Texas Commission on Law Enforcement, 2002). For the ease of discussion, the masculine pronoun will be used when referring to a sex offender.

the commitment law, this example illustrates the precarious position the offender is faced with when asked to disclose his level of control.

*Behavioral Inference.* A second model for assessing volitional impairments developed from clinical assessments of criminal responsibility. Arguably, all sex offenses are instances of uncontrolled behavior in that the offender failed to conform his behavior to legal and societal standards. Of key importance in SVP applications, similar to insanity evaluations, is proving that the behavior resulted from an uncontrollable impulse, rather than an impulse that was not controlled. As a clear, if extreme example, Rogers and Shuman (2000) discussed the “police at the elbow” test. If the defendant would have committed the crime, even in the presence of law enforcement, sufficient impairment in volitionality is demonstrated. The conclusion begs the question, “Is profoundly poor judgment equal to a lack of control?” Fortunately for the clinician conducting the SVP evaluation, most offenders involved in these cases have long histories of offending, allowing for a close examination of the pattern of crimes. A pattern suggesting similar crimes, committed impulsively, over a period a time, without regard for punishment *may* suggest impaired volitional control, rather than an instance of poor judgment. Experience from insanity evaluations (Rogers & Shuman, 2000) and SVP evaluations suggest this distinction is primarily one of clinical judgment.

In applying the SVP standard, experts’ ratings of volitional impairment were significantly correlated with indicated support of SVP commitment ( $r = .37, p < .01$ ). Given the importance of volitional impairment to the commitment standard, however, this correlation is rather modest. Reassuringly, very few ( $n = 3$ ) clinicians recommended SVP commitment for offenders who had no loss of control. Importantly, these clinicians were not very likely ( $M = 30.0\%$  likelihood) to support an SVP commitment. For forensic experts, the lack of volitional control appeared less

important in making their SVP recommendations than ratings of future sexual violence ( $r = .54$ ,  $p < .001$ ) or the presence of a mental disorder ( $r = .52$ ,  $p < .001$ ). Importantly, ratings of volitional impairment were unrelated to the presence of a mental disorder ( $r = .12$ ,  $p > .05$ ), indicating that clinicians apparently failed to conform their ratings to the legal definition of mental disorder (*Crane, 2002; Hendricks, 1997; Janus, 1998; Schopp, 1998*).

Clinicians on average saw offenders as having “some” inability to control behavior ( $M = 4.76$ ). In the absence of a relationship between mental disorder and volitionality, it is unclear how clinicians were assessing lack of control. Presented with limited information on which to base their ratings, clinicians appeared to use their clinical judgment to make ratings. Interestingly, ratings of the offender’s ability to control his behavior did not differ depending on assessment method. Furthermore, no variables were found that significantly predicted volitionality ratings, including offender’s outcome status, presence of a mental disorder, and likelihood ratings of violence. Clinician characteristics such as years practicing, risk assessments conducted, or gender also failed to predict ratings of volitional control.

A possible avenue for the assessment of volitional impairment is in the related clinical construct of impulsivity. Rachlin et al. (1984) argued that volitional impairment occurs when an impulse becomes uncontrollable rather than uncontrolled. They articulated their belief that when behavior is ego-dystonic, as occurs with compulsive acts, volitional impairment is evident. Therefore, it is important to assess impulsive behavior as a component of volitional control.

In addition, Prentky, Knight, Lee, and Cerce (1995) examined impulsivity among rapists and found that rapists high in impulsivity were at a three times greater risk of committing a future sexual offense than rapists low in impulsivity. Their results suggest that impulsivity is an important component in assessing future risk. As impulsivity is conceptually related to

volitionality, the SVP standards' lack of control requirement is clinically as well as legally pertinent.

### *Professional Practice Issues*

The current design provided a very rigorous test of forensic experts' predictive abilities in differentiating types of violence as required by SVP standards. Previous investigations (see e.g., Gardner et al., 1996; Lidz et al., 1993; Mossman, 1994; Quinsey et al., 1998), often address violence as a general construct. In contrast, this investigation required experts to make near-neighbor comparisons, distinguishing sexual violence, a necessary condition of SVP findings, from other forms of violence that are insufficient for an SVP determination. In keeping with the risk assessment literature to date, however, the discussion of professional practice issues begins with a brief discussion of experts' ability to accurately predict general violence.

Predictive accuracy achieved by clinicians in this study exceeded earlier estimates (Gardner et al., 1996; Lidz et al., 1993; Monahan, 1981). Over 80% of the experts accurately predicted future violence in the current group of offenders. Importantly, 69 of the 107 (64.5%) doctoral-level clinicians indicated a greater than 50% likelihood (i.e., more likely than not) of future sexual or nonsexual violence offender. As previously indicated, estimates of at least 50% represent a logical lower-bound threshold for predictions. In terms of general violence prediction, the experts in the current study achieved considerably greater-than-chance accuracy.

*Professional Training.* Research suggests mixed results regarding the effects of professional training on predictions of violence. Early work by Quinsey and Ambtman (1979) indicated that professionals' and laypersons' predictions do not differ. More recent work by Lidz et al. (1992) indicated that experienced clinicians tend to be more conservative in their judgments. Ziskin (1995) undertook a review of three decades of research regarding experience



and accuracy and concluded that “there is a substantial body of scientific and professional literature indicating a lack of relationship between experience and accuracy of psychological assessment” (p. 523). For example, among the more recent studies, Ziskin related the findings of Walters, White, and Greene (1988) who demonstrated that experts were no more accurate than graduate students in identifying malingering on the MMPI.

Consistent with the majority of past research (see also Garb, 1998), surprising little difference was found between graduate students’ and professionals’ ratings in the current study. Doctoral-level clinicians were virtually identical to graduate students in overall hit rates (50.0% versus 45.9%). Clinicians and graduate students did not differ in their confidence levels or indicated support of SVP commitment for the offender. Extrapolating from the results of Lidz et al. (1992), graduate students were expected to provide higher estimates of future violence potential and indicate greater support for SVP commitment. These hypotheses were not borne out by the data.

Graduate students were more optimistic than doctoral-level clinicians regarding the amenability of the offender to treatment ( $d = .30$ ). Reasons for this difference are unclear. Perhaps graduate students are unaware of high recidivism rates even for “treated” sex offenders (see e.g., Seto & Barbaree, 1999) Perhaps doctoral-level clinicians were more jaded by the clinical lore dictating that personality pathology, such as that exhibited by offenders in this study (i.e., antisocial and psychopathic traits), is immutable.

Of paramount importance in conducting SVP evaluations is the proper application of the relevant legal standard (Heilbrun, 2001; Heilbrun, Rogers, & Otto, in press) . Graduate students did not understand and apply the standard as well as professionals. In contrast to professionals, graduate students did not rely significantly on the mental abnormality requirement in making

their recommendations. A stepwise multiple regression indicated that graduate students' support of SVP commitment was predicted by their ratings of sexual violence potential, ability to control dangerous behavior and their ratings of nonsexual violence. By basing SVP decisions on both violence and sexual violence potential, graduate students demonstrated less ability to accurately apply the standard. Furthermore, the finding may suggest a bias toward "locking up" violent offenders, regardless of their specific risk of sexual violence and mentally disordered status—two of the three required prongs of civil commitment.

*The Effects of Confidence.* Similar to previous research investigating confidence about professional judgments (Oskamp, 1965; Granhag & Stroemwall, 1999; Suarez, Chalk, Russell, Kim, O'Brien, et al., 2001), experts became more confident in their judgments with the addition of data. Oskamp (1965) asked participants to render conclusions regarding a subject's personality characteristics at four different time periods. Between ratings, participants reviewed additional data concerning the subject. At each rating, the participants' confidence in their conclusions increased but their accuracy did not. Similar results were found in deception studies (Granhag & Stroemwall, 1999) and studies of diagnostic accuracy (Suarez et al., 2001). Despite the lack of incremental validity, additional information appeared to provide false confidence in resultant conclusions.

Research suggests that confident judgments are afforded more weight by juries than judgments offered with lower levels of certainty (Penrod & Cutler, 1995), although much of this research investigated eyewitness confidence (Fox & Walters, 1987; Wells, Lindsay, & Tousignant, 1980). The available data on the relationship between confidence and accuracy are more limited (McNiel, Sandberg, & Binder, 1998). Contrary to expectations, McNiel et al. found that clinicians' confidence was strongly related to accuracy. The authors investigated physicians'

ability to predict civil psychiatric patients' short-term risk for violence. Upon admission, physicians were asked to rate each patient's risk for committing violent assault within the ensuing seven days. In addition, physicians rated their confidence level for each prediction. Clinicians with a high degree of confidence (i.e., 75-100%) correctly classified most dangerous patients (96%; PPP = .75) compared to clinicians with low levels of confidence (< 50%) who correctly classified relatively few (PPP = .36) dangerous patients.

In stark contrast to the McNiel et al. study, clinicians in the present investigation with low levels of confidence (< 50%) achieved *greater* accuracy than clinicians with high levels of confidence (> 75%). Surprisingly, doctoral-level clinicians with low levels of confidence achieved accurate hit rates nearly 50% greater than highly confident clinicians (hit rates = 60.7 versus 41.7, respectively). Sample differences are likely partially responsible for disparate results between McNiel et al. and the current study. Steadman and Coccozza (1974) and Monahan (1981) aptly demonstrated that long-term predictions of violence are difficult to render with accuracy. More recent research (Lidz et al., 1993; McNiel & Binder, 1991) has demonstrated that short-term predictions of violence are more accurate. Clinicians in the present study were asked to make long-term predictions, a more difficult task than the predictions made by the McNiel et al.'s sample. In addition, several civil psychiatric patients present relatively little risk of violence as they lack violent histories, substance abuse, or other risk factors, allowing for non-predictions to have a high level of accuracy (NPP = .98 for highly confident clinicians). In contrast, the current study presented offenders with high likelihoods of violent behavior, with none failing to be violent at a long-term follow-up.

At minimum, the conflicting results of the two studies suggest that the relationship between accuracy and confidence for violence predictions is unknown. The counter-intuitive

finding in the present study that judgments rendered with low confidence were more accurate casts doubt on Litwack's (2001) assertion that confident judgments should be afforded more weight than less confident judgements or those based on actuarial techniques.

#### Recommendations for Conducting SVP Evaluations

Psychologists choosing to conduct SVP evaluations are given few guidelines on how to conduct their assessments. Clearly, they must address the legal standard and render certain judgments. However, with an evolving standard (*Kansas v. Hendricks*, 1997; *Kansas v. Crane*, 2002) and imperfect psychological measures, professionals are faced with an arduous task. This section offers tentative recommendations for conducting SVP evaluations based on past research along with the current results.

The current study's results suggest that no assessment method is more accurate in making predictions of future sexual violence. Making very fine-grained distinctions between violence and sexual violence, experts made accurate predictions approximately half the time. At least three possible conclusions can be drawn. First, because each assessment method is as good as another, clinicians may use any one of the assessment methods. Second, they all are inferior and should be abandoned. Third, a *multimethod* approach to risk assessment represents the most prudent by combining measures for conducting SVP evaluations (Heilbrun, 2001; but see Ziskin, 1995).

At a minimum, a comprehensive risk assessment in the context of a sexually violent predator evaluation must accomplish three goals: (a) determine the likelihood of sexually violent behavior in the future, (b) establish the presence of a *legal* mental disorder, and (c) establish an impairment in volitional control (see e.g., *Crane*, 2002; *Hendricks*, 1997; Janus, 1998; Schopp, 1998). A multimethod approach to risk assessment is best able to achieve these goals, addressing

the legally relevant questions posed to the forensic mental health expert (see e.g., Heilbrun et al., in press; Rogers & Shuman, 2000) The following section discusses the utility of the multimethod approach in addressing each of these elements.

### *Determining Future Violence*

The accurate prediction of future sexual violence is of central importance in conducting SVP evaluations. Specialized assessment measures, such as the Static-99 (Hanson & Thornton, 1999) and the Sexual Violence Risk-20 (SVR-20; Boer et al., 1997; © Mental Health Policy and Law Institute, Burnaby, Canada), have been developed specifically to predict future sex offenses. Despite their tailored purpose, their superiority in aiding clinical decision makers was not demonstrated in the current investigation. Because unstructured clinical judgments tended to underestimate the risk of potential violence, the actuarial methods appeared to slightly outperform the sole use of unstructured clinical judgment.

Methodological advances (Borum, 1996; Harris et al., 1993; Hart, 1998; Monahan et al., 2001; Webster et al., 1997) have improved the science of risk assessment; however, predicting future violence continues to be challenging. Beyond the use of specialized risk assessment guides (e.g., Violence Risk Appraisal Guide [VRAG; © American Psychological Association, Washington, DC], Historical Clinical Risk –20 [HCR-20; © Mental Health Policy and Law Institute, Burnaby, Canada], etc.), comprehensive assessments must consider the “totality of circumstances” (Hart, 1998, p. 126), taking into consideration static and dynamic variables as well as environmental variables. SVP laws may be in their infancy; however, forensic mental health assessment is not. Important lessons are learned from other areas of forensic practice, such as forensic assessment of criminal responsibility. For example, Rogers and Shuman (2000) stated that in many cases, “corroborative interviews should be established as the *standard of*

*practice* for insanity evaluations” (p. 180, emphasis in the original). As outlined earlier, the legal standards imposed for insanity evaluations can reasonably be extrapolated to SVP evaluations. Rogers and Shuman further recommended interviewing friends and family as well as reviewing collateral data sources such as police reports and laboratory test findings. In addition, they assert, “Victims ... have a unique perspective on the defendant’s criminal behavior” regarding the offender’s behavior (p. 180).

A completely different stance is advocated by Quinsey et al. (1998) in suggesting that actuarial techniques should be used in isolation because other sources of data (i.e., clinical data) contaminates a professional’s resultant decision. In light of this conflicting advice, how should the conscientious clinician approach an SVP evaluation?

Data presented here suggest the most prudent approach is a sequential multimethod model in which statistical data (e.g., actuarial or structured clinical guides) is used to anchor the assessment. The statistical data is then augmented by unstructured clinical methods such as interviews and review of collateral information (e.g., victim interviews or statements). The current study found no support for Quinsey et al.’s assertion regarding actuarial superiority. Nonetheless, current results do suggest that anchoring assessments with statistical data inoculates the professional against the potentially biasing effects that may occur during the course of a comprehensive assessment. For example, these data suggest that clinicians beginning with unstructured clinical judgment were most vulnerable to the biasing effects of the emotionally-laden information. However, those clinicians who utilized unstructured clinical judgment were best able to discriminate between sexually and non-sexually violent offenders. A reasonable conclusion is that combining the approaches (i.e., structured methods followed by unstructured) yields the maximum benefit afforded by both approaches.

### *Legal Mental Disorder*

The importance of the legal mental disorder requirement in SVP statutes cannot be understated as it is the fundamental principle upon which civil commitment laws rest (Morse, 1978; Janus, 1998). SVP statutes requiring a finding of a mental abnormality or personality disorder implicitly require the finding of a *legal* mental disorder. Regardless of the presence of a clinical mental disorder, a legal mental disorder must impair the offender's ability to control his behavior. Without the legal mental disorder requirement, nothing separates civil commitment from criminal punishment (Janus, 1998). As articulated in *Hendricks* (1997), the mental disorder requirement "adequately distinguishes Hendricks from other dangerous persons who are perhaps more properly dealt with exclusively through criminal proceedings" (p. 2081).

Importantly, clinicians relying solely on actuarial methods are unable to address fully the presence and severity of several mental disorders or abnormalities. Of the two actuarial methods used in the current study, one (i.e., Static-99) does not address the presence of mental abnormalities at all. The VRAG addresses it only in a cursory manner, allowing for ratings of the presence or absence of a personality disorder, schizophrenia, and psychopathy. Moreover, the VRAG (Harris et al., 1993; Quinsey et al., 1998) provides no guidance in assessing for the mental disorder nor requires that the clinician performs the assessment.

An additional conceptual difficulty in adopting the VRAG for use in SVP evaluations is that schizophrenia receives a negative rating on the VRAG, indicating that schizophrenic offenders are at a lower risk of violence than offenders free from psychosis. However, Rogers and Cavanaugh (1981) and Rogers and Shuman (2000) found that the presence of psychosis is most often related to volitional impairment in criminal responsibility assessments. The failure of the VRAG to assess the presence and severity of psychotic symptoms is inconsistent with the

demands of SVP evaluations. Clearly, actuarial measures provide inaccurate coverage of important areas of potential mental abnormality and do not allow for a causal finding between mental disorder and volitionality.

Structured clinical guides (i.e., HCR-20 and SVR-20) provide slightly better coverage of mental illness variables, allowing for ratings of “major mental illness” in addition to psychopathy and personality disorders. Further, the HCR-20 allows for ratings of “active mental illness.” Monahan (1992) articulated the importance of *active* symptoms of psychopathology in demonstrating the association between florid psychotic symptoms and violence. The HCR-20, but not the SVR-20, recognizes the importance of active symptoms of psychopathology in relation to violence. Neither measure, however, accounts for specific symptoms of psychopathology that may be related to violence, such as well-developed delusions (de Pauw & Szulecka, 1988) or behavioral dyscontrol often found in mania (Benjaminsen, Gotzsche, Norrie, Harder et al., 1996; Tardiff, 1998). In summary, these structured clinical guides also fail to account for symptoms of mental disorder and their relationship to volitional impairment.

Comprehensive clinical assessments are necessary to gauge the presence and severity of psychological symptoms (Rogers, 2001; Rogers & Shuman, 2000) that may be related to violence potential (see e.g., de Pauw & Szulecka, 1988) and volitional impairment (see e.g., discussion of threat/control-override symptoms in Swanson, Borum, Swartz, & Monahan, 1996). Ratings of mental illness or abnormality provided by actuarial and structured clinical guides provide categorical diagnostic data only. A much more detailed investigation of signs and symptoms is necessary for the clinician to understand the individual’s mental abnormality and assess its impact on his or her functioning. Comprehensive assessments of this kind require the



use of clinical interviews (unstructured, structured, or semi-structured) along with additional sources of data.

### *Volitional Impairment*

The final prong of SVP commitment statutes, that requiring volitional impairment, is integrally linked to the mental illness requirement. After *Hendricks*, the degree of volitional impairment required by law was left slightly obscured. However, the more recent *Crane* (2002) ruling addressed the point directly and reiterated that lack of control, of a “serious” but not total nature, is indeed an essential element in SVP commitments. No variables on either the actuarial methods or the structured clinical methods allow the clinician to draw conclusions regarding the volitionality of the offender’s behavior. Without the inclusion of clinical judgment (e.g., the blended approach), clinicians cannot address this important prong. To be fair, actuarial and other structured risk assessment measures were not designed for use with SVP evaluations. Nonetheless, the inability to assess volitionality is an important limitation in adopting their use for SVP assessments. Given our current methods, volitionality cannot be assessed without the use of clinical judgment rendered following interviews, record reviews, collateral information, etc.

Several key symptoms of psychopathology provide potentially important information in assessing volitional impairment. As reviewed in the previous section, violence and other forms of behavioral dyscontrol are related to mania, impulsivity, and some symptoms of psychosis. In addition, certain traumatic brain injuries may impair the affected individual’s ability to control his or her behavior (Brower & Price, 2001; McAllister, 1992). Impulsivity is empirically associated with violent behavior (Boer et al., 1997). As described earlier, impulsive sex offenders are at a much higher risk of sexual recidivism than non-impulsive offenders (Prentky et

al., 1995). Of the measures utilized in the current study, the HCR-20 specifically calls for ratings of impulsivity. Indirectly, both the SVR-20 and the VRAG account for impulsivity by their inclusion of the PCL-R. However, inspection of total scores on the PCL-R does not allow the clinician to independently consider the impulsivity of the offender in his or her risk evaluation.

The degree to which impulse is synonymous with volitional impairment is a controversial question (Miller, 1992; Rachlin et al., 1984). A decision must be made regarding the point at which an impulse becomes an uncontrollable impulse. Clinicians turning to the DSM-IV (2000) for clarification regarding impulse control disorders are provided little guidance in making this distinction. The DSM-IV states the essential feature of impulse control disorders is “the failure to resist an impulse, drive, or temptation to perform an act that is harmful to the person or to others” (p. 609). It offers no guidance in assessing whether the individual *could have* resisted the impulse. Miller (1992) argued that volitional standards should include a measure of impulse control to avoid simple descriptions rather than explanations of behavior.

A second important consideration in volitional impairment is the type and severity of psychotic symptoms. Despite the lower risk value received by schizophrenics on the VRAG (Harris et al., 1993), research has found that certain psychotic symptoms are related to violent behavior (see e.g., de Pauw & Szulecka, 1988; Rogers, Gillis, Turner, & Frise-Smith, 1990) and ability to control that behavior (Link & Stueve, 1994; Swanson et al., 1996). Specifically, Link and Stueve (1994) and Swanson et al. (1986) identified the presence of certain psychotic symptoms that were particularly associated with violent behavior. Symptoms such as hallucinations or delusions that threatened a patient’s sense of safety led to violent outbursts, presumably of a self-protective nature. These “threat/control-override” symptoms impaired the

patient's capacity to control his violent behavior. Additionally, Rogers et al. (1990) found that the presence of command hallucinations may have a profound impact on a patient's behavior.

As the preceding analysis outlines, current risk assessment methodology is insufficient to meet the specific requirements of SVP legislation. Importantly, the legal standard demands more than accurate predictions of risk. In the absence of existing methods that meet the legal need, a multimethod approach to SVP evaluations must be adopted. Actuarial and structured clinical guides demonstrated a slight advantage over unstructured clinical judgment in predicting future violent behavior. However, important components of SVP evaluations, such as the presence of a legal mental disorder leading to volitional impairment, can best be addressed via the traditional clinical approach. Only by combining the strengths of the structured methods with the comprehensiveness of clinical methods can the standard be sufficiently addressed.

Considerations in future work with risk assessment, particularly for legally relevant purposes, must include assessing for the presence of a legal mental disorder as well as volitional impairments. Lessons from other forensic mental health assessments, such as competency to stand trial and criminal responsibility suggest that specialized assessment methods will be developed. A useful direction for researchers is to develop guides to help clinicians address the legal standard similar to the guides in use for assessing competency and insanity. An important component of an SVP evaluation guide will surely be guidelines for identifying and assessing volitionality in chronic offenders.

#### Next Steps Toward Validating an SVP Measure

An important component in improving methodology is developing empirically validated and forensically-relevant tests. A fundamental principle in classical test theory is that no test is valid in all purposes or situations (American Psychological Association [APA], 1985). The

validity of an interpretation of a test must be demonstrated with the population for whom the test will be used (APA, 1985; Foster & Cone, 1995). A reasonable extrapolation requires that any SVP evaluation method be (a) validated on individuals likely to be candidates for SVP commitment, and (b) validated for the express purpose of addressing the relevant legal standard. Of interest, each of the actuarial and structured clinical guides utilized in the current study were developed in Canada, where laws governing dangerous offenders (Canadian Criminal Code, 1987) have a broader scope than the U.S.'s SVP legislation. Additionally, the range of violent behaviors included in the original validation is much too broad to be helpful in making SVP commitment decisions (Boer et al. 1997; Harris et al., 1993; Litwack, 2001). A violent outcome category for test validation purposes that equates a non-contact assault with a homicide (VRAG; Harris et al., 1993) or obscene phone calls with violent rape (SVR-20; Boer et al., 1997) does not allow for valid interpretations to be made regarding the degree of violence needed for involuntary commitment, particularly highly specific SVP commitments (Litwack, 2001).

Heilbrun, Rogers, and Otto (in press) emphasized the importance of validating forensic assessment instruments on relevant populations. To this end, an SVP "test" must be validated on sex offenders, particularly sex offenders who commit acts of sexual violence that warrant commitment. This requirement presents some difficulty as the degree of sexual violence necessary has not been operationalized (see e.g., Community Protection Act of WA, 1990; *Hendricks*, 1997). Nonetheless, the lack of operationalization is not insurmountable. A priori decisions can be made regarding the types of sexual violence *unlikely* to meet the legal standard, such as exhibitionism, distribution of pornography, and sexual threats.<sup>9</sup> A potentially fruitful option for establishing relevant criteria is involving judges in the operationalization of sexual

violence. Their definition of sexual violence warranting SVP commitments is valuable information to include in test validation. In addition, an effective measure must also allow for prediction of *degree* of violence. To be maximally useful in the forensic context, the measure must distinguish individuals at risk for violent rape from individuals engaging in frotteurism.

An appropriate population for SVP test validation, therefore, is sex offenders with a range of offenses. Distinguishing sex offenders from non-sex offenders, while potentially important in other contexts, is less of a concern here. Individuals with no sex offenses do not become candidates for SVP commitment, and therefore, are not part of the target population for validation purposes (APA, 1985). Important variables to consider are those items that accurately discriminate sex offenders at risk for various levels of violence. Ideally, the test would have both high Negative Predictive Power, with low scores indicating risk of less violent acts, and high Positive Predictive Power, with high scores indicating risk of extremely violent acts.

Heilbrun et al. (in press) also articulated the need for forensic assessment instruments to meet the relevant legal standard. A test, such as one described above, meets a single prong of current SVP statutes (i.e., likelihood of future sexual violence; *Hendricks, 1997; Janus, 1998*). To be forensically relevant, the test must also assess the presence of a mental illness or abnormality and the degree to which the mental abnormality impairs the offender's ability to control his behavior (see e.g., *Crane, 2002*). The earlier discussion provided suggestions for assessing these constructs, including the assessment of mania, certain psychotic symptoms, and impulsivity. Few empirical studies have addressed volitionality (see e. g., *Howard & Conway, 1986*) or its relationship to impulsivity (*Rachlin et al., 1984*). Conceptually, impulsivity is related

---

<sup>9</sup> These examples are taken from Boer et al., 1997. In validating the SVR-20, each of these acts was considered acts of sexual violence.

to volitional impairment (i.e., an uncontrollable impulse) and may provide avenues for its assessment.

Recently, Webster and Jackson (1997) published an entire volume on the topic of impulsivity. Attesting to the importance of impulsivity's relationship to violent behavior, nearly half is devoted to violence. In this volume, Harris and Rice (1997) explored the construct of impulsivity in mentally disordered offenders. They acknowledge two definitions of impulsivity roughly corresponding to irresistible impulse and impulses that are not resisted (p. 367-368). No attempt is made to disentangle the two.

Distinguishing between *lack* of control and *inability* to control is a complex task. Currently, clinicians and researchers appear to be using the construct of impulsivity to subsume both these definitions. In the absence of empirical approaches to volitional assessment, forensic mental health professionals rely on clinical guidelines to assess this construct. For example, Rogers and Shuman (2000) cited the "police at the elbow" criterion as an extreme example of volitional impairment. Valid approaches to volitional assessment are clearly needed. A potentially fruitful line of research is the expansion of experimental paradigms such as those utilized by Newman and others to investigate impulsivity and behavioral inhibition in psychopaths (see e.g., Blaszczynski, Steel & McConaghy, 1997; Newman, Wallace, Schmitt, & Arnett, 1997; Schmitt, Brinkley, & Newman, 1999). By expanding this research to mentally disorder offenders in general, and sex offenders in particular, important information can be learned regarding these offenders' *ability* to control their impulses.

In proposing a model for a forensic assessment instrument specifically for use with SVP evaluations, it is clear that much exciting work remains in improving our methodology.

Additional research is needed in the prediction of sexual violence with instruments being

developed specifically on sex offenders that are able to predict extreme acts of sexual violence. Furthermore, little empirical evidence exists regarding the assessment of volitionality. Without this requisite knowledge, SVP evaluations fail to meet the demands of the law. Importantly, they also may fail to meet the demands of *Daubert* (*Daubert v. Merrell Dow Pharmaceuticals*, 1993), requiring that scientific techniques be testable, generally accepted, and possess a known error rate.

#### Limitations of the Current Study and Directions for Future Research

The current study was designed to provide a rigorous examination of forensic experts' ability to predict highly specific dangerous behavior and apply a relevant legal standard based on their judgments. The ecological validity of the study was bolstered by its inclusion of risk assessment data on six actual criminal defendants. In maximizing external validity, design and methodological decisions that were made may have affected internal validity.

An important component of internal validity in experimental design is the inclusion of a control or comparison group (Kazdin, 1998) allowing for significant findings to be attributed to the independent variables. In the current quasi-experimental design no comparison group of non-recidivators was included. Instead, the current design focused on discriminability between types of violence. Nonsexually violent offenders constituted the comparison group for the majority of the research questions. However, the lack of a non-violent control group did limit conclusions that could be drawn regarding clinician's ability to predict non-violent outcomes. Utilizing nonsexually violent offenders as the comparison group allowed the researcher to maximize power in both the sexually and nonsexually violent outcome categories, the key prediction in SVP evaluations.

Two additional considerations became apparent during data analysis and provide interesting avenues for future investigation. Clinicians in the present design were asked whether or not the offender suffered from a mental disorder or abnormality. As stated previously, to adequately address the legal standard, the offender must have a mental disorder that impairs volitionality. To parallel the standard, questions regarding *which* mental disorder they believed the offender had was not central. In future research regarding clinicians' decision making processes, particularly in SVP evaluations, it would be interesting to investigate which mental disorders are associated with SVP commitment recommendations.

Importantly, the way in which the question regarding support for SVP commitment was phrased may have been misleading. A potential limitation regarding estimated support of SVP commitment is that the question read, "What is the likelihood that you would support this person's commitment as an SVP?" A clearer question might be phrased, "What is the likelihood this person meets the criteria for commitment?" The original wording may have inadvertently tapped people's personal convictions rather than their ability to apply the relevant legal standard.

The present design focused on clinicians' ability to make predictions and apply a relevant legal standard. Because the refined sample consisted of forensic experts, reasonable extrapolations can be made from their final judgments to potential expert testimony. In future research, questions requiring "official testimony" from the clinician could build upon this information. Specifically, questions such as, "In court, would you testify that this individual is likely to commit future acts of sexual violence?" and "How confident would you be in that testimony?" Although the current procedure allowed for close inspection of the decision-making process, the extrapolation to expert testimony is tentative. A potentially interesting empirical



question asks whether experts alter their conclusions or confidence level when they are delivered as expert testimony.

### Summary and Conclusions

In reviewing the literature on violence risk assessment, the frustration of researchers and clinicians becomes evident. Some researchers hotly debate the actuarial versus clinical prediction question, while others refer to the debate dismissively as a dead horse (Monahan et al., 2002). Still other professionals champion improved methodology, calling for increased accuracy of prediction by identifying relevant variables (Hanson & Harris, 2000), and context-specific predictions (Skeem et al., 2000). Most recently, one researcher (Hart, 2002) advocated for the abandonment of violence prediction entirely.

Research to date has demonstrated that clinicians and techniques can achieve moderate predictive accuracy, particularly under very specific parameters. Contrary to Quinsey et al.'s (1998) critique of clinical judgement, the current research does not clearly support one method of prediction (e.g., actuarial) over any other (e.g., clinical judgement). The current project was also unable to demonstrate the superiority of any particular method in predicting violence and sexual violence, or in discriminating between them. Investigators concerned with the accurate prediction of violent behavior continue to refine and revisit their current methodologies and the resulting tools (see e.g., Harris, Rice, & Cormier, 2002). Others (see e.g., Hart, 2002) have responded to the lack of empirical support of prediction techniques by advocating the abandonment of prediction entirely, an approach simply not viable for forensic mental health professionals involved in SVP evaluations.

Two general conclusions are drawn from the current research on SVP determinations. First, given our current knowledge and methodology, a sequential, multimethod approach to SVP

evaluations appears prudent. Secondly, additional research leading to improved methodology is needed to address the legal standards imposed by SVP statutes (see e.g., *Hendricks, 1997; Crane, 2002*).

The multimethod approach recommended in the current research combines the strengths of structured approaches to risk assessment with the comprehensiveness and flexibility of unstructured approaches. This combined approach is similar to Webster et al.'s (1994) original recommendation that clinical data be used to adjust judgments derived from actuarial data. It is also similar to Hart's recent (2002) advocacy for the *anchored narrative approach* to risk assessment.

The second conclusion drawn from the present study is that improved methodology is required in order to adequately address the components of SVP statutes. Importantly, defining and assessing volitionality is central to many psycholegal constructs, including SVP evaluations. Measures validated on a relevant population of sex offenders will increase prediction accuracy. In addition, the validation of SVP evaluation guides has important *Daubert* implications (*Daubert v. Merrell Dow Pharmaceuticals, 1993*). To meet the demands of *Daubert*, these guides must be tested and peer reviewed. Current methods of risk assessment have not been adequately validated for SVP evaluations, rendering them potentially inadmissible.

In closing, risk assessment knowledge and methodology have vastly improved since the pioneering work of Steadman and Coccoza (1974) and Monahan (1981). Leaders in the field (e.g., Harris, Rice et al., 1993; 1998; 2002; Hart, 1998; 2002; Monahan et al., 2001) continue to make small but appreciable advances. As the science of risk assessment evolves, advances will likely include improvement in prediction of highly specific acts of violence, such as sexual violence, and the validation of forensic assessment instruments designed to address the

psycholegal construct of a sexually violent predator.

APPENDIX A  
CONFIDENTIAL INMATE SUMMARY (EXAMPLE)

Name: Vincent Williams  
Age: 30  
Release date: April 30, 2002

---

### Background History

Mr. Williams is the seventh of ten children and was reared in an intact family. Mr. Williams related that both of his parents and eight of his siblings have substance abuse problems, primarily alcoholism. When Mr. Williams was still residing in the family home, his father was involved in an incestuous relationship with Mr. Williams' older sister. This relationship produced two children who were raised as Mr. Williams' siblings. According to Mr. Williams, his father went to prison for two years as a result of this relationship, but since his release his mother and father have reconciled and now enjoy a more stable relationship. In addition, he also reports both his parents are now substance-free.

At 15, Mr. Williams quit school and moved out of the family home to reside with his first common law wife. This relationship lasted for two years and produced one child. The child is currently being raised by Mr. Williams' parents. He then entered into a common law relationship for approximately two years. He describes their separation as amicable, noting that it "just didn't work out." Nine years ago, Mr. Williams married Emily. The couple now have two young children. Emily describes herself as a supportive wife and looks forward to helping her husband "through his difficulties". Both Mr. Williams and his wife describe their relationship as happy, stable, and supportive.

Mr. Williams' work history is variable. He has held a series of jobs, including construction, fishing, and factory work. He was unemployed at the time of his arrest and had no education higher than the eighth grade.

### Substance abuse history

Mr. Williams began drinking heavily at age 14 and continues to do so. He admitted smoking marijuana on a daily basis during his late teens. He denies marijuana and other recreational drug use at this time. Mr. Williams admitted that he has an alcohol problem and that alcohol contributes to his criminal activities and marital difficulties. He admitted that, "when I'm drinking, I'm not faithful."

Following his previous incarceration for sexual assault of two girls under the age of 14, Mr. Williams participated in substance abuse treatment. He acknowledged that these crimes also were committed under the influence of alcohol. After his release, however, he returned to drinking heavily.

### Psychological assessment results

Results on the intelligence screening measure indicates that Mr. Williams is functioning within the average range of overall intellectual ability although his verbal skills appear to be somewhat hampered by his lack of schooling. His personality profile suggests antisocial tendencies and narcissistic personality features. Feelings of guilt, inferiority and hopelessness are frequent. Although there is no evidence of major mental illness, there is considerable current distress (anxiety, nervousness and depression) related to his present circumstances.

During the assessment, Mr. Williams emphasized his heavy drinking as being responsible for his crimes. However, he did not project all his problems onto alcohol consumption but

showed insight that his attitudes toward women have to change. He admitted what he did to his victim was wrong and acknowledged that she did not give permission. Nonetheless, he stated that he didn't "think she'd mind". He explained that he knew the victim socially and that she was known to be promiscuous. Therefore, he stated that he was surprised when she pressed charges against him.

#### Course of treatment during incarceration

Three months into Mr. Williams' incarceration, he began to attend AA and an inmate-run group for sex offenders. He is enrolled in school full-time and plans to achieve his GED. Mr. Williams, along with several other inmates, contributed personal funds in order to hire an outside facilitator for their sex offender group. Mr. Williams continues to participate in AA and sex offender treatment. He is described by group leaders as cooperative and motivated to make changes. He is also enrolled in school and is performing well.

#### Future plans

After release, Mr. Williams plans to move back home with his wife and children. His wife is eager to have him home and has expressed her willingness to remain loyal and supportive. Mr. Williams does not currently have any job prospects, but intends to begin looking for employment once released. Mr. Williams has also stated his intention to continue alcohol treatment in the community and has recognized the likelihood of his reoffending should a relapse occur.

#### Summary

Mr. Williams has served almost three years in prison for sexually assaulting a female acquaintance. This assault occurred while he was on parole for committing sexual assaults against two girls under the age of 14. He admits that on each occasion, he was under the influence of alcohol.

Mr. Williams acknowledges his serious alcohol problem and appears motivated to remain abstinent once released. He has also recognized that his attitudes and beliefs toward women must change. Mr. Williams appears to be a motivated and remorseful inmate who is serious about participating in remediation programs. In addition to alcohol and sex offender treatment, Mr. Williams has also been working on upgrading his education. Mr. Williams has a supportive home to which to return and is anxious to resume his life with his wife and children.

APPENDIX B  
VICTIM STATEMENT

At approximately 8:30 on a Tuesday night in January, I was coming home from work and pulled into the parking lot of my apartment building. As I got out of the car, I noticed a man I knew only slightly by the name of Vincent Williams. I was about to speak to him when he put his hand around my neck and told me not to make any noise. I thought I was going to pass out or throw up. Although I was struggling to get away, I wasn't able to make any noise because of the way he was holding me. Even so, he kept telling me not to scream. He had this deep threatening voice. He pulled me over to the area where people work on their cars. It's enclosed so even if anyone came by they wouldn't have seen me. I felt so helpless; as if my hands were tied. He pushed me to the ground and ripped my clothes off. I almost passed out. I was crying and hysterical. He got on top of me, pushed my legs apart and forced his penis inside me. I don't know for how long, although it felt like hours. I felt like I couldn't breathe, like being bound and gagged. He was hurting me. I think he called me a bitch. I was so scared that he would kill me. I was terrified. Then all of a sudden, he withdrew and rolled over and forced me to masturbate him. I was still terrified that he was going to kill me. Then he got up and pulled up his pants, told me not to make any noise and stay where I was. I don't know how I managed to get my skirt and coat back on or how I got back to my apartment. As soon as my husband took one look at me he called the police.



APPENDIX C  
INFORMATION/CONSENT FORM

## Informing legal decisions: Utility of risk assessment methods

Dear Colleague:

Assessing risk of future violent behavior continues to be an important yet controversial role of psychologists. The significance of their role has been underscored in the last decade with over one-third of states passing “Sexually Violent Predators” laws.

With minor variations, these statutes share four common elements: (a) a past act of sexually harmful conduct, (b) a current mental disorder or abnormality, (c) a finding of risk of future sexually harmful conduct, and (d) some connection between the mental abnormality and the danger. Very recently, the Supreme Court in *Kansas v. Crane* (2002) clarified that the individual committed must also have “serious difficulty in controlling behavior”. As with other psycholegal constructs, the ultimate determination of an individual as a “sexually violent predator” is made by the court. To inform their decisions, the courts consistently have turned to forensic psychologists as experts on risk assessment.

The present research study is designed to investigate several models of risk assessment. In the next few pages, you will be reviewing risk assessment data about a male criminal defendant. Following review of that information, you will be asked to make several judgments regarding his risk for future violence.

Your participation in this study is very important. We need a broad representation of psychologists and advanced doctoral students to address essential issues about risk assessment.

This study has been approved by the University of North Texas’ Institutional Review Board (940-565-3940). Your participation is, of course, completely voluntary and you can withdraw your participation at any time. By completing the survey, you are indicating your voluntary participation in this study. The ratings will take only 15-20 of your time. There are no identified risks of participating in this study. Please do not place your name or any identifying information on the questionnaires, we wish to keep this process completely anonymous. Finally, upon completion of your ratings, you will receive a small token of our appreciation. If you have any questions, please do not hesitate to ask the principal investigator, Becky Jackson, or her major professor, Dick Rogers.

Thank you for you participation,

Becky Jackson, M.S.

APPENDIX D

DEMOGRAPHIC INFORMATION AND REFERRAL QUESTION (EXAMPLE)

Name: Vincent Williams  
Age: 30  
Release date: April 30, 2002

---

#### Referral question

Vincent Williams, a 30-year-old male inmate, has been serving a 3-year sentence for sexually assaulting a 22-year-old female acquaintance.

Mr. Williams is about to be released from prison. The state in which you practice has petitioned to have Mr. Williams civilly committed under the state's sexually violent predator statute. The petition argues that Mr. Williams has not been rehabilitated while in prison and that he is unable to control his sexually deviant behavior. You are being asked to render conclusions regarding Mr. Williams likelihood of future sexual and nonsexual violent behavior. You will also be asked your opinion regarding the appropriateness of the state's petition for civil commitment.

To facilitate your risk assessment, you were provided a description of the client's offense history as well as a brief summary of the sexually violent predator statute.

#### Offense History

##### Index offense(s)

Sexual assault

##### Juvenile Record

Vandalism

Assault with a weapon causing bodily injury

Theft

##### Adult Criminal Record

Sex with a female under 14 (2 counts)

Consuming liquor in a public place

Probation violation

Possession of stolen property

#### Sexually Violent Predator Law

The current sexually violent predator statute effective in your state requires the demonstration of four elements: (a) a past act of sexually harmful conduct, (b) a current mental disorder or abnormality, (c) a finding of risk of future sexually harmful conduct, and (d) some connection between the mental abnormality and the danger.

APPENDIX E  
DEPENDENT VARIABLE RATINGS

Do you believe this person is likely to become violent in the next 5 years?

\_\_\_\_ Yes \_\_\_\_ No

What is the likelihood that this individual will commit future acts of *general* violence? \_\_\_\_%

How confident are you in this prediction? \_\_\_\_ %

How likely is this individual to commit future acts of *sexual* violence? \_\_\_\_ %

How confident are you in this prediction? \_\_\_\_ %

How strongly would you support this person's commitment under a *Sexually Violent Predator Act*? \_\_\_\_ %

Does this individual have a mental abnormality or personality disorder?

\_\_\_\_ Yes \_\_\_\_ No

If yes, what impact does it have on his ability to control his dangerous behavior?

1	2	3	4	5	6	7
No difficulty						Complete Inability

How amenable do you believe this person to be to treatment?

1	2	3	4	5	6	7
very untreatable			moderately treatable			very treatable

(For second and third ratings)

Did you change your ratings after reviewing this additional information?

If yes, why?

APPENDIX F  
DEMOGRAPHIC INFORMATION

What is your level of training?

- Ph.D.
- PsyD
- Graduate Student
- Other (Please Specify)\_\_\_\_\_

How many years have you been practicing after completing your doctoral training?

Years

What percentage of your practice is forensic in nature (students, please ignore)

%

How many risk assessments have you conducted?

What is your gender?

Male       Female

What is your current state (or country) of residence?

Is there a Sexually Violent Predator commitment law in the state in which you work?

Yes       No



APPENDIX G

DISTINGUISHING SEXUALLY FROM NONSEXUALLY VIOLENT OFFENDERS

UTILIZING A COMBINED ASSESSMENT APPROACH

Table 1

*Likelihood Estimates of Nonsexual and Sexual Violence for Sexually Violent Offenders by Order of Clinical Decision Making Method*

Decision Making Method ( <i>n</i> )	Nonsexual Violence Estimate <i>M</i> ( <i>SD</i> )	Sexual Violence Estimate <i>M</i> ( <i>SD</i> )	<i>d</i>
Actuarial/Structured (9)	60.78 (22.23)	54.11 (26.80)	.27
Actuarial/Unstructured (4)	28.75 (20.16)	38.75 (25.94)	.43
Clinical/Actuarial (11)	50.00 (25.50)	43.18 (24.21)	.27
Clinical/Unstructured (9)	54.44 (27.09)	50.00 (24.11)	.17
Unstructured/Actuarial (5)	49.00 (29.66)	43.00 (28.64)	.21
Unstructured/Clinical (10)	41.90 (30.11)	53.00 (23.24)	.42

*Note:* Cell sizes in each condition ( $\leq 11$ ) are less than optimal.

A remarkable finding elucidated by Table 1 is the considerable underestimation of violence risk provided by clinicians utilizing actuarial methods combined with unstructured clinical judgment. Clinicians who received these methods, in either order, rated violence potential to be less than 50%, although the absolute magnitude is less for those receiving unstructured clinical information after reviewing actuarial information. When clinicians viewed actuarial information first, they rated sexual offenders' risk of nonsexual violence to be  $M = 50.15$  and sexual violence  $M = 43.77$ . Following additional information of an unstructured nature, they reduced their ratings to  $M_s = 28.75$  and  $38.75$  respectively. Interesting, although the addition of unstructured judgment caused a reduction in ratings of violence potential, they appeared better able to discriminate between offenders likely to become sexually violent. During Phase II, they viewed sex offenders as more likely to be sexually violent than nonsexually violent (i.e., a reverse trend from ratings based on actuarial data alone).

Table 2

*Likelihood Estimates of Nonsexual and Sexual Violence for Non Sexually Violent Offenders by Order of Clinical Decision Making Method*

Decision Making Method ( <i>n</i> )	Nonsexual Violence Estimate <i>M</i> ( <i>SD</i> )	Sexual Violence Estimate <i>M</i> ( <i>SD</i> )	<i>d</i>
Actuarial/Structured (12)	55.50 (16.27)	57.58 (18.12)	.12
Actuarial/Unstructured (9)	48.89 (24.85)	54.00 (21.60)	.22
Clinical/Actuarial (8)	46.88 (34.53)	51.88 (24.78)	.17
Clinical/Unstructured (14)	63.21 (23.34)	48.57 (20.61)	.67
Unstructured/Actuarial (5)	59.60 (11.10)	36.60 (12.05)	1.99
Unstructured/Clinical (10)	38.10 (26.64)	36.60 (19.95)	.01

*Note:* Cell sizes ( $\leq 14$ ) in each condition are less than optimal.

Clinicians utilizing unstructured clinical judgment followed by actuarial data were best able to identify nonsexually violent offenders and rate their risk of nonsexual violence ( $M = 59.60$ ) considerably higher than those offenders' risk of sexually violent behavior ( $M = 36.60$ ). Importantly, the cell sizes, particularly of the Unstructured/Actuarial group were much lower than is optimal. Despite the large effect size ( $d = 1.99$ ), conclusions must be tempered.

Tentatively, it appears that the combination of unstructured/actuarial performs well in identifying violence risk potential among nonsexually violent offenders. As seen in Table 1, however, their ability to predict sexual risk among sex offenders was not very meaningful ( $d = .21$ ). Lastly, the utilization of unstructured clinical judgment, followed by structured clinical judgment appeared to result in *underestimations* of risk and little discriminatory ability among violent offenders.

## REFERENCES

- Addington v. Texas*, 441 U.S. 418 (1979).
- American Bar Association. (1998). *National benchbook on psychiatric and psychological evidence and testimony*. Washington, DC: Author.
- American Law Institute (1962). *Model penal code: Proposed official draft*. Philadelphia: Author.
- American Psychiatric Association. (1974). *Clinical aspects of the violent individual*. Washington, DC: Author.
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (Text Revision). Washington, DC: Author.
- American Psychological Association (1985). *Standards for educational and psychological testing*. Washington, DC: Author.
- Bank, S. C. (2002). Principle: Select and employ a model to guide data gathering. In K. Heilbrun, G. R. Marczyk, & D. DeMatteo (Eds.), *Forensic mental health assessment: A casebook* (pp. 54-64). New York, NY: Oxford University Press.
- 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.
- Barefoot v. Estelle*, 463 U.S. 880, (1983).
- Barker, E. & Buck, M. F. (1977). LSD in a coercive milieu therapy program. *Canadian Psychiatric Association Journal*, 22, 311-314.

- Barker, E., & Mason, M. (1968). Buber behind bars. *Canadian Psychiatric Association Journal*, 13, 61-72.
- Barker, E., & McLaughlin, A. J. (1977). The total encounter capsule. *Canadian Psychiatric Association Journal*, 22, 355-360.
- Baxstrom v. Herold*, 383 U.S. 107 (1966).
- Belfrage, H. (1998). Implementing the HCR-20 scheme for risk assessment in a forensic psychiatric hospital: Integrating research and clinical practice. *Journal of Forensic Psychiatry*, 9, 328-338.
- Benjaminsen, S., Gotzsche, L. K., Norrie, B., and Harder, L. (1996). Patient violence in a psychiatric hospital in Denmark: Rate of violence and relation to diagnosis. *Nordic Journal of Psychiatry*, 50, 233-242.
- Blaszczynski, A., Steel, Z., & McConaghy, N. (1997). Impulsivity in pathological gambling: The antisocial impulsivist. *Addiction*, 92, 75-87.
- Boer, D. P., Hart, S. D., Kropp, P. R., & Webster, C. D. (1997). *Manual for the Sexual Violence Risk – 20: Professional guidelines for assessing risk of sexual violence*. Vancouver, BC: The British Columbia Institute Against Family Violence.
- Borum, R. (1996). Improving the clinical practice of violence risk assessment: Technology, guidelines, and training. *American Psychologist*, 51, 945-956.
- Borum, R., Otto, R., & Golding, S. (1993). Improving clinical judgment and decision making in forensic evaluation. *Journal of Psychiatry and Law*, 21, 35-76.
- Brower, M. C., & Price, B. H. (2001). Neuropsychiatry of frontal lobe dysfunction in violent and criminal behavior: A critical review. *Journal of Neurology, Neurosurgery, and Psychiatry*, 71, 720-726.

- California Welfare and Justice Code, sec. 6600 (1996).
- Canadian Criminal Code (1987). *Dangerous Offender Act*, Part XXIV.
- Committee on Ethical Guidelines for Forensic Psychologists (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior*, 15, 655-665.
- Community Protection Act, Rev. Code of WA § 71.09 (1990).
- Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579 (1993).
- de Pauw, K. W., & Szulecka, T. K. (1988). Dangerous delusions: Violence and the misidentification syndromes. *British Journal of Psychiatry*, 152, 91-96.
- Douglas, K. S., Ogloff, J. R. P., Nicholls, T. L., & Grant, I. (1999). Assessing risk for violence among psychiatric patients: The HCR-20 violence risk assessment scheme and the Psychopathy Checklist: Screening Version. *Journal of Consulting and Clinical Psychology*, 67, 917-930.
- Douglas, K. S., & Webster, C. D. (1999). The HCR-20 violence risk assessment scheme: Concurrent validity in a sample of incarcerated offenders. *Criminal Justice and Behavior*, 26, 3-19.
- Falk, A. (1999). Sex offenders, mental illness and criminal responsibility: The constitutional boundaries of civil commitment. *American Journal of Law and Medicine*. Retrieved from [www.findarticles.com](http://www.findarticles.com).
- Foster, S. L., & Cone, J. D. (1995). Validity issues in clinical assessment. *Psychological Assessment*, 7, 248-260.
- Fox, S. G., & Walters, H. A. (1986). The impact of general versus specific expert testimony and eyewitness confidence upon mock juror judgment. *Law and Human Behavior*, 10, 215-228.

- Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment*. Washington, D.C.: American Psychological Association.
- Gardner, W., Lidz, C. W., Mulvey, E. P., & Shaw, E. C. (1996). Clinical versus actuarial predictions of violence in patients with mental illnesses. *Journal of Consulting and Clinical Psychology, 64*, 602-609.
- Goldberg, L. R. (1968). Simple models or simple processes? Some research on clinical judgments. *American Psychologist, 23*, 483-496.
- Granhag, P. A., & Stroemwall, L. A. (1999). Repeated interrogations: Stretching the deception detection paradigm. *Expert Evidence, 7*, 163-174.
- Grann, M., Belfrage, H., & Tengstrom, A. (2000). Actuarial assessment of risk for violence: Predictive validity of the VRAG and the Historical part of the HCR-20. *Criminal Justice and Behavior, 27*, 97-114.
- Grisso, T. (1986). *Evaluating competencies: Forensic assessments and instruments*. New York: Plenum Press.
- Grisso, T. (1991). Clinical assessments for legal decisionmaking: Research recommendations. In S. A. Shah & B. J. Sales (Eds.), *Law and mental health: Major developments and research needs* (pp. 49-81). Washington, DC: National Institute of Mental Health.
- Grisso, T., & Appelbaum, P.S. (1992). Is it unethical to offer predictions of future violence? *Law and Human Behavior, 16*, 621-633.
- 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. Solicitor General of Canada. [www.sgc.gc.ca](http://www.sgc.gc.ca)

- Hanson, R. K. (2000). *Risk Assessment*. Prepared for the Association for the Treatment of Sexual Abusers.
- 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., & Harris, A. J. R. (2000). Where should we intervene? Dynamic predictors of sexual assault recidivism. *Criminal Justice and Behavior, 27*, 6-35.
- Hanson, R. K. & Thornton, D. (1999). *Static-99: Improving actuarial risk assessments for sex offenders*. Solicitor General of Canada. [www.sgc.gc.ca](http://www.sgc.gc.ca)
- Harris, G. T., & Rice, M. E. (1997). Mentally disordered offenders: What research says about effective service. In C. D. Webster & M. A. Jackson (Eds.), *Impulsivity: Theory, assessment, and treatment* (pp. 361-393). New York, NY: Guilford Press.
- Harris, G. T., Rice, M. E., & Cormier, C. A. (2002). Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients. *Law and Human Behavior, 26*, 377-394.
- Harris, G. T., Rice, M. E., & Quinsey, V. L. (1993). Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument. *Criminal Justice and Behavior, 20*, 315-335.
- Hart, S. D. (1998). The role of psychopathy in assessing risk for violence: Conceptual and methodological issues. *Legal and Criminological Psychology, 3*, 121-137.
- Hart, S. D. (1999). Assessing violence risk: Thoughts and second thoughts. *Contemporary Psychology, 44*, 486-488.
- Hart, S. D. (August, 2002). *Violence risk assessment: An anchored narrative approach*. Presented at the American Psychological Association Annual Convention, Chicago.



- Heilbrun, K. (1997). Prediction versus management models relevant to risk assessment: The importance of legal decision-making context. *Law and Human Behavior, 21*, 347-359.
- Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York, NY: Kluwer Academic/Plenum Publishers.
- Heilbrun, K., Rogers, R., & Otto, R. K. (2002). Forensic assessment: Current status and future directions. In J. Ogloff (Ed.), *Psychology and law: Reviewing the discipline*.
- Howard, G. S., & Conway, C. G. (1986). Can there be an empirical science of volitional action? *American Psychologist, 41*, 1241-1251.
- Janus, E. S. (1998). *Hendricks* and the moral terrain of police power civil commitment. *Psychology, Public Policy, and Law, 4*, 297-322.
- Janus, E. S. (2000). Sexual predator commitment laws: Lessons for law and the behavioral sciences. *Behavioral Sciences and the Law, 18*, 5-21.
- Janus, E. S., & Nudell, L. J. (2000). Defending sex offender commitment cases. In A. Schlink and F. Cohen (Eds.), *The sexual predator* (pp. 1-29). Kingston, NJ: Civic Research Institute.
- Kan. Stat. Ann 59-29 (a); SB 671 (1994).
- Kansas v. Crane*, 534 U.S. 407 (2002).
- Kansas v. Hendricks*, 521 U.S. 346 (1997).
- Kennedy, H. G., & Grubin, D. H. (1992). Patterns of denial in sex offenders. *Psychological Medicine, 22*, 191-196.
- Kroner, D. G., & Mills, J. F. (2001). The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions. *Criminal Justice and Behavior, 28*, 471-489.

- Langevin, R. (1988). Defensiveness in sex offenders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 269-290). New York, NY: Guilford Press.
- Lessard v. Schmidt*, 413 F. Supp. 1318 (E.D. Wis. 1976).
- Lidz, C. W., Mulvey, E. P., Apperson, L. J., Evanczuk, K., & Shea, S. (1992). Sources of disagreement among clinicians' assessments of dangerousness in a psychiatric emergency room. *International Journal of Law and Psychiatry*, 15, 237-250.
- Lidz, C. W., Mulvey, E. P., & Gardner, W. (1993). The accuracy of predictions of violence to others. *Journal of the American Medical Association*, 269, 1007-1011.
- Link, B. G., & Stueve, A. (1994). Psychotic symptoms and the violent/illegal behavior of mental patients compared to community controls. In J. Monahan & H. J. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 137-159). Chicago: University of Chicago Press.
- Litwack, T. R. (2001). Actuarial versus clinical assessments of dangerousness. *Psychology, Public Policy, and Law*, 7, 409-443.
- Lieb, R. & Matson, S. (1998). *Sexual predator commitment laws in the United States: 1998 update*. Washington State Institute for Public Policy.
- Lynett, E. (1990). The effects of a stated diagnosis of antisocial personality disorder, victim statements, and offender's antisocial history on psychiatrists' opinions in dangerous offender determinations. *Unpublished doctoral dissertation*.
- Lynett, E. & Rogers, R. (2000). Emotions overriding forensic opinions? The potentially biasing effects of victim statements. *The Journal of Psychiatry and Law*, 28, 457.
- McAllister, T. W. (1992). Neuropsychiatric sequelae of head injuries. *Psychiatric Clinics of North America*, 15, 395-413.

- McNiel, D. E., & Binder, R. L. (1991). Clinical assessment of the risk of violence among psychiatric inpatients. *American Journal of Psychiatry*, *148*, 1317-1321.
- McNiel, D. E., Sandberg, D. A., & Binder, R. L. (1998). The relationship between confidence and accuracy in clinical assessment of psychiatric patients' potential for violence. *Law and Human behavior*, *22*, 655-669.
- Meehl, P. E. (1954/1996). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Northvale, NJ: Jason Aronson, Inc.
- Melton, G. B., Petrila, J., Poythress, N. G., & Slobogin, C. (1997). *Psychological evaluations for the courts* (2<sup>nd</sup> ed.). New York: Guilford Press.
- Miller, G. H. (1992). Insanity standards. *Psychiatric Annals*, *22*, 626-631.
- Monahan, J. (1981). *The clinical prediction of violent behavior*. Washington, DC: National Institute of Mental Health.
- Monahan, J. (1992). Mental disorder and violent behavior. *American Psychologist*, *47*, 511-521.
- Monahan, J. (2000). Violence risk assessment: Scientific validity and evidentiary admissibility. *Washington and Lee Law Review*, *57*, 901-918.
- Monahan, J. & Steadman, H. J. (2001). Violence risk assessment: A quarter century of research. In L. E. Frost & R. J. Bonnie (Eds.), *The evolution of mental health law* (pp. 195-211). Washington, DC: American Psychological Association.
- Monahan, J. & Steadman, H. J. (1994). *Violence and mental disorder: Developments in risk assessment*. Chicago, IL: The University of Chicago Press.
- Monahan, J., Steadman, H. J., Silver, E., Appelbaum P. S., Robbins, P. C., Mulvey, E. P., et al. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence*. New York, NY: Oxford University Press, Inc.

- Morse S. (1978). Crazy behavior, morals, and science: An analysis of mental health law. *Southern California Law Review*, 51, 527-554.
- Morse, S. J. (1998). Fear of danger, flight from culpability. *Psychology, Public Policy, and Law*, 4, 250-267.
- Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology*, 62, 783-792.
- Mossman, D. (2000). Commentary: Assessing the risk of violence- Are accurate predictions useful? *Journal of the American Academy of Psychiatry and Law*, 28, 272-281.
- Mulvey, E. P. & Lidz, C. W. (1998). Clinical prediction of violence as a conditional judgment. *Social Psychiatry and Psychiatric Epidemiology*, 33, S107-S113.
- Norko, M. A. (2000). Commentary: Dangerousness- a failed paradigm for clinical practice and service delivery. *The Journal of the American Academy of Psychiatry and the Law*, 28, 282-289.
- New Jersey, P.L. c.71, Senate Bill 895 (1998).
- Newman, J. P., Wallace, J. F., Schmitt, W. A., & Arnett, P. A. (1997). Behavioral inhibition system functioning in anxious, impulsive, and psychopathic individuals. *Personality and Individual Differences*, 23, 583-592.
- North Dakota House Bill 1047, (1997).
- Oskamp, S. (1965). Overconfidence in case study judgments. *Journal of Consulting Psychology*, 29, 261-265.
- Otto, R. K. (1992). Prediction of dangerous behavior: A review and analysis of “second-generation” research. *Forensic Reports*, 5, 103-133.

- Otto, R K. (August, 2001). *Setting specific effects on test results in forensic cases*.  
Presented at the American Psychological Association Annual Convention, San Francisco, CA.
- Overall, J. E., & Klett, C. J. (1972). *Applied multivariate analysis*. New York: McGraw-Hill.
- Penrod, S., & Cutler, B. (1995). Witness confidence and witness accuracy: Assessing their forensic relation. *Psychology, Public Policy, and Law, 1*, 817-845.
- Prentky, R. A., Knight, R. A., Lee, A., & Cerce, D. (1995). Predictive validity of lifestyle impulsivity for rapists. *Criminal Justice and Behavior, 22*, 106-128.
- Pruesse, M. G., & Quinsey, V. L. (1977). The dangerousness of patients released from maximum security: A replication. *Journal of Psychiatry and Law, 5*, 293-299.
- Quinsey, V. L., & Ambtman, R. (1979). Variables affecting psychiatrists' and teachers' assessments of the dangerousness of mentally ill offenders. *Journal of Consulting and Clinical Psychology, 47*, 353-362.
- Quinsey, V. L., Harris, G. T., Rice, M. E. & Cormier, C. A. (1998). *Violent offenders: appraising and managing risk*. Washington, DC: American Psychological Association.
- Quinsey, V. L., Pruesse, M., & Fernley, R. (1975). Oak Ridge patients: Prerelease characteristics and postrelease adjustment. *Journal of Psychiatry and Law, 3*, 63- 77.
- Rachlin, S., Halpern, A. L., & Portnow, S. L. (1984). The volitional rule, personality disorders, and the insanity defense. *Psychiatric Annals, 14*, 139-147.
- Rice, M. E. & Harris, G. T. (1995). Violent recidivism: Assessing predictive validity. *Journal of Consulting and Clinical Psychology, 63*, 737-748.
- Rice, M. E. & Harris, G. T. (1997). Cross-validation and extension of the violence risk appraisal guide for child molesters and rapists. *Law and Human Behavior, 21*, 231-241.

- Rice, M. E., Harris, G. T., & Cormier, C. (1992). Evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. *Law and Human Behavior, 16*, 399-412.
- Rice, M. E., Harris, G. T., Lang, C., & Bell, V. (1990). Recidivism among male insanity acquittees. *The Journal of Psychiatry and Law, 18*, 379-403.
- Rogers, R. (1990). Models of feigned mental illness. *Professional Psychology, 21*, 182-188.
- Rogers, R., & Cavanaugh, J. L. (1981). Application of the SADS diagnostic interview to forensic psychiatry. *Journal of Psychiatry and Law, 9*, 329-344.
- Rogers, R., & Dickey, R. (1991). Denial and minimization among sex offenders: A review of competing models of deception. *Annals of Sex Research, 4*, 49-63.
- Rogers, R., & Ewing, C. P. (1989). Ultimate opinion proscriptions: A cosmetic fix and a plea for empiricism. *Law and Human Behavior, 13*, 357-374.
- Rogers, R., Gillis, J. R., Turner, R. E., & Frise-Smith, T. (1990). The clinical presentation of command hallucinations. *American Journal of Psychiatry, 147*, 1304-1307.
- Rogers, R. & Shuman, D. W. (2000). *Conducting insanity evaluations* (2<sup>nd</sup> ed.). New York, NY: The Guilford Press.
- Schall v. Martin*, 467 U.S. 253, 264 (1984).
- Schmitt, W. A., Brinkley, C. A., & Newman, J. P. (1999). Testing Damasio's somatic marker hypothesis with psychopathic individuals: Risk takers or risk averse? *Journal of Abnormal Psychology, 108*, 538-543.
- Schopp, R. F. (1998). Civil commitment and sexual predators: Competence and condemnation. *Psychology, Public Policy, and Law, 4*, 323-376.

- Seto, M. C., & Barbaree, H. E. (1999). Psychopathy, treatment behavior, and sex offender recidivism. *Journal of Interpersonal Violence, 14*, 1235-1248.
- Sewell, K. W., & Salekin, R. T. (1997). Understanding and detecting dissimulation in sex offenders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2<sup>nd</sup> ed., pp. 328-350). New York, NY: Guilford Press.
- Sjostedt, G., & Langstrom, N. (2001). Actuarial assessment of sex offender recidivism risk: Cross-validation of the RRASOR and the Static-99 in Sweden. *Law and Human Behavior, 25*, 629-645.
- Skeem, J. L., Mulvey, E. P., & Lidz, C. W. (2000). Building mental health professionals' decisional models into tests of predictive validity: The accuracy of contextualized predictions of violence. *Law and Human behavior, 24*, 607-628.
- Sreenivasan, S., Kirkish, P., Garrick, T., Weinberger, L. E., & Phenix, A. (2000). Actuarial risk assessment models: A review of critical issues related to violence and sex-offender recidivism assessments. *The Journal of the American Academy of Psychiatry and the Law, 28*, 438-448.
- Steadman, H. J. (2000). From dangerousness to risk assessment of community violence: Taking stock at the turn of the century. *Journal of the American Academy of Psychiatry and Law, 28*, 265-271.
- Steadman, H. J., & Coccozza, J. (1974). *Careers of the criminally insane*. Lexington, MA: Lexington books.
- Stone, A. A. (1975). Comment on Peszke M.A.: Is dangerous an issue for physicians in emergency commitment? *American Journal of Psychiatry, 132*, 829-831.

- Strand, S., Belgrage, H., Fransson, G., & Levander, S. (1999). Clinical and risk management factors in risk prediction of mentally disordered offenders—more important than historical data? A retrospective study of 40 mentally disordered offenders assessed with the HCR-20 violence risk assessment scheme. *Legal and Criminological Psychology, 4*, 67-76.
- Suarez, G. A., Chalk, C. H., Russell, J. W., Kim, S. M., O'Brien, P. C., et al. (2001). Diagnostic accuracy and certainty from sequential evaluations in peripheral neuropathy. *Neurology, 57*, 1118-1120.
- Swanson, J. W., Borum, R., Swartz, M. S., & Monahan, J. (1996). Psychotic symptoms and disorders and the risk of violent behavior in the community. *Criminal Behavior and Mental Health, 6*, 317-338.
- Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics* (4<sup>th</sup> ed.). Needham Heights, MA: Allyn & Bacon.
- Tardiff, K. (1998). Unusual diagnoses among violent patients. *Psychiatric Clinics of North America, 21*, 567-576.
- Texas Commission on Law Enforcement (2002). Officer standards and education: Sex offender. Retrieved from [www.utexas.edu](http://www.utexas.edu).
- Thornberry, T., & Jacoby, J. (1979). *The criminally insane: A community follow-up of mentally ill offenders*. Chicago: University of Chicago Press.



- Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence (Version 2)*. Burnaby, BC: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Webster, C. D., Eaves, D., Douglas, K. S., & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk*. Vancouver: Simon Fraser University and British Columbia Forensic Psychiatric Services Commission.
- Webster, C. D., Harris, G. T., Rice, M. E., Cormier, C., & Quinsey, V. L. (1994). *The violence prediction scheme: Assessing dangerousness in high risk men*. Toronto, Canada: University of Toronto, Centre of Criminology.
- Webster, C. D., & Jackson, M. A. (1997). *Impulsivity: Theory, assessment, and treatment*. New York, NY: Guilford Press.
- Weiner, I. (1999). Writing forensic reports. In A. K. Hess and I. B. Weiner (Eds), *The handbook of forensic psychology* (pp. 501-520). New York, NY: John Wiley & Sons, Inc.
- Wells, G. L., Lindsay, R. C., & Tousseignant, J. P. (1980). Effects of expert psychological advice on human performance in judging the validity of eyewitness testimony. *Law and Human Behavior*, 4, 275-285.
- Werner, P. D., Rose, T. L., & Yesavage, J. A. (1983). Reliability, accuracy and decision-making strategy in clinical predictions of imminent dangerousness. *Journal of Consulting and Clinical Psychology*, 51, 815-825.
- Werner, P. D., Rose, T. L., Yesavage, J. A., & Seeman, K. (1984). Psychiatrists' judgments of dangerousness in patients on an acute care unit. *American Journal of Psychiatry*, 141, 263-266.
- Wisconsin Stat. Chapter 980 (1994).

Ziskin, J. (1995). *Coping with Psychological and Psychiatric Testimony: Volume 1* (5<sup>th</sup> ed.).  
Los Angeles: Law and Psychology Press.