Psychology, Public Policy, and Law
2006, Vol. 12, No. 4, 357–393

1076-8971/06/$12.00 DOI: 10.1037/1076-8971.12.4.357

SEXUALLY VIOLENT PREDATORS IN THE COURTROOM
Science on Trial

Robert A. Prentky
Justice Resource Institute

Eric Janus
William Mitchell College of Law

Howard Barbaree
University of Toronto

Barbara K. Schwartz
Justice Resource Institute

Martin P. Kafka
Harvard University Medical School

Adjudication of sexually violent predator commitment laws places demands on science. In the current article, the authors discuss the determination of mental abnormality and its reliance on medical nosological systems. Second, the authors examine the determination of current risk by reviewing three common concerns: (a) mechanistic estimations of risk, (b) mitigation of risk as a function of age, and (c) estimation of contemporaneous (dynamic) risk. The authors focus specifically on determinations of risk posed by the nexus of mental abnormality with prior history of sexually violent acts. Third, the article examines relevant, though sometimes nonstatutory, considerations, namely, the standards and expectations for treatment provided in high-security civil commitment programs. Potentially important dynamic or time-varying factors that may mitigate risk, such as offender age and treatment, are considered. Recommendations to promote "good science" and to avoid "bad science" are included with respect to determinations of mental abnormality, risk of reoffending, and treatment.

Keywords: sexually violent predator, science, risk, mental disorder, treatment

In most realms of the law, science is supportive and elucidative but does not comprise the key elements of the case. In sexually violent predator (SVP) commitment laws, however, science is integral because the targets of these laws are people with some mental disorder or abnormality that gives rise to future risk of harmful behavior, justifying the deprivation of physical liberty and the imposition of treatment. The integral role of science in SVP statutes leads to unique dangers to both law and science.

Our concern in this article is to address the most critical problems that occur at the intersection of law and science in the SVP context. We have twin concerns:

Robert A. Prentky and Barbara K. Schwartz, Research Department, Justice Resource Institute, Bridgewater, Massachusetts; Eric Janus, William Mitchell College of Law; Howard Barbaree, Department of Psychiatry, University of Toronto, Toronto, Ontario, Canada; Martin P. Kafka, Harvard University Medical School.

We wish to acknowledge the invaluable assistance of Deborah Cavanaugh and Ann Pimental of the Justice Resource Institute Research Department and Kristopher Lee of the William Mitchell College of Law in the preparation of this article.

Correspondence concerning this article should be addressed to Robert A. Prentky, Research Department, Justice Resource Institute, 63 Main Street, Suite 6, Bridgewater, MA 02324. E-mail: rprentky@jri.org

that "good science" will be unrecognized or misunderstood by the law and that the pressures of the law will not only use but encourage "bad science." Both concerns are potential sources of injustice and both threaten the integrity of science and the law. Good science, in this context, refers to the faithful and rigorous adherence to the findings, the limitations, and the conclusions of published, peer-reviewed articles in scientific journals. Bad science refers to the intentional or uninformed distortion, misinterpretation, or selective reporting of findings from scientific articles.

SVP laws lock up sex offenders judged to be mentally disordered and dangerous after the expiration of their criminal sentences (Janus, in press). Spawned in the early 1990s in Washington and Minnesota, these statutes are now found in 16 states and are being considered in several more (New York, Delaware, New Mexico, Ohio, and Vermont). SVP laws were developed to address a pressing political and public safety imperative, incapacitating patently dangerous sex offenders beyond the legally mandated end of their criminal sentence. Fundamental constitutional restrictions prevented states from simply lengthening previously imposed criminal sentences. States adopted a traditional legal form—civil commitment—and adapted it to a new purpose: confining sex offenders deemed too dangerous to release from prison (Janus, in press).

Despite some state-to-state variation, all SVP laws share four basic elements. To be committed, an individual must (a) have some mental disorder or abnormality that (b) causes or is associated with (c) an elevated risk of future sexual misconduct. In addition, all SVP laws profess that (d) a purpose of commitment is treatment (Lieb & Gookin, 2005). SVP laws raise many questions of constitutional law and morality. At issue, on the one hand, is the safety of the public from violent sexual crime and, on the other, personal liberty. When they were enacted, the SVP laws pressed against the ill-defined constitutional boundaries of standard civil commitment (Janus, 1998). The new laws expand the range of mental disorders that could serve as predicates for civil commitment. Like conventional civil commitment laws, SVP laws require a finding of future dangerousness. Prediction in the SVP setting is more problematic, however, because of the absence of acute psychiatric symptoms tied closely to the predicted harm and because most targets of SVP commitment have been imprisoned for years or even decades following their most recent episodes of violence. Thus, SVP laws ask for long-term predictions of future violence that are often based on long-distant predicate violence. Moreover, compared with the pharmacological treatments available for psychiatric illnesses, the efficacy of psychological treatment for sexual offenders remains much more uncertain.

Constitutional Challenges

The constitutionality of SVP laws has been challenged in a series of cases. Although the challenges failed to overturn the laws, they did clarify the constitutional boundaries for the permissible use of SVP laws. First, the courts rejected claims that the ability to predict dangerousness in sex offenders was so weak as to violate categorically due process protections (Janus & Prentky, 2003). Rather, the courts suggested that commitment is appropriate only for a very narrow group of the most dangerous offenders. The courts have been wary, however, of defining too closely how the most dangerous will be identified. Several courts have

articulated at a rather abstract level some constraints on the likelihood of future violence that is necessary to justify commitment (e.g., "highly likely"), and several courts have designated the kind of violence (e.g., "predatory") necessary (Janus & Prentky, 2003). At least one court has specified the conditions under which danger is to be assessed (i.e., with treatment and under supervision), and a few courts have begun to set some standards for the validity of the assessments of risk (Janus & Prentky, 2003). The risk assessment or danger predication aspect of SVP cases, however, remains highly discretionary. Real decision making is situated with juries or trial courts, and few appellate courts intervene to set standards that might lead to some accountability or uniformity in risk thresholds.

Second, courts confirmed that dangerousness alone is constitutionally inadequate to support commitment (*Foucha v. Louisiana*, 1992; *Kansas v. Hendricks*, 1997). Some form of mental disorder is a constitutional predicate for civil commitment. Indeed, it seems clear that it is the presence of a mental abnormality that saves SVP laws from being unconstitutional preventive detention. The Supreme Court has never explained, however, why the presence of a mental disorder has such central constitutional significance (*Kansas v. Crane*, 2002; *Kansas v. Hendricks*, 1997). Thus, it is not surprising that despite the constitutional centrality of the mental disorder requirement, the precise boundaries set by the Constitution remain somewhat obscure. It is in large measure this very obscurity that gives mental health diagnosis testimony such importance in SVP cases and opens the door to abuses of science.

Finally, several courts have addressed the question of treatment in the SVP context. As discussed above, it now appears more certain that treatment must be one of the purposes for civil commitment if the confinement is to be constitutional (Janus, 2003). All SVP laws claim that treatment is a purpose, and all SVP programs make some effort to provide something they call treatment. It also seems clear that amenability to treatment is not a question that is addressed at the threshold of commitment. In other words, as long as the state in some sense intends to provide treatment, it need not prove, as a condition precedent to commitment, that any particular individual will benefit from treatment (Janus, 2003). For this reason, testimony regarding treatment does not often play a role in the judicial consideration of a petition for commitment. Nonetheless, testimony about an individual's response to treatment may play a central role in release decisions, and further litigation regarding the constitutionality of SVP programs may turn on compliance with professional standards for confinement and treatment.

Science in the Courtroom

From DNA testing to accident reconstruction and in toxic torts, paternity, and child custody, the law invites and relies on both the hard and soft sciences. In each of these contexts, as in the SVP arena, the interface between law and science brings benefits as well as problems. Our goal in this article is to identify the interface problems in the SVP area and make suggestions for maximizing the benefits. A brief overview of the reasons for the interface problems is helpful and suggests the reasons why the interface is particularly problematic in the SVP area.

In any context, the science–law interface must negotiate the potential for

breakdown in three basic areas: translation, boundaries, and evaluation (Schopp, Scalora, & Pearce, 1999). The SVP laws entail serious threats to the interface in each of these categories. The translation of scientific into legal categories is a critical and potentially hazardous step. In the legal context, the categories have normative significance, representing moral or value judgments about (in the context we are dealing with) what kinds of circumstances justify the deprivation of liberty. Scientific laws and categories, on the other hand, are largely descriptive and gain their validity not because they are normatively sound, but because they are found to be useful as descriptors or predictors of some presumptive objective reality. Thus, to confuse legal and scientific categories is to commit what some philosophers call the *naturalistic fallacy*, thoughtlessly equating what is with what ought to be.

The problem of translation is amplified in the SVP context because the legal categories and thresholds are themselves poorly defined, in large measure because of the lack of clarity about the normative values underlying these laws. As we describe more fully below, there is a fundamental vagueness about the degree and nature of the risk that is normatively sufficient to justify the deprivation of liberty (Janus & Prentky, 2003). Similarly, there is little clarity about what features of a person's psychological makeup justify the extraordinary use of civil commitment (Janus, 2001). Without clarity on the legal side, the translation between science and law is largely unmoored and subject to manipulation.

The lack of normative clarity in the law also impairs the process of setting and patrolling the boundaries between science and law. Mental health experts have no legitimate expertise in defining risk thresholds or in defining the normative legal standards for a constitutionally sufficient mental disorder. Yet because of the vagueness of the legal standards, it is tempting for legal decision makers to treat mental health testimony as if it had normative, as well as descriptive, import (Schopp et al., 1999).

Finally, judicial decision makers must be able to evaluate the science that is proffered in the courtroom. They must determine the degree of fit between the testimony and the relevant legal inquiry, as well as the soundness of the science underlying the testimony, to determine the proper weight to give the expert's opinion (Janus & Prentky, 2003). In the SVP context, several factors coalesce to undermine the effectiveness of this evaluative process. The high political salience of sexual predator policy combines with the real harm caused by sexual violence to elevate concern for false negative judgments over concern about false positives. Added to this is the opacity of psychiatric diagnosis and the statistical complexity involved in evaluating and interpreting actuarial risk assessment data. Exacerbating these factors is the increasing tendency, described below, for experts to stretch or distort the science—to introduce bad science—in response to the strong advocacy pressures inherent in SVP proceedings (Janus, 2004).

Science can have a positive impact on the SVP courtroom. Properly translated, kept within its appropriate boundaries, and well evaluated, it can bring much needed accuracy and accountability to the consequential balance between safety and liberty. Without proper controls, science can obscure findings and compound the injustice of a constitutionally extraordinary deprivation of liberty. The danger is greatest when the science imported into the courtroom is bad science. This problem is amplified because bad science, having escaped from the

constraints that govern good science, is highly vulnerable to manipulation. Bad science is, of course, nothing new, and the courtroom is certainly not the only place that we witness it. The esteemed Union of Concerned Scientists launched a Sound Science Initiative to counter the use of "junk science" by policymakers, by the government, and by the media (Cole, 1996). We suggest in this article, however, that the intrusion of bad science occurs with disturbing frequency in SVP cases.

Perlin (1991) characterized the legal system's handling of mental health issues as *pretextual*. He adopted that term to refer to the use of legal "fictions" by courts and litigants "to falsely interpret the true meaning of legislation" (Perlin, 1993, p. 631). He warned that the tolerance of pretextuality "creates ambivalence toward concepts of law and justice. Toleration of 'sleight of hand' in the law's theoretical bases breeds cynicism and fosters an atmosphere of systemic manipulation by litigants, legislators, litigators, and courts" (Perlin, 1993, p. 632).

We would argue that the misuse of science in the SVP courtroom is a variation of pretextuality. It provides a legitimizing cover, allowing the state to cast the constitutionally doubtful preventive detention of dangerous individuals as constitutionally safe civil commitment. In doing so, as Perlin observed, one undercuts the integrity of the ultimate goal of reducing sexual violence by jeopardizing the critical benefits that accrue from an empirically informed understanding of the roots, risk factors, and management strategies related to sexual violence (Perlin, 1993).

The clearest example of this is the role that mental disorder diagnosis plays in SVP cases. As we detail below, it is the existence of some form of mental disorder that saves SVP commitments from being considered unconstitutional preventive detention and transforms them into legitimate civil commitment (Janus, in press). Thus, testimony about mental disorder serves a critical gatekeeping function in SVP cases. If those evaluated under the SVP laws are truly mentally disordered, the legitimacy of their selection for this extraordinary form of legal incarceration is strengthened. The distinction between those who are subject to preventive detention and those who are not at least has the appearance of some validity; it is not simply a political choice by those with the power to make the choice. When the law relies on bad science, such as individually defined mental disorder, this external touchstone is severely compromised. Far from providing transparency and accountability, the presence of bad science further obscures the arbitrariness of the legal process. The key choices and the key distinctions that ultimately determine who is subject to lifetime confinement are made not on the basis of law, not by judges duly appointed, and not on the basis of externally validated science.

This brings us to the last of the dangers arising from the misuse of science in the SVP courtroom. There is an important and looming possibility that science itself will be distorted by the pressures of the SVP process. The introduction of new mental disorders and the distortion of standard mental disorder categories undercuts the legitimacy of science and limits its ability to provide a sound and objective touchstone in the fight to understand and reduce sexual violence.

In the current article, we set forth and discuss some of the most critical misuses of science in the SVP courtroom. In choosing the issues on which to focus, we have returned to the principal elements of the civil commitment statutes.

First, we begin with a discussion of the determination of mental abnormality and its reliance on medical nosological systems. Second, we examine the determination of current risk. This section is broken down into three areas of common concerns: (a) mechanistic estimations of risk, (b) mitigation of risk as a function of age, and (c) estimation of contemporaneous (dynamic) risk. Third, we examine a relevant (though sometimes nonstatutory) consideration: the standards and expectations for treatment provided in high-security civil commitment programs. Our coverage in this article is not comprehensive. There are, undoubtedly, many other areas of misuse that we have intentionally or inadvertently omitted.

Mental Abnormality

The courts have been clear that dangerousness alone cannot support civil commitment. In some way, not fully explained, it is the additional presence of a mental disorder that transforms the deprivation of liberty from unconstitutional preventive detention to constitutional civil commitment. A key point of contention in SVP cases focuses on the legal definition that makes mental disorder constitutionally sufficient to support commitment.

In order for a mental disorder to support SVP commitment, there must be some sort of nexus or connection between the disorder and a risk of sexual reoffending. Typical statutory language requires proof of a mental disorder that "predisposes a person to commit sexual acts. . ." (Ariz. Rev. Stat. § 36.3701). In *Hendricks*, the Supreme Court upheld the Kansas SVP Act against constitutional challenge, in part because the statute required such a connection. The Court stated:

> The Kansas Act . . . requires a finding of future dangerousness, and then *links that finding* to the existence of a "mental abnormality" or "personality disorder" that *makes it* difficult, if not impossible, for the person to control his dangerous behavior. (*Kansas v. Hendricks*, 1997; emphasis added)

Second, the Supreme Court has stated that the "nature" and "severity" of the mental disorder "must be sufficient to distinguish the dangerous sexual offender whose serious mental illness, abnormality, or disorder subjects him to civil commitment from the dangerous but typical recidivist convicted in an ordinary criminal case" (*Kansas v. Crane*, 2002). This key point deserves emphasis. The Supreme Court is instructing that civil commitment must be the exception, not the rule. Individuals subject to SVP commitments must be "distinguished" by their mental disorders from "dangerous but typical" recidivists.

Third, one way in which a mental disorder can satisfy both the nexus requirement and the distinction requirement is by impairing the individual's ability to control his sexual impulses, a notion referred to as *volitional impairment* (VI). In *Kansas v. Crane* (2002), the Supreme Court concluded, "It is enough to say that there must be proof of serious difficulty in controlling behavior." The Court was not more specific:

> The Constitution's liberty safeguards in the area of mental illness are not always best enforced through precise bright-line rules. States retain considerable leeway in defining the mental abnormalities and personality disorders that make an individual eligible for commitment; psychiatry, which informs but does not control

> ultimate legal determinations, is an ever-advancing science, whose distinctions do not seek precisely to mirror those of the law.

Although an extended discussion of VI is undertaken elsewhere (e.g., Mercado, Schopp, & Bornstein, 2005) and is unnecessary for our purposes, three points are usefully made. First, the notion of VI has some relation to the defense of criminal nonresponsibility (e.g., Morse, 1994), which, in some variants, contains a volitional prong. Second, the concept, both in the criminal context and in the SVP context, is notoriously opaque. La Fond (2000) pointed out that the reason for the joint recommendation of the American Bar Association and the American Psychiatric Association (that the volitional prong of the American Law Institute's 1962 insanity defense standard be dropped) was that VI was impossible for a clinician to assess reliably. Third, the Crane case (consistent with the established jurisprudence of civil commitment laws over the years) makes clear that VI is not the only type of impairment that might support commitment. To date, however, no other form of impairment has received the Court's imprimatur in the SVP context.

In some states, criminal acts may be excused on proof of irresistible impulses or an impairment in ability to conform one's behavior to the law when caused by mental illness (psychosis). The use of a volitional standard in the insanity defense has a long legacy, dating back over 100 years in the United States (*Parsons v. Alabama,* 1886). In 1939, the Minnesota Supreme Court appeared to import that idea into the civil commitment arena, issuing a construction of the Minnesota sex offender commitment law that narrowed its application to those individuals who demonstrated an "utter lack of power to control sexual impulses" (*State ex rel Pearson v. Probate Court*, 1939). In *Kansas v. Hendricks* (1997), the U.S. Supreme Court affirmed the centrality of VI (cf. Mercado et al., 2005). The Kansas Supreme Court was strongly influenced by Justice Thomas's numerous references in *Hendricks* to dyscontrol, in holding that "The controlling issue is whether it is constitutionally permissible to commit Crane as a sexual predator absent a showing that he was unable to control his dangerous behavior" (*In re Crane*, 2000).

Closer scrutiny, however, reveals that the volitional dysfunction standard—whether in the criminal insanity context or the SVP civil commitment context—is highly problematic. In its amicus brief in support of the petitioner (State of Kansas) in *Kansas v. Crane*, the Association for the Treatment of Sexual Abusers (2001) argued that the SVP "cannot control" standard is "meaningless and unworkable" (p. 3) and that the ancestral standard, the "irresistible impulse insanity test, has been largely rejected by both the medical and legal professions" (p. 2). As has been articulated on numerous prior occasions, it is problematic, and perhaps impossible, to distinguish between impulses that are "irresistible" and impulses that simply are not resisted (p. 2).

The volitional dysfunction standard as applied in insanity defenses is rarely appropriate in the SVP context. To begin with, very few sex offenders are psychotic (e.g., Långström, Sjöstedt, & Grann, 2004; McElroy et al., 1999). Moreover, because the overarching motivation for the modern SVP laws is to continue the secure confinement of sex offenders after they have served prison sentences, the overwhelming majority of SVP candidates have been convicted and

thus are not insane. Thus, if VI is to have any meaning in the SVP context, it might be a weaker version of the volitional prong of the insanity defense.

The standard nosological system used by all mental health professionals and the system that is typically referenced in SVP hearings is the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; *DSM–IV*; American Psychiatric Association, 1994; text rev.; *DSM–IV–TR*; American Psychiatric Association, 2000). The *DSM–IV–TR* is almost universally relied on as the authoritative support for expert opinions on mental abnormality or personality disorder. The classification of a syndrome as a mental disorder in the *DSM–IV–TR* must be regarded as the primary standard for medical validity in the SVP context.

However VI is conceptualized, it is a diagnostic distinction that is fraught with ambiguity and unreliability. The *DSM–IV–TR* recognizes the precarious nature of diagnosing VI in its cautionary note 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" (American Psychiatric Association, 2000, p. xxiii). It remains entirely unclear how to translate VI into a psychologically meaningful construct that can discriminate among sex offenders, separating the very few with a constitutionally adequate mental disorder from all other high-risk offenders. This discrimination is not aided by mechanistic (actuarial) assessment. As Jackson, Rogers, and Shuman (2004) observed, ". . .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" (p. 126). In offering a particularly SVP-relevant example, Jackson et al. (2004) pointed out that the "mere presence [of a paraphilia] is not sufficient to meet the legal criteria [for commitment]. The paraphilia must also impair the offender's ability to control his behavior" (p. 126). The legitimacy of the SVP process appears to hinge on the scientific integrity of this diagnostic decision.

Another feature of mental disorder that we can glean from the adjudication of SVP commitments is that it helps if the disorder is a medically recognized diagnosis. On this point, however, it must be said that the courts have been somewhat coy and apparently self-contradictory. On the one hand, the courts have insisted that the definition of mental disorder in the SVP context is a legal, rather than a medical or psychological, prerogative. On the other hand, in reviewing the legality of commitments in individual cases, the courts frequently count as a legitimizing consideration the fact that the identified mental disorder was included in the *DSM*. In *Hendricks*, for example, the Supreme Court pointed out that, "The mental health professionals who evaluated Hendricks diagnosed him as suffering from pedophilia, a condition the psychiatric profession itself classifies as a serious mental disorder" (*Kansas v. Hendricks*, 1997, p. 360).

The contradiction may be more apparent than substantive. Most likely, the courts will insist on some sort of medical validity for a diagnosis. Otherwise, the use of the civil commitment form for liberty deprivation appears illegitimate. Medical validity, however, itself, is not sufficient for commitment. Indeed, the *DSM–IV–TR* itself warns that its clinical diagnoses are "not sufficient to establish the existence for legal purposes of a mental disorder. . ." (American Psychiatric Association, 2000, p. xxiii). The law seeks to carve out a subset of those who are

in some medical sense mentally disordered, and the boundaries of that subset are legally defined.

Keeping in mind the statutory requirements that the mental disorder must be coupled with proof of dangerousness and must render the individual incapable, or nearly incapable, of controlling his dangerous, impulsive behavior, we are left with a number of reasonable, potential diagnostic candidates in the *DSM–IV–TR*. Most notable are (a) impulse control disorders, (b) V Code for adult antisocial behavior, (c) mood disorders and attention-deficit/hyperactivity disorder, combined subtype, (d) sexual disorders, and (e) personality disorders.

Impulse Control Disorders

Impulse control disorders include kleptomania, pathological gambling, pyromania, trichotillomania, and intermittent explosive disorder. Although the impulse control disorders enumerated in the *DSM–IV–TR* certainly address the second requirement of impaired self-control, all fail the first requirement (i.e., there is no empirically established link between any of these disorders and sexual dangerousness).

V Code for Adult Antisocial Behavior

The V Codes are for conditions not attributable to a mental disorder; hence they would be excluded on that ground.

Mood Disorders and Attention-Deficit/Hyperactivity Disorder

Mood disorders and attention-deficit/hyperactivity disorder have been reported as prevalent among males with paraphilic disorders (Grant, 2005; Kafka & Hennen, 2002; Kafka & Prentky, 1998), including pedophiles (Raymond, Coleman, Ohlerking, Christenson, & Miner, 1999), as well as other sexual offender subtypes (Ahlmeyer, Kleinsasser, Stoner, & Retzlaff, 2003; Dunsieth et al., 2004). Attention-deficit/hyperactivity disorder, combined subtype, has been associated with conduct disorder (common in adolescent sexual offenders; Kavoussi, Kaplan, & Becker, 1988) and is a precursor to antisocial personality disorder (APD) in adult sexual (and nonsexual) offenders.

Mood disorders, including dysthymia, major depression, and bipolar disorder, are associated with antisocial behaviors, impulsive aggression, and violence, that is, VI (O'Connor, McGuire, Reiss, Hetherington, & Plomin, 1998). In addition, mood disorders and attention-deficit/hyperactivity disorder have been reported to be specifically associated with sexual appetitive disinhibition or hypersexuality (Kafka & Hennen, 2003). This disinhibition or dysregulation, which includes dimensions of sexual preoccupation and the increased frequency of enacted sexual behaviors, can be specifically associated with sexual offending (Kafka, 2003).

Sexual Disorders

Although, in theory, any of the sexual disorders and personality disorders could satisfy the statutory requirement as long as there is some defensible connection to risk of reoffense, there are only two sexual disorders (pedophilia and paraphilia not otherwise specified [NOS]—nonconsent) and two personality

disorders (APD and personality disorder NOS or mixed) that are frequently used. We address each of these in turn.

Sexual disorders: Paraphilias. *Paraphilias* are fantasies, urges, and behaviors that reflect atypical, nonnormative, or deviant expressions of sexual gratification. To be classified as a paraphilia, according to the *DSM–IV–TR*, the behavior must exhibit three elements: first, there must be a clearly specifiable deviant mode of sexual gratification; second, there must be evidence of a pattern of arousal in response to this deviant mode of gratification that is recurrent and intense; and third, the behavior must have persisted for at least 6 months. All three elements are essential, regardless of how infrequent or unusual the paraphilia is.

The diagnostic validity and the operational criteria used to diagnose paraphilic disorders, including pedophilia, remain controversial (Polaschek, 2003), despite the common use of these diagnoses in incarcerated populations of sexual offenders. For example, there are no structured clinical interviews or rating instruments for the diagnosis of specific paraphilic disorders that have been adequately tested for reliability and validity (Miller, Amenta, & Conroy, 2005).

The *DSM–IV–TR* includes, as examples, a number of the more commonly observed paraphilias: exhibitionism, voyeurism, fetishism, transvestic fetishism, frotteurism, pedophilia, sexual masochism, and sexual sadism. Of these paraphilias, there are only three—pedophilia, sadism, and possibly frotteurism—that potentially satisfy the dangerousness nexus requirement.

There are numerous other less frequently observed paraphilias. The *DSM–IV–TR* attempts to cover these additional atypical expressions of sexual arousal with the catchall diagnosis of paraphilia NOS. The *DSM–IV–TR* offers a variety of examples of paraphilias that might fall into this NOS category, such as necrophilia (dead bodies), zoophilia (animals), klismaphilia (enemas), coprophilia (feces), and urophilia (urine). There are numerous others, perhaps as many as one hundred that have been catalogued by Money (1986). In each instance there is a unique, highly distinctive stimulus or mode of expression that is sexually arousing.

Paraphilia—Pedophilia. Although pedophilia, unquestionably, is a diagnosis that, in appropriate cases, can satisfy the required elements for a mental disorder, the universal application of the diagnosis of pedophilia for all child molesters fails to differentiate among them, producing one highly heterogeneous group, all classified as pedophiles. Enumerated problems with this diagnosis may be found elsewhere (O'Donohue, Regev & Hagstrom, 2000).

Paraphilia NOS—Nonconsent. Because the *DSM–IV–TR* offers little explicit diagnostic guidance with respect to rapists, it has become a common practice among some examiners to apply a newly coined diagnosis, paraphilia NOS—nonconsent (Abracen & Looman, 2006; Becker, Stinson, Tromp, and Messer, 2003; Levenson, 2004; Zander, 2005). Becker et al. (2003) reported that over half of their sample of 120 SVPs (56%) were diagnosed with paraphilia NOS. As Miller et al. (2005) observed,

> Numerous evaluators have utilized the diagnosis "paraphilia not otherwise specified" to apply to rapists. However, the definition of this appellation is so amorphous that no research has ever been conducted to establish its validity. How such

> a diagnosis would differentiate a class of rapists who suffer from a mental abnormality is very unclear. (p. 39)

With the exception of those rapists who can be properly diagnosed with sexual sadism, representing only 2% to 5% of rapists (Quinsey, Chaplin, & Varney, 1981), a substantial proportion of rapists do not meet the criteria for any paraphilia (Miller et al., 2005). Another way to examine this question is to study known paraphilics. Abel, Becker, Cunningham-Rathner, Mittleman, and Rouleau (1988) obtained extensive, confidential self-reports from 561 outpatient paraphilics, 60% of whom were assaultive. These 561 paraphilics reported a total of 195,408 victims. Of these 195,000 victims, only 0.2% were victims of rapists, less than the proportion of victims of zoophiles (0.5%).

The apparent justification for such a diagnosis comes from the language in the *DSM–IV–TR*'s definition of paraphilia referring to "children or other nonconsenting persons" (American Psychiatric Association, 2000, p. 566). *DSM–IV–TR*'s recognition that some paraphilias may be acted out with nonconsenting partners has been twisted to suggest that the presence of nonconsent, by itself, is sufficient to diagnose a paraphilia. To be true to the *DSM–IV–TR* criteria, however, such a diagnosis would require that the nonconsent itself be the specific stimulus for the intense sexual urges (American Psychiatric Association, 2000).

Sexual arousal is rarely, if ever, associated simply with a partner saying "no." In the loose construction of paraphilia nonconsent, the underlying premise is that the offender is cognizant of what the victim wants and is aroused by the knowledge that the victim does not want sex. Arguably, however, the substantial majority of sex offenders care only about what they want, at least at the time of the sexual assault, and a matter such as victim nonconsent fails to register in any important way other than as an obstacle to achieving compliance. Indeed, because by definition all victims of sexual crimes are nonconsenting, all sexual offenders with multiple offenses (spanning at least 6 months) could be diagnosed with paraphilia NOS—nonconsent. To the extent that this category becomes a wastebasket for sex offenders, it is taxonomically useless (i.e., it provides no discrimination).

Money (1986) might well have argued in favor of such a diagnostic category. Some rapists qualify in Money's terms for paraphilic rape, or what he called "raptophilia" (Latin) or "biastophilia" (Greek; p. 54). Money defined paraphilic rape as a "syndrome, in which the stark terror, screaming, yelling, and struggling" of the victim is integral to the offender's sexual arousal (p. 54). In other words, the deviant stimulus, in Money's words, is the "stark terror" of the victim. This goes well beyond mere nonconsent. If "stark terror, screaming, yelling, and struggling" were sexually arousing, it satisfies the diagnostic criteria for sexual sadism (p. 54). Indeed, Money (1995) maintained that paraphilic rape was included in sexual sadism.

Spitzer, one of the architects of the *DSM–IV–TR*, convened a Paraphilia Subcommittee in the mid-1980s to make recommendations for changes for the *DSM–III* revision. The Paraphilia Subcommittee's recommendation of a new paraphilia called "paraphilic coercive disorder" was not endorsed by the American Psychiatric Association. The rationale and the thinking of the subcommittee, however, are of interest. According to Spitzer (personal communication, January

23, 2004), it was the subcommittee's understanding that this proposed paraphilia applied to only a small subgroup of sex offenders and, most important, that the diagnosis required evidence "that the coercive element of the sexual assault was sexually arousing" and that sexual sadism was ruled out as a preferable diagnosis.

For nonconsent to be a paraphilia, expressions of verbal or physical resistance would have to evoke intense sexual arousal, or the offender would have to evidence a pattern of consistently selecting victims incapable of giving consent (e.g., drugged, unconscious, dead, mentally retarded, etc.). A rule-in for such a diagnosis might well be that consenting sexual partners are inhibitory to sexual arousal. Currently, no such criterion exists.

Personality Disorders

APD. In *Foucha v. Louisiana* (1992), the Supreme Court hinted broadly that APD is not, even when combined with dangerousness, a sufficient predicate for civil commitment. Writing for the Court, Justice White warned that if the state could civilly commit someone just for having "an antisocial personality that sometimes leads to aggressive conduct," then it could routinely commit prisoners nearing release from prison because many of them, like *Foucha*, could be diagnosed as having antisocial personalities and being dangerous (p. 85). For Justice White, this would be unacceptable, because it would be "only a step away" from substituting "confinements for dangerousness for our present system."

It is possible to argue, as did Cornwell (1998), that anyone with APD could be involuntarily committed using the mental abnormality standard as it was applied to *Hendricks*. If the standard is a personality disorder that leaves individuals impaired in their ability to control their dangerous impulses, then APD could arguably qualify. As Cornwell (1998) noted, "states may well argue that APD is, generally speaking, a sufficient ground for psychiatric commitment because it is characterized, *inter alia*, by chronic impulsivity, irresponsibility, aggressiveness, and unlawful behavior" (p. 397). Such a definition for constitutionally adequate mental disorder, however, would cast an exceedingly broad net—anywhere between 50% and 75% of the prison population might qualify for civil commitment on the basis of an APD diagnosis—and thus violate a key premise of the Supreme Court's holdings, that SVP commitments must target a small subgroup that is somehow distinguished from the ordinary dangerous recidivist.

This analysis notwithstanding, one could envision a constellation of antisocial traits that fulfill the requirements for APD and come closer to satisfying the more particularized constitutional requirements for civil commitment. For example, consider an individual for whom the core APD traits include a reliably increased likelihood of sexual assault (as opposed to any other nonsexual criminal act). Such traits might well include elements of hostile or negative masculinity, "con" attitudes, misogynistic attitudes, sexual entitlement, sexual preoccupation, or other indices of current hypersexuality (e.g., Knight & Sims-Knight, 2003; Malamuth, 2003). One could argue that this particular manifestation of APD satisfied the nexus with risk and hence met that statutory standard for mental disorder.

Psychopathy. Psychopathy, a heuristic diagnostic construct distinct from, but highly associated with, APD, is not an acknowledged psychiatric diagnosis in

the *DSM–IV–TR*. Psychopathy, however, is consistently associated with violent behavior and to a lesser degree with sex offender recidivism. Although the use of this diagnosis would deviate from the legal norm by relying on nosology not presented in the *DSM–IV–TR*, psychopathy is indisputably a strong predictor of general and violent recidivism (e.g., Serin, Malcolm, Khanna, & Barbaree, 1994). Thus, a reliable diagnosis of psychopathy, despite its absence in the *DSM–IV–TR*, may constitute a defensible mental abnormality.

There is complex evidence, however, for the predictive relationship of psychopathy with sexual recidivism. In one study of rapists, only the psychopathic offenders who were also sexually deviant according to Item 1 on the Sexual Violence Risk-20 (Boer, Hart, Kropp, & Webster, 1997) evidenced high sexual recidivism rates (82%; Hildebrand, de Ruiter, & de Vogel, 2004). Among psychopathic, nondeviant sex offenders, the rate dropped to 25% (lower than the deviant, nonpsychopathic group; 30%). Serin et al. (1994) found that the relationship between psychopathy and sexual deviance (assessed by penile plethysmograph, or PPG) was stronger for child molesters ($r = .42$) than for rapists ($r = .32$). Although Hanson and Morton-Bourgon (2005) reported that sexual recidivism was predicted by the Hare Psychopathy Checklist (Hare, 2003) with a Cohen's d value of 0.29, a d of 0.29 based on 13 studies is regarded as small (J. Cohen, 1988). Consistent with the literature, Hanson and Morton-Bourgnon (2005) reported a Cohen's d value of 0.58 when psychopathy predicted violent (nonsexual and sexual) recidivism.

When an appropriate clinically defensible diagnosis is not available, defendants are occasionally shoehorned into a legitimate diagnosis by ignoring elements required for classification, such as ignoring the conduct disorder before age 15 requirement for a diagnosis of APD, or worse, placing defendants into newly created categories, such as paraphilia NOS—nonconsent, that have no known empirical support. Force-fitting a diagnosis or creating a new *DSM* diagnosis to justify commitment is clearly unethical for psychologists (Ethical Standards, 9.01; American Psychological Association, 2002, p. 1071). As Behnke (2005), American Psychological Association Ethics Director, noted, "Principle A: Beneficence and Nonmaleficence, exhorted psychologists 'to benefit those with whom they work and take care to do no harm.' Promoting welfare and safeguarding from harm are thus values central to our profession. Rendering a diagnosis has direct relevance to each" (p. 80). Behnke also observed, "In few areas of practice does a psychologist exercise greater authority and influence than to render a diagnosis, for in so doing the psychologist comes to know and convey information that may profoundly affect that individual's life" (p. 80). This is patently the case in SVP civil commitment hearings, wherein the diagnosis is tied directly to indeterminate deprivation of liberty.

In addition, this issue raises a profound scientific concern. Classification, as Meehl (1996) noted, "is a problem in applied mathematics" (p. 266). At least in the world of hard science, classification reveals order and brings clarity. For the resulting categories or subgroups to have any meaning, there must be a common understanding of who belongs in the group, how to put people into the group, and once constituted, whether the newly formed group is valid. When new categories that have never been subjected to empirical scrutiny are relied on, all of the precepts of good science are violated, from the most elementary (can cases be

assigned to this new category reliably) to the more recondite (what does assignment to this category inform us about reoffense risk). The third prong of the SVP laws is reasonably clear: The mental abnormality must be of a kind that increases the likelihood of a reoffense. A newly created category would have no empirical track record providing evidence for such a linkage. From the court's standpoint, diagnoses that have no empirical foundation and no guidance with respect to decision making quite effectively undermine, or eliminate, the precise basis for the medical authority of a system such as the *DSM*. Perhaps worse, we are conferring on unvalidated diagnoses the presumptive medical authority of the *DSM*.

Opining about a defendant's diagnosis without having ever met the defendant also appears to violate the code of ethics for psychologists (9.01[b]), which states that "psychologists provide opinions of the psychological characteristics of individuals only after they have conducted an examination of the individuals adequate to support their statements or conclusions" (American Psychological Association, 2002, p. 1071). When an examination is not conducted, psychologists must explain why. In the case of a liberty-interest, high-stakes matter such as an SVP hearing, failure to conduct a proper examination would certainly seem to be of questionable ethicality.

Moreover, from the standpoint of the court, the examiner cannot be examined or cross-examined with respect to the proper fit of an undisclosed diagnosis. The court can only accept on faith that the defendant has what the examiner claims he has and that what he has is statutorily relevant. Failure to disclose that which serves as the basis for one's expert opinion should go to admissibility.

Over 3 decades ago, Blashfield (1973) demonstrated that coverage and reliability of diagnoses were inversely related for subtypes of schizophrenia. The inverse relation between reliability and coverage is not unique to schizophrenia. Reliability increases as categories are narrowly and precisely defined, though coverage will be sacrificed. When categories are broad, imprecisely defined, with vague criteria for classification, coverage can be maximized, but reliability will be poor. Such is clearly the case with "wastebasket" diagnoses, such as paraphilia NOS—nonconsent. The category is so broad as to embrace virtually all sex offenders with an offense history lasting 6 months or longer. Hence, coverage is almost complete. On the other hand, absent any clear definitional criteria for classifying nonconsent as a paraphilia, reliability will be very poor.

Assessment of Risk

Courts are asking in more detail what the legal standards for dangerousness mean and what sort of evidence is legally available, and ought to be required, to prove those standards. In large measure, the legal challenges in these more recent cases have addressed the use of actuarially derived (as distinct from clinical) risk assessments. Clinical judgments of dangerousness, that is, judgments that ultimately rest on the arbitrary opinion of a mental health professional, are a routine part of the judicial landscape. In contrast, *actuarial risk assessment* (ARA), empirically derived mechanical rules for combining information to produce a quantitative estimate of risk, is uncommon in the legal arena and has provoked a reaction. Critics of ARA have focused their objections on the admissibility of

ARA-derived expert testimony. Pointing to a variety of shortcomings, they argue that the relatively new ARA techniques do not merit admissibility under prevailing legal standards. These challenges have met with mixed success (Janus & Prentky, 2003).

Dangerousness is one of two constitutionally required components of civil commitment and the unambiguous justification for the civil commitment of sex offenders (i.e., "we are protecting society from the most dangerous offenders"). Although preventive detention would be legally and ethically problematic even with perfect knowledge about the future, the imperfection of risk assessment exacerbates constitutional and ethical concerns because it raises the likelihood that nonrecidivists and low-risk individuals will be among the group suffering long-term loss of liberty. The same is true for the more utilitarian concerns about resource allocation and efficacy. The central justification for spending vast sums of money on SVP programs is that the most dangerous offenders are incapacitated. Public policy is not well served if, because of inaccurate assessments of risk, extraordinary resources are squandered on those who pose low risk to public safety.

As a result, the demand for specialized risk assessments has been urgent, producing a "cottage industry of forensic psychologists" (Grisso, 1987, p. 831) and vigorous empirical efforts to develop actuarial and other structured approaches to supplement the traditional clinical assessment of the risk posed by these most dangerous sex offenders.

What Makes Risk Assessment So Difficult?

Risk assessment has been aptly referred to as "the mother of all uncertainties" (Bailar & Bailer, 1999, p. 273). Uncertainty in behavioral risk assessment can be traced to two main sources. First, humans (experts or otherwise) have a limited ability to assess future risk of harmful behavior. Clinical risk assessment is, by definition, an exercise in human judgment. The susceptibility of human judgment to error has been the subject of considerable empirical scrutiny (cf. Garb, 1998). Numerous sources of error in clinical judgments have been described (e.g., Garb, 1998; Grove, Zald, Lebow, Snitz, & Nelson, 2000) and will not be addressed here. In part, limitations in clinical judgment stem from the fact that the future is inherently unknowable, and in part, from inherent shortcomings in human judgment. Although the former limitation is inescapable, the limits of human judgment can, to some extent, be ameliorated through empirical research. Thus, the quality of risk assessment is variable, and improvement is not only possible, it is critical.

We must not ignore, moreover, the natural human proclivity to overvalue vivid or emotionally laden information. Vivid detail, for experts as well as juries, can be misleading and frequently detracts from the objectivity and reliability of an opinion. Vivid detail can have a potent psychologically biasing effect that is explained by the availability heuristic. As Jackson, Rogers, and Shuman (2004) demonstrated experimentally, "certain types of information, such as emotionally-evocative victim statements, can bias the professional conducting the forensic evaluation" (p. 125). Focusing on highly evocative, vivid information can lead to an overattribution of predictive efficacy and erroneous opinions. This form of bias

is distinguished here because it appears to be especially common, given the nature of the behaviors (i.e., facts) under scrutiny in SVP evaluations and hearings.

Notably, a major source of problems in risk assessment lies in the legal system. The risk thresholds for invoking SVP commitments are vague, and courts have failed to set standards that can be implemented reliably, relying instead on unoperationalized terms, such as "likely." Frequently, the liberty deprivation decision boils down to a credibility judgment between the clinical assessments of two (or more) competing expert witnesses. As a result, there is no assurance that risk thresholds are uniform or that risk assessments are performed using equivalent standards and procedures. Thus, the legal system fails to take advantage of the increase in clarity and reliability conferred by science and exacerbates the weaknesses of the decision-making process by inviting arbitrariness to join the mix.

Mechanistic Assessments of Risk

Prevailing wisdom dictates that our ability to predict violent behavior using traditional clinical methods falls well below a threshold of accuracy that justifies the use of such predictions in legal proceedings (e.g., Cocozza & Steadman, 1976; Ennis & Litwack, 1974; Ewing, 1983, 1985, 1991; Megargee, 1981; Monahan, 1981).

In response to the increasing need for clinicians to assist with judgments about sexual dangerousness under the new statutes, researchers began working in earnest to develop reliable, valid ARA instruments. This research focusing on risk assessment was a response, at least in part, to the aforementioned widespread doubts about the ability of mental health professionals to predict dangerousness.

The pace of developments issuing from the demands imposed by this recent wave of civil commitment laws has been rapid, with empirically driven revisions to, and support for, existing static scales, a recent wave of scholarship on dynamic risk factors (Beech, Friendship, Erikson, & Hanson, 2002; Douglas & Skeem, 2005; Hanson & Harris, 2001; Hudson, Wales, Bakker, & Ward, 2002; Thornton, 2002), and adaptations of existing scales for special populations, such as juveniles (Epperson, Ralston, Fowers, DeWitt, & Gore, 2006; Prentky, Harris, Frizzell, & Righthand, 2000; Prentky & Righthand, 2003; Righthand et al., 2005; Worling, 2004; Worling & Curwen, 2001).

Most scholars have concluded that the predictive efficacy of actuarial methods of risk assessment is superior to clinically derived assessments of risk (e.g., Dawes, Faust, & Meehl, 1992; Grove & Meehl, 1996; Grove et al., 2000; Meehl, 1954; Monahan et al., 2001; Swets, Dawes, & Monahan, 2000). Monahan and his colleagues observed that "the general superiority of statistical over clinical risk assessment in the behavioral sciences has been known for almost half a century" (Monahan et al., 2001, p. 7). In large measure, the superiority of ARA arises from the elimination or reduction of the many sources of error referred to earlier. To be sure, existing ARA scales are far from perfect, and some have significant methodological problems (Litwack, 2002). There is no reason to believe, however, that clinical judgments are systematically better than even the weakest of the actuarial assessment scales. Problems that are present in a poorly designed actuarial scale are likely to be equaled, or exceeded, in clinical assessments (Grove & Meehl, 1996).

ARA

Actuarial scales are developed using statistical analyses of groups of individuals (in the present case, released sex offenders) with known outcomes during a follow-up period (either arrested for, or convicted of, a new sexual offense or not identified as having committed a new sexual offense). These analyses tell us which items (predictor variables) do the best job of differentiating between those who reoffend and those who did not reoffend within a specified time period. Because some of these variables inevitably do a better job than others, these analyses can also tell us how much each item should be weighted. The variables are then combined to form a scale, and the scale is tested on many other groups of offenders (cross-validation). When the scale has been used on many samples with a sufficiently large number of offenders, the scores derived from the scale may be expressed as estimates of the probability that individuals with that score will reoffend within a specified time frame.

The accuracy of the estimate is a function of the similarity of the assessed individual to the members of the reference group that were used to derive the estimate. Consequently, estimates from ARA scales developed for adult male sex offenders are not appropriate for use with very young offenders (i.e., juveniles) and female offenders. Similarly, these estimates may not be appropriate for use with exclusive, endogamous incest offenders or older offenders (e.g., age 50 or older), depending on the relative proportion of offenders in the reference groups that are exclusive incest offenders or older offenders. Estimates may also not be appropriate for use with severely developmentally or cognitively impaired offenders. In sum, the more the individual departs from the reference group in potentially risk-relevant ways, the more unreliable and inaccurate the derived estimate is likely to be (L. J. Cohen, 1981a).

The Probative Value of Base Rates

Although we recognize the historical, seemingly intractable, debate over the proper role (if any) of base rates and other probabilistic evidence in the courtroom (e.g., L. J. Cohen, 1981b; Koehler, 1996), base rates, for better or worse, have come to assume center stage in many civil commitment hearings. Rather than base-rate data being ignored, as it traditionally has been in violence prediction, such evidence is often introduced, albeit with marginal probative value. As Koehler (1996) clearly noted, the problem is not ignoring base rates as much as failing "to consider how the ambiguous, unreliable, and unstable base rates of the real world should be used" (p. 1). In addressing the problem of ambiguous or frankly erroneous base rates, Rogers (2000) pointed to the review by Hiday (1990) of violent conduct among civilly committed inpatients. Within this relatively narrowly defined population, the base rates, according to Hiday, ranged from 7.5% to 66.7%. Closer to home, Barbaree (1997) reported that base rates for adult sexual offenders generally range from 0.10 to 0.40. In fact, some base rates may be lower than 0.10 (e.g., exclusive endogamous incest offenders or elderly offenders), and some base rates may be higher than 0.40 (e.g., career sex offenders with many offenses spanning many years).

The estimated base rate for sexual recidivism reported by Hanson and Bussiere (1998), which was based on a meta-analysis of 61 studies with a total

sample of 23,393 sex offenders, was 13.4% over 4 to 5 years. In a subsequent meta-analysis of 95 studies with a sample of 31,216 sex offenders, Hanson and Morton-Bourgon (2005) reported an aggregate sexual recidivism rate of 13.7% over 5 to 6 years. Although the (partially overlapping) samples in these meta-analyses obviously are quite large, they are a highly mixed assemblage of offenders that included nuisance (i.e., exhibitionists) along with hands-on offenders. From the standpoint of estimating base rates for evaluating the risk posed by a particular individual, these figures of 13%–14% are not overly helpful. Base rates vary considerably in a population as markedly heterogeneous as sex offenders. Moreover, collapsing across the diverse methods of data collection, data calculation, and data reporting in 61 (or 95) separate studies inevitably results in a high degree of methodological variability (Prentky, Lee, Knight, & Cerce, 1997). As Davis and Follette (2002) pointed out, "base rates representing averages across an entire population . . . may seriously misestimate the likelihood of violence" (p. 149).

This does not mean that aggregate base rates from the meta-analyses convey nothing. Aggregate base rates of 13% or 14% tell us that, as a group, the known sexual recidivism rates of sexual offenders are quite low. Swets (1992) pointed out that, "Probabilities of rare events . . . are estimated with little reliability. Benefits and costs are notoriously difficult to judge when human lives are at stake and are often in conflict when a balance must be struck between an individual's and society's concerns" (p. 525). The problems inherent in low base-rate prediction of dangerousness have been addressed numerous times (e.g., Grove & Meehl, 1996; Monahan & Steadman, 1994; Swets, Dawes, & Monahan, 2000; Wollert, 2006).

The task in an SVP hearing is to estimate the base rate for one individual, the defendant. As Meehl (1954) long ago advised, when a variety of base-rate reference groups are available, the "best class is always defined . . . (as) the smallest class . . . for which the N is large enough to generate stable, relative frequencies" (p. 22). Thus, at minimum, base-rate estimates should be drawn from reference groups that exclusively comprise (a) rapists; (b) extrafamilial child molesters; (c) exclusive, endogamous incest offenders; (d) juvenile sex offenders; or (e) hands-off offenders (e.g., exhibitionists).

We would argue, however, that even these five reference groups may not be sufficiently homogeneous to provide optimal base rate estimates. We reported (Prentky, Knight, Lee, & Cerce, 1995), for instance, that within a higher risk sample of 109 rapists, all of whom had been civilly committed as sexually dangerous, a simple dimension of lifestyle impulsivity significantly discriminated among the offenders with respect to sexual reoffense (i.e., the high impulsivity rapists were almost three times more likely to commit another sexual offense). Although it is impractical to achieve L . J. Cohen's (1981b) recommendation that the evaluee should "share all the relevant characteristics" of the reference group from which the base rate came (p. 329), we certainly have ample empirical data to identify the "relevant characteristics" and to derive base rates from groups of individuals possessing those characteristics. When considering the potential magnitude of the false positive and false negative errors in judgment that are routinely a part of the expert opinions offered in SVP trials, maximum precision in base-rate estimation should be axiomatic for the court and demanded of all examiners.

Unfortunately, incorporating an informed estimate of a defendant's base rate into a summary opinion about the risk posed by the defendant is a rare event among examiners. As Kahneman (2003) has elegantly demonstrated, there is a highly significant negative correlation between judgments of probability and base rates (−0.63) and an almost perfect positive correlation between judgments of probability and similarity (0.98), confirming "a bias of base-rate neglect in this prediction task" (pp. 708-709). Kahneman's (2003) studies support the attribute substitution hypothesis, namely "that respondents offer a reasonable answer to a question that they have not been asked" (p. 709). Thus, if the defendant looks like (similarity) a class of sex offenders who pose high risk, the judged probability must be high, regardless of the base rate.

In commenting on the insensitivity of people to base rates, Quinsey (1996) posed two responsible factors: remembering the base rate and considering the base rate important. We would qualify Quinsey's observation only by noting that remembering the base rate implies knowing the base rate. The problem in the SVP context does not appear to be relying on irrelevant information as much as failing to use relevant information, ignoring population heterogeneity, and basing conclusions on selective and erroneous base rates.

A related and relevant consideration is the regression fallacy. Regression toward the mean, or toward average, is a natural statistical correction that is observed across all domains of human performance. Failure to take regression toward the mean into account is a major source of error in judgments of violence (Monahan, 1981). In Kahneman and Tversky's (1973) terms, the representativeness heuristic explains that we tend to make nonregressive predictions that highly unusual outcomes will continue as if they were the mean (i.e., average). Reoffending sexually is a relatively unusual outcome, whereas not reoffending sexually is the average or typical outcome. We are inclined, however, to make nonregressive predictions that assume the opposite (i.e., reoffending sexually is the expected or average outcome).

The Risk-Mitigating Significance of Age

SVP legislation tends to be applied to older sex offenders because such legislation is generally applied to higher risk offenders after they have achieved a lengthy criminal record and because such legislation is most often applied after the offender has served a lengthy criminal record. Although the legislation is applied to older offenders, this legislation does not recognize important maturational changes that may mitigate risk for recidivism. There are good theoretical reasons why we expect recidivism in sex offenders to decrease with age. First, one of the most robust findings in the field of criminology is that the prevalence and incidence of criminal behavior by adults decreases steadily with increasing age (Hirschi & Gottfredson, 1983; Sampson & Laub, 2003; Wolfgang & Ferracuti, 1982). In a classic review article, Hirschi and Gottfredson (1983) documented a pattern of crime rates decreasing with age for offenders who (a) lived in different centuries, (b) came from different countries, (c) differed with respect to age and gender, (d) were at large in the community or incarcerated, and (e) committed different types of crimes. Even habitual criminals, defined by their intractability, may begin to decline in middle age (Blumstein, Cohen, Roth, & Visher, 1986). In

cross-sectional data on violent crime, including rape, compiled by the Federal Bureau of Investigation in 1980, 1994, and 2001 (Office of Juvenile Justice and Delinquency Prevention, 2004), prevalence rates increase to a peak in the mid-to-late teens and then steadily decrease over age until the end of life (as cited in Wollert, 2006).

Second, there are particularly good reasons to question the notion that sexually motivated behaviors of any type, deviant or conventional, would continue at the same levels throughout a man's middle years and into old age. Such an expectation is at variance with the known facts of human endocrinology and sexuality. Specifically, empirical studies indicate that bio-available testosterone, which is necessary or at least important in maintaining libido, peaks in early adulthood and thereafter decreases through the remainder of the lifespan (e.g., Denti et al., 2000). There is also evidence that testosterone receptor sites may become less sensitive with age, so that the threshold concentration of testosterone necessary to maintain libido may increase with age (e.g., Baker & Hudson, 1983). Furthermore, there is a general decline in male sexual behavior through the lifespan, including intercourse and masturbation (e.g., Rowland, Greenleaf, Dorfman & Davidson, 1993); and sex offenders show substantial reductions in the strength of sexual arousal through the lifespan, from the mid-teens to old age (Barbaree, Blanchard, & Langton, 2003; Blanchard & Barbaree, 2005; Kaemingk, Koselka, Becker, & Kaplan, 1995).

Accordingly, we would expect to see reductions in recidivism in sex offenders as they age. There are four scientific studies that have specifically examined changes in recidivism rates in sex offenders over a large range of age-at-release, and these studies confirm substantial reductions in recidivism over the lifespan (Barbaree et al., 2003; Fazel, Sjöstedt, Långström, and Grann, in press; Hanson, 2002; Thornton, in press).

The samples of sex offenders that have been used in the development and validation of the actuarial risk assessment instruments have included a preponderance of younger offenders. The average years of age at release in these samples are in the mid-30s. Therefore, it could be reasonably argued that the use of the actuarial instruments is inappropriate in estimating risk in the aging sexual offender. Professional standards guiding the use of psychological tests warn against the use of tests if such use may be discriminatory on the basis of age, race, culture, and so forth. Clearly, if recidivism risk decreases with age and if the actuarial instruments estimating risk were developed with young sex offenders, then the use of these instruments with older offenders could be considered to be discriminatory. On this basis, it could be argued that actuarial instruments should not be used with older offenders and to do so might be considered to be a breech of the standards of professional practice.

In response to these concerns, Hanson (2005) studied the validity of the Static-99 (Hanson & Thornton, 1999) with older sexual offenders, using data from eight samples ($N = 3{,}425$) drawn from Canada, the United States, and the United Kingdom, and followed them after release from prison for an average of 7 years. Results indicated that when controlling for Static-99 scores, recidivism risk increased slightly between ages 18 and 30 years, then declined thereafter, such that after age 41, recidivism risk was lower than at any previous age, and rates declined thereafter with further increases in age. Average recidivism rates steadily

declined from 14.8% in offenders under age 40, to 8.8% for offenders in their 40s, 7.5% for offenders in their 50s, and 2% for offenders age 60 or older. Among offenders age 60 or older, the sexual recidivism rates were low even for those who scored in the *moderate–high* range (4.8%) and the *high* range (9.1%) on the Static-99. The interaction between age and Static-99 was not significant, meaning that the amount of age-related decrease in risk was the same for all levels of risk. Although the overall recidivism rates were lower for the older offenders, Static-99 was equally effective in ranking the relative risk of both the younger and older offenders. Therefore, the Static-99 may be valid in older offenders in ranking offenders as to risk, but it is clear that the risk levels (% likelihood) suggested by the Static-99 are too high for older offenders. Hanson (2005) concluded that for offenders under 40, there was little justification for using age to reduce the expected Static-99 recidivism rates. For offenders over 60, Hanson acknowledged that the Static-99 substantially overestimates expected risk. In the 40–60 age range, Hanson indicated that some adjustment may be necessary but declined to offer specific advice as to how to do it.

In sum, most individuals subject to civil commitment under the SVP laws, having already served their criminal sentence, are older. Their ages usually fall within the age range of 40–60, in which some adjustment to risk based on age is appropriate, but in which there are no commonly accepted methods for examiners to follow in adjusting risk. One response would be simply to refrain from using actuarial instruments for older offenders. Alternatively, examiners could substitute structured clinical judgment, such as the SVR-20 (Boer et al., 1997), for ARA. In structured clinical judgment, offenders are ranked in broad risk categories (low, moderate, high) on risk-relevant dimensions but an estimated probability of reoffense is not provided.

A second approach is to adjust recidivism risk downward after age 40 on the basis of empirically derived criteria. If we assume a linear decrease in recidivism rate after age 40 (most of the empirical studies would support such an assumption), adjustments to actuarially derived risk could be made depending on two additional factors: the number of years the individual's age-at-release exceeds 40 and the rate of decline in recidivism rates on an annual basis. Hazard rates reflect the decline in recidivism rates as a proportion of the recidivism rate in Year x, in Year x + 1, etc. Hazard rates have been reported by the empirical studies cited above. Barbaree et al. (2003) reported a hazard rate of .95 when controlling for actuarial risk. Hanson (2005) and Thornton (in press) reported hazard rates of .98 when controlling for risk using the Static-99 and the number of previous offenses, respectively. Therefore, the most conservative adjustment would use a hazard rate of .98, indicating a reduction in recidivism risk of approximately 2% per year after age 40.

For example, if actuarial assessment estimates that an offender's risk for recidivism over the next 5 years is 40%, and he is 52 years of age (12 years over 40) at his expected release date, then his age-adjusted estimate or risk to reoffend sexually would be calculated as follows: Age-adjusted risk (%) = [Actuarially derived risk (%) $\times$ $.98^{\text{Number of years over 40}}$], and in the present example, 40% $\times$ $.98^{12}$ = 31.39%. In this calculation, each subsequent estimate or risk is calculated as .98 of the risk given for the previous year.

Methodologies for Assessing Risk

Examiners are occasionally asked to opine about the approach or method used in conducting their evaluation (cf. Hanson, 1998), such as guided clinical judgment, research guided clinical judgment, or clinically adjusted actuarial approach. Thus, it seems only proper to recognize formally with convenient names two methods that reflect bad science.

H. G. Wells method of risk assessment. The H. G. Wells method of risk assessment employs the services of a time machine to transport the examiner back in time to events as they were occurring 20 or 30 years ago. Having accurately captured the risk posed at that time, the examiner reenters his time machine and returns to present day, risk in hand. Translated, when we estimate current risk based solely on long distant events, often for individuals that have been in prison for decades, the result is likely to be inaccurate.

Long distant events are indeed informative; they tell us about how individuals behaved in the past, and to a limited extent that information is useful today. It would be erroneous, however, to assume that risk based on past conduct accurately reflects current risk (Douglas & Skeem, 2005). When an individual has been down for many years, change, whether it is based solely on the passage of time (i.e., aging process) or on potentially risk-relevant considerations (e.g., the inmate got married while in prison, the inmate has been in treatment while in prison, the inmate has acquired significant vocational skills while in prison, the inmate attended school and acquired a degree while in prison, a parent or significant other has died while the inmate has been in prison, the inmate's health has deteriorated while in prison, etc.), must be taken into consideration. These considerations are not necessarily mitigatory. Change may also reflect increased or sustained risk (e.g., the inmate's prison adjustment has been very poor, marked by numerous disciplinary reports, fighting, sexually inappropriate or assaultive behavior, etc.). In sum, from a scientific standpoint, it is irresponsible to ignore obvious risk-relevant changes that may have occurred over the past 10, 20, or 30 years.

Cherry picking method of risk assessment. The cherry picking method refers to selectively harvesting information that supports and confirms an a priori opinion. Putative risk predictors are screened, either by formal reference to the empirical literature or by informal reference to one's memory, and those that pertain to the case at hand are selected. The flaw of such an approach is the reliance on individual predictors, each of which may have no association with recidivism or, at best, a weak, nonsignificant association with recidivism.

With good reason, Hanson and Bussiere (1998) admonished readers of their meta-analysis: "The predictive accuracy of most of the variables was also small (.10–.20 range), and no variable was sufficiently related [to recidivism] to justify its use in isolation" (p. 358). Many of these individual items, commonly referenced in SVP hearings as support for commitment, have extremely low, or in some cases negative, correlations with reoffense. Both meta-analyses, Hanson and Bussiere (1998) and Hanson and Morton-Bourgon (2005), found that a negative clinical presentation was unrelated to sexual recidivism. The correlations from Hanson and Bussiere (1998) and Cohen's *d* values from Hanson and Morton-Bourgon (2005) were very small for victim empathy, denial, and low motivation

for treatment. Poor progress in treatment, assessed in Hanson and Morton-Bourgon (2005), was also unrelated to recidivism. Although Hanson and Bussiere reported a slightly higher correlation (.17) with "failure to complete treatment," by the authors' own guide a correlation of .17 is considered small. If an item with a correlation of .17 was used in isolation to predict sexual recidivism, the usefulness of the item would be marginal at best. Additionally, from a practical perspective, the use of "failure to complete treatment" as a predictor implies that there is a clearly delineated procedure for completing treatment. Such an expectation would seem unwarranted for most of the SVP treatment programs.

Adjustment of ARA-Derived Risk Assessments

Extending risk estimates. Extending or projecting risk estimates beyond the time period reported for the particular scale is not permissible. Essentially, the examiner is making the assumption that the shape of the failure curve will remain the same if extended with additional data. In other words, if the reported estimates stop at 15 years and the examiner wants to extrapolate to 25 years, the examiner is assuming that the shape of the failure curve would remain essentially the same if another 10 years of follow-up data had been gathered. Not only is that an untestable and possibly erroneous assumption, the magnitude of the potential error will vary according to the defendant that the adjustment is applied to. For example, if the defendant is 45 years old, the extension may take the defendant to age 55, placing him in a bracket of age-mitigated risk reduction. In this case, the extrapolation is likely to be incorrect even if the failure curve remains essentially the same.

Adjusting ARA scale estimates for an individual's estimated base rate. Risk estimates from ARA scales are based on the known base rates for samples (reference groups) used to derive the estimates. What if the estimated base rate for the defendant differs markedly from the base rates of the reference groups? If the estimated base rate that appears most appropriate for the defendant is higher, the ARA scale will underestimate risk. If the estimated base rate is lower, the ARA scale will overestimate risk. In these cases, can the scale estimates be adjusted for the defendant? The answer is, in principle, "yes, as long as the adjustment is empirically based." The insurance industry, as an example, uses time-varying factors to modify or adjust risk for clients all the time. In the insurance case, the time-varying factors are built into the mechanistic formula. In the SVP case, the integration of time-varying, dynamic risk factors into static risk is in its most puerile state. One example is the empirically derived modification of static risk with the time-varying factor of age, using hazard rates from Cox regression.

Adjustment of ARA scores. *Adjustment* is a familiar and somewhat unfortunate term that is commonly used by examiners in SVP hearings, often to give scientific legitimacy to clinical judgment. Examiners opine that they used an adjusted actuarial method, which included the selective identification of those risk predictors from a scale that apply to the case at hand (e.g., an examiner may report three or four risk factors taken from the Static-99 without completing the full scale and without identifying the origin of the risk factors). An alternative approach is to adjust the results of the ARA scale using presumptively risk-relevant contemporaneous information. Both methods are highly problematic. In the former case,

ARA provides little more than empirical window dressing for clinical judgment. In the latter case, the process of adjustment may constitute little more than adding ARA input to clinical judgment. Because there is no standardized, uniform procedure for adjusting static risk with time-varying factors, we are vulnerable to examiner bias and subject to the same opacity that often characterizes clinical judgment in the courtroom. The net result is that, once again, we may be dressing up clinical judgment with actuarial science.

Treatment Under Sexually Dangerous Person or SVP Civil Commitment Law

Although all of the SVP laws profess a treatment purpose, it seems to be distinctly subordinate to the public safety–oriented purpose of incapacitation. The Supreme Court, although acknowledging that treatment plays some key constitutional role in legitimizing SVP laws, has so far telegraphed a wide tolerance for meager treatment efforts. The Court apparently envisions a (limited) role for treatment (Janus, 1997, 1998). The Court seemed to adopt a middle ground position, indicating that the state may be obliged to provide treatment that is available for disorders that are treatable (cf. Janus, 1998). By the same token, the Court clearly rejected the proposition that effective treatment is required to justify civil commitment. In the Court's SVP jurisprudence, the role played by treatment in civil commitment is quite ambiguous. If we conclude that *Hendricks* is a police power commitment, then in police power commitments, treatment may be a right of those committed, but it is not a justification for commitment. The principal objective is to protect society, not to rehabilitate.

Perhaps responding to this weak judicial concern, the implementation of these laws has resulted in two areas of marked concern: providing patently suboptimal or inadequate prison-based clinical programs under the guise of treatment and rendering indefensible judgments about change or lack of change as a function of treatment, to justify continued detention.

Demonstrating Treatment Efficacy

Although numerous studies have demonstrated that cognitive–behavioral interventions can reduce sexual recidivism in the general population of sex offenders, until a sufficient number of those who are civilly committed are released into the community, it will be difficult to ascertain the efficacy of treatment in these SVP programs. With limited exceptions, very few of those committed are being released. Those discharged or released range from 0 in North Dakota, New Jersey, and Iowa to 1 in Minnesota, 4 in Massachusetts, 6 in Missouri, and fewer than 20 in Washington, Kansas, Illinois, and Florida (Lieb & Gookin, 2005). The only states that have released a sufficient number of committed offenders to permit a follow-up are Arizona (221), California (67), and Wisconsin (56).

Beyond the empirical demonstration of efficacy, it must be acknowledged that there are numerous roadblocks to effective treatment in civil commitment settings. First, the vast majority of those committed have not volunteered to be in the program. Even if a relatively few, selected individuals are motivated to participate, the programs are often dominated by angry, litigious individuals who are in

constant direct and indirect conflict with the staff. The staff in turn are often confused about their role. Is it their job to make an honest attempt to treat these individuals in the most effective way possible, thus enhancing their chances of release? Or alternatively, is it their first responsibility to help ensure that their charges continue to remain committed as SVPs? The conservative approach is to set impossibly high standards for release and then to maintain that individuals will be released as soon as they complete treatment. In this catch-22 world that is absent of specific concrete goals and objectives leading inexorably to program completion, as long as an individual is committed, his treatment team will prepare a treatment plan for the coming year, and for as long as the individual has recommended tasks listed on his treatment plan, he has not completed treatment. This conundrum goes to heart of this type of treatment for SVPs. Are therapists clinical babysitters hired to dress up the program or are they functional change agents? One litmus test is confidentiality. When there is essentially no true confidentiality and clinicians require full disclosure polygraphs and document in reports any disclosures that may be used in court to ensure continued detention, the former is the case. When confidentiality is protected and clinicians play no adverse role in the adjudication of release, the latter is the case. There is certainly precedent for this latter approach. Treatment is routinely offered in prison with progress being reported to parole boards without turning over treatment records. Prison administrators have little or no interest in whether an inmate is paroled or not. That is simply not their concern. Therapists may be requested to testify by a treatment participant. Alternatively, Wisconsin has developed a system that separates treatment evaluation by the therapists from treatment progress as evaluated by an independent panel.

Effective treatment not only requires that participants acquire a number of skills by which to manage their sexual deviance but also that they be given the opportunity to practice these skills in realistic situations. Thus, some type of community transition program is mandatory. However, unless such programs already exist, as was the case in Minnesota, locating a site for such a facility is close to impossible. Thus, few of the programs have made any provision for the carefully monitored gradual release of a program participant into the community. Understandably, staff are highly reluctant to risk a reoffense by a sexual predator on a pass to the local community. However, this is the only way that a treated offender can demonstrate that he has internalized his treatment.

Texas Model

Texas has adopted an interesting alternative to involuntary commitment to a secure facility. Faced with the overwhelming expense of committing just 1% of their sex offender population, the Texas Legislature devised an SVP statute (Texas Health and Safety Code-841) that mandates community-based treatment and specialized supervision for released sex offenders who are so designated by the courts. The treatment is funded by the state, and violation of conditions can result in a felony conviction and subsequent return to prison. This system has the potential to avoid the problems of distorted or simply inaccurate risk assessments that might institutionalize a person for life. In a well-run program, the worst outcome of a false positive assessment is that the released sex offender will have

free treatment and closer supervision than he might otherwise receive. The underlying assumption is that the offender can be controlled in the community, and if the individual fails to meet his obligations, he is then dealt with through the criminal justice system. Such an approach also avoids the manipulation of treatment under patently suboptimal conditions for justifying civil commitment and offers extraordinary monetary savings to the state. We offer no evaluative judgment about the efficacy of the Texas program, only about the design of the program. The Texas model is vulnerable to a charge of pretextuality if the conditions for release are difficult or impossible to satisfy.

Best Practice Methodology: Recommendations

Mental Disorder and Mental Abnormality

The mental disorder prong plays a central role in legitimizing SVP commitments. It is, moreover, a key point of contact between law and science. If it lacks legitimacy, as we contend it does, it poses a grave danger to the legitimacy of both the law and the science and thus to the foundation of this statutory management strategy for controlling sexual violence.

The tolerance by the legal system for nonstandard and nonauthoritative diagnoses suggests strongly that the legal system's reliance on diagnostic testimony is largely pretextual. This conclusion is strengthened by the central role that the opaque and confusing concept of VI plays.

We have separated our recommendations by intended audience:

Professional associations. Because the abuses or potential abuses in rendering diagnostic opinions represent the results of activities (almost exclusively) of psychologists and because the problems as identified are applicable in all states with SVP legislation, we strongly recommend that the American Psychological Association appoint a task force charged with the responsibility of developing standards of practice for psychologists in the SVP courtroom.

Courts. Testimony must be based on well-founded, empirically defensible diagnostic principles, and the opined classification must have a scientifically based relationship to reoffense risk in sex offenders, in general, and in the respondent, in particular. In sum, testimony must follow best practices supported by the empirical literature. Courts should demand that examiners (a) identify legitimate *DSM* diagnoses, (b) provide a defense of the case-specific appropriateness of that diagnosis, and (c) provide a defense of the case-specific third prong risk relevance of that diagnosis. Courts should focus on the clear language from *Crane* and *Hendricks*. Only those diagnoses that clearly and unmistakably distinguish the subject of commitment from the typical dangerous recidivists should be accepted. Courts must develop a set of criteria for judging this distinctiveness criterion. Also, the courts must develop a definition of VI that is behaviorally anchored and that truly distinguishes serious and pathological impairment of self-control from the lack of self-control that characterizes most law breakers.

Researchers. We recommend development of a taxonomic system capable of drawing risk-relevant distinctions among sexual offenders, with clear operational criteria for subtype assignment. Both reliability and validity of diagnostic decisions would be improved dramatically, and along with it, there would be

intelligible input to the SVP process. Further, the diagnostic categories should have clear ties to the legal standards.

Examiners. Structured diagnostic interview rating instruments would improve the reliability of Axis I and II psychiatric diagnoses as applied to sex offenders. As a caveat, it is important to emphasize that psychiatric diagnoses (and treatment) change slowly on the basis of new scientific data. Thorough diagnostic assessment would need to incorporate clinical information garnered through contemporary rating instruments based on the current scientific literature. Retrospective chart review, although occasionally rich in details, is not a substitute for direct clinical assessment that targets specific diagnostic issues.

Risk Assessment

The ethicality of risk assessment decision making has been the subject of some discussion (Grisso, 2000; Grisso & Appelbaum, 1992; Litwack, 2002), though, in our estimation, not sufficient discussion, particularly in the realm of indefinite detention, such as SVP proceedings. In the SVP courtroom, the stakes typically are very high, involving liberty interests (for the defendant), safety interests (for victims or potential victims), and competing claims on scarce treatment and prevention resources (for policy makers). Although human error is inevitable, the search for the closest approximation of the truth must be uncompromising. To that end, those who undertake the task of rendering judgments about presumptive dangerousness must exercise utmost care and use methods and procedures that reflect the best practice. As above, we have separated our recommendations by intended audience:

Professional associations. Again, we strongly recommend that the appropriate professional associations, such as the American Psychological Association, appoint a task force charged with the responsibility of developing standards of practice for psychologists and other mental health professionals in the SVP courtroom. Standards of practice pertaining to the evaluation of the SVP respondent's risk to reoffend should address the use of actuarial assessment instruments, adjustments to the actuarial estimate, the use of dynamic risk assessment, the format of the written report, and the language of the eventual testimony. We further recommend a rigorous certification procedure that would establish and test the knowledge base and the opinion-based efficacy of expert witnesses in SVP proceedings.

Courts. Courts should impose a set of requirements that will ensure full transparency with respect to the method and formulation of risk-based opinions. Examiners should be required to disclose and defend the estimated base rate for the respondent, disclose both an estimate of static risk and an estimate of dynamic risk, and discuss the relevance of risk-mitigating and risk-aggravating time-varying factors. Broadly, SVP courts must require examiners to be explicit as to the scientific basis for their opinions (Goodman-Delahunty, 1997).

Legislators. SVP statutes should institute standardized, state-wide screening procedures for selecting candidates for commitment. Screening should be done by an independent entity with no vested interest.

Researchers. Researchers should develop a mechanistic integration of static and dynamic risk predictors that yields a uniformly applied procedure for miti-

gation or aggravation of static risk. They should develop separate life tables for rapists, extrafamilial child molesters, incest-only offenders, and noncontact offenders. And, they should begin examining other stable dynamic risk factors, such as lifestyle impulsivity, which may have a high degree of risk relevance, be easily assessed, and become treatment targets.

Examiners. The best practice method concludes that reliance solely on clinical judgment is improper and, under forensic circumstances, arguably unethical. Grisso and Appelbaum (1993) endorsed the ethicality of risk assessments that are "based on actuarial indicators offering probability estimates of future violence for persons manifesting various measurable characteristics" (p. 483). In their lengthy review of the large number of situations in which actuarial prediction is demonstrably superior to clinical prediction, Grove and Meehl (1996) concluded tersely that, "To use the less efficient of two prediction procedures in dealing with such matters ['high stakes' predictions] is not only unscientific and irrational, it is unethical" (p. 320).

The same best practice method, however, would not, in our opinion, rely exclusively on the results of an ARA and would never knowingly exclude potentially critical, risk-relevant information that is not reflected in the ARA. This last conclusion is made in full awareness of the potential for the contamination of the findings of ARA with clinical adjustments. Risk-relevant augmentation should never be used to change the numeric score of the ARA scale. Such information is used to defend and support one's conclusions that the current risk posed by an individual is greater or lesser than the risk suggested by the ARA scale. Unfortunately, there is no way at the present time to accomplish this task in a uniform, systematic way. We decry, however, the manipulative use of the word *adjustment* to wrap a clinical opinion in the mantle of science. We strongly recommend that the word *adjustment* only be used when there is a clearly and precisely delineated procedure that is systematically applied.

Best practice recognizes the criticality of time varying, individuating, or dynamic risk factors in estimating the current risk of individuals who often committed their known sexual offenses many years, even decades, prior to the evaluation. Indeed, we would argue that best practice, at this point, conceptualizes relapse (sexual reoffense) as a nonlinear, dynamic event (Hufford, Witkiewitz, Shields, Kodya, & Caruso, 2003). Thorough consideration of dynamic risk factors must be incorporated into an overall risk assessment in the narrative conclusions and recommendations provided in the report (e.g., "although the client has a score of 5 on the Static-99, the following dynamic risk factors mitigate the current estimate of risk for this individual").

Incorporating dynamic risk into the conclusions and recommendations is consistent with those of Quinsey, Rice, and Harris (1995), who argued that an actuarial estimate is not merely "an additional piece of information to combine with a clinical appraisal of dangerousness" and that the essential task is "to anchor clinical judgment by having the clinician start with an actuarial estimate of risk and then to alter it by examining dynamic variables. . ." (p. 100). This is not a semantic difference. If the score from an ARA scale is simply incorporated into a clinical judgment, absent any systematic, transparent procedure for doing so that is recommended by the authors of that scale, we run the risk of nullifying the advantage of objectivity achieved by the use of the scale. As Quinsey et al. (1995)

noted, anchoring clinical judgment with mechanistic estimation of risk is not new. Such a procedure was recommended by Gottfredson, Wilkins, and Hoffman (1978) almost 30 years ago for discretionary decisions in parole and sentencing. In sum, best practice demands the recognition and consideration of empirically supported dynamic risk factors in the overall assessment of risk posed by SVP respondents. This admonition, although well recognized in some quarters (e.g., Beech, Fisher, & Thornton, 2003; Douglas & Skeem, 2005; Dvoskin & Heilbrun, 2001), has not generally been heeded by examiners assisting in the adjudication of SVP status.

Treatment

Effective delivery of treatment in SVP programs is critical for two fundamental reasons. First, despite the rather tolerant approach taken by many courts, courts may lose patience with treatment programs that lack integrity. Second, effective treatment programs are the key to maintaining some reasonable control over the growing populations, and concomitant growing expense, of SVP programs. To that end, we offer specific recommendations based on well-established professional standards, to advance the efficacy of treatment that is provided under the umbrella of SVP commitment. As above, we have separated our recommendations by intended audience:

Legislators and corrective services practitioners. Treatment must be provided in the least restrictive therapeutic environment. Programs must establish objective, measurable, obtainable treatment goals. Absent clear goals, it is impossible to draw reliable conclusions about the risk-relevant impact of treatment. Comprehensive intake assessment must identify individualized treatment goals. Treatment programs must respect confidentiality within the security limits of the program. To avoid obvious conflict, clinical staff must be removed from all tasks related to commitment or release. Programs must offer meaningful educational and vocational components. Programs must include reintegration programs that transition men back to the community and viable aftercare programs in the community. Ongoing feedback mechanisms for evaluating the successes and failures of the program are critical. Finally, empiricism must be integrated with clinical practice at each step.

Examiners. There are two overarching principles for consideration of treatment in court. First, treatment must be regarded as simply one of a number of important dynamic risk factors that must be considered and not as the ultimate weapon to keep detainees in prison. Second, change, or lack thereof, as a function of treatment must be risk-relevant. There are many clinical issues focused on in treatment, such as denial, minimization, lack of empathy, low motivation, and length of treatment, that do not appear to be risk relevant (Hanson & Bussiere, 1998; Hanson & Morton-Bourgon, 2005).

Conclusion

We are witnessing, in dramatic relief, the unfolding of two outcomes in SVP hearings: (a) the loss of the integrity of the adjudicative process with testimony that is, at best, opaque, and, at worst, dissimulative; and (b) the loss of the integrity of science itself. Like the king who appeared before his startled minions

with transparent lack of attire, the fact finder is frequently confronted by expert testimony that is energetic in spirit but naked in substance. Like the king, the experts sport a well-tailored suit of opinions, which often are diaphanous. The net result is an expensive process that has a veneer of legal and scientific polish but ignores both the norms of a just society and the solid guidance that good science could furnish. Both are critical ingredients for any successful battle against the scourge of sexual violence.

The mental disorder prong plays a central role in legitimizing SVP commitments. It is, moreover, a key point of contact between law and science. If it lacks legitimacy, as we contend it does, it poses a grave danger to the legitimacy of both the law and the science and thus to the foundation of this statutory management strategy for controlling sexual violence. The tolerance by the legal system for nonstandard and nonauthoritative diagnoses suggests strongly that the legal system's reliance on diagnostic testimony is largely pretextual. This conclusion is strengthened by the central role that the opaque and confusing concept of VI plays.

Our fundamental recommendation is clear and crisp: SVP courts must insist on good science. This is more than the generic plea that could be sounded across the litigation landscape. More is at stake here than the allocation of rights among private parties. Vast sums are being spent on SVP proceedings and programs. If these resources are guided by pretextualism, by a pseudoscience shaved and extruded to fit a politically expedient approach to public safety, we can expect less than optimal return on our investment. What is worse, however, is a distinct risk that the application of science itself will be distorted, that the opaque and transient needs of SVP cases will warp its findings, and that the fact finder will be faced with highly unreliable testimony cloaked in the mantle of science.

We fully recognize that implementing this broad recommendation will require the cooperation of all participants in the SVP process. Courts, attorneys, and examiners must understand that the interface between the law and science requires careful translation, enforcement of professional boundaries, and competent evaluation of scientific testimony. The adversarial nature of the process must not be allowed to cross the bright line of good, or at the very least acceptable, science.

Poor science may survive judicial scrutiny because good science will embarrass the only politically viable stance, maintaining the appearance of being tough on sex offenders. More likely, judges are simply ill equipped to discriminate good from bad science, particularly in the statistical arena. Judges may also occasionally excuse poor science, because they truly believe that mistakes in diagnosis, risk assessment, and treatment do not adversely affect the end goal, protecting the public from sexual violence. In truth, the opposite is the case. Bad science hinders, and good science advances, our common interest in protecting society from sexual offenders.

References

Abel, G. G., Becker, J. V., Cunningham-Rathner, J., Mittleman, M., & Rouleau, J. L. (1988). Multiple paraphilic diagnoses among sex offenders. *Bulletin of the American Academy of Psychiatry and the Law, 16,* 153–168.

Abracen, J., & Looman, J. (2006). Evaluation of civil commitment criteria in a high risk

sample of sexual offenders. *Journal of Sexual Offender Civil Commitment: Science and the Law, 1,* 124–140.

Ahlmeyer, S., Kleinsasser, D., Stoner, J., & Retzlaff, P. (2003). Psychopathology of incarcerated sex offenders. *Journal of Personality Disorder, 17,* 306–318.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57,* 1060–1063.

Arizona Revised Statutes Annotated § 36–3701 (1998).

Bailar, J. C., & Bailer, A. J. (1999). Risk assessment—The mother of all uncertainties: Disciplinary perspectives on uncertainty in risk assessment. *Annals of the New York Academy of Sciences, 895,* 273–285.

Baker, H. W. G., & Hudson, B. (1983). Changes in the pituitary–testicular axis with age. *Monographs on Endocrinology, 25,* 71–83.

Barbaree, H. E. (1997). Evaluating treatment efficacy with sexual offenders: The insensitivity of recidivism studies to treatment effects. *Sexual Abuse: A Journal of Research and Treatment, 9,* 111–128.

Barbaree, H. E., Blanchard, R., & Langton, C. M. (2003). The development of sexual aggression through the life span: The effect of age on sexual arousal and recidivism among sex offenders. *Annals of the New York Academy of Sciences, 989,* 59–71.

Becker, J. V., Stinson, J., Tromp, S., & Messer, G. (2003). Characteristics of individuals petitioned for civil commitment. *International Journal of Offender Therapy and Comparative Criminology, 47,* 185–195.

Beech, A. R., Fisher, D. D., & Thornton, D. (2003). Risk assessment of sex offenders. *Professional Psychology: Research and Practice, 34,* 339–352.

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

Behnke, S. (2005). Diagnoses, record reviews and the new Ethics Code: Ethical Standard 9.01 guides psychologists in rendering diagnostic opinions. *The Monitor, 36,* 80–81.

Blanchard, R., & Barbaree, H. E. (2005). The strength of sexual arousal as a function of the age of the sex offender: Comparisons among pedophiles, hebephiles, and teleiophiles. *Sexual Abuse: A Journal of Research and Treatment, 17,* 441–456.

Blashfield, R. K. (1973). Evaluation of the *DSM–II* classification of schizophrenia as a nomenclature. *Journal of Abnormal Psychology, 82,* 382–389.

Blumstein, A., Cohen, J., Roth, J. A., & Visher, C. A. (1986). *Criminal careers and "career criminals"* (Vol. 1). Washington, DC: National Academy Press.

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.* Burnaby, British Columbia, Canada: Simon Fraser University, Mental Health, Law, and Policy Institute.

Brief for the Association for the Treatment of Sexual Abusers As Amicus Curiae in Support of Petitioner, Kansas v. Crane, No. 00–957, 2001 WL 670067 (U.S.). Available from http://web2.westlaw.com

Cocozza, J., & Steadman, H. (1976). Prediction in psychiatry: An example of misplaced confidence in experts. *Rutgers Law Review, 29,* 1084–1101.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, L. J. (1981a). Are there any a priori constraints on the study of rationality? *Behavioral and Brain Sciences, 4,* 359–367.

Cohen, L. J. (1981b). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences, 4,* 317–331.

Cole, N. (1996). Debunking junk. *Nucleus, 3*(2), 3–4.

Cornwell, J. K. (1998). Understanding the role of the police and *parens patriae* powers in involuntary civil commitment before and after *Hendricks. Psychology, Public Policy, and Law, 4,* 377–413.

Davis, D., & Follette, W. C. (2002). Rethinking the probative value of evidence: Base rates, intuitive profiling, and the "*post*diction" of behavior. *Law and Human Behavior, 26,* 133–158.

Dawes, R. M., Faust, D., & Meehl, P. E. (1992). Statistical prediction versus clinical prediction: Improving what works. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 351–367). Mahwah, NJ: Erlbaum.

Denti, L., Pasolini, G., Sanfelici, L., Benedetti, R., Cecchetti, A., Ceda, G. P., et al. (2000). Aging-related decline of gonadal function in healthy men: Correlation with body composition and lipoproteins. *Journal of the American Geriatrics Society, 48,* 51–58.

Douglas, K. S., & Skeem, J. L. (2005). Violence risk assessment: Getting specific about being dynamic. *Psychology, Public Policy, and Law, 11,* 347–383.

Dunsieth, N. W., Nelson, E. B., Brusman-Lovins, L. A., Holcomb, J. L., Beckman, D., Welge, J. A., et al. (2004). Psychiatric and legal features of 113 men convicted of sexual offenses. *Journal of Clinical Psychiatry, 65,* 293–300.

Dvoskin, J. A., & Heilbrun, K. (2001). Risk assessment and release decision-making: Toward resolving the great debate. *Journal of the American Academy of Psychiatry and the Law, 29,* 6–10.

Ennis, B. J., & Litwack, T. R. (1974). Psychiatry and the presumption of expertise: Flipping coins in the courtroom. *California Law Review, 62,* 693–752.

Epperson, D. L., Ralston, C. A., Fowers, D., DeWitt, J., & Gore, K. S. (2006). Actuarial risk assessment with juveniles who offend sexually: Development of the Juvenile Sexual Offense Recidivism Risk Assessment Tool–II (J-SORRAT-II). In D. Prescott (Ed.), *Risk assessment: Theory, controversy, and emerging strategies of youth who have sexually abused* (pp. 118–169). Oklahoma City, OK: Wood and Barnes Publishing.

Ewing, C. (1983). "Dr. Death" and the case for an ethical ban on psychiatric and psychological prediction of dangerousness in capital sentencing proceedings. *American Journal of Law and Medicine, 8,* 407–428.

Ewing, C. (1985). *Schall v. Martin*: Preventive detention and dangerousness through the looking glass. *Buffalo Law Review, 34,* 173–226.

Ewing, C. (1991). Preventive detention and execution: The constitutionality of punishing future crimes. *Law and Human Behavior, 15,* 139–163.

Fazel, S., Sjöstedt, G., Långström, N., & Grann, M. (in press). Risk factors for criminal recidivism in older sexual offenders. *Sexual Abuse: A Journal of Research and Treatment.*

Foucha v. Louisiana, 504 U.S. 7 (1992).

Garb, H. N. (1998). *Studying the clinician.* Washington, DC: American Psychological Association.

Goodman-Delahunty, J. (1997). Forensic psychological expertise in the wake of *Daubert. Law and Human Behavior, 21,* 121–140.

Gottfredson, D. M., Wilkins, L. T., & Hoffman, P. B. (1978). *Guidelines for parole and sentencing: A policy control method.* Toronto, Ontario, Canada: Lexington Books.

Grant, J. E. (2005). Clinical characteristics and psychiatric comorbidity in males with exhibitionism. *Journal of Clinical Psychiatry, 66,* 1367–1371.

Grisso, T. (1987). The economic and scientific future of forensic psychological assessment. *American Psychologist, 42,* 831–839.

Grisso, T. (2000, March). *Ethical issues in evaluations for sex offender re-offending.* Paper presented at the Symposium on Sex Offender Re-Offense Risk Prediction, Sinclair Seminars, Madison, WI.

Grisso, T., & Appelbaum, P. S. (1992). Is it unethical to offer predictions of future violence? *Law and Human Behavior, 16,* 621–633.

Grisso, T., & Appelbaum, P. S. (1993). Structuring the debate about ethical predictions of future violence. *Law and Human Behavior, 17,* 482–485.

Grove, W. M., & Meehl, P. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical-statistical controversy. *Psychology, Public Policy, and Law, 2,* 293–323.

Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment, 12,* 19–30.

Hanson, R. K. (1998). What do we know about sex offender risk assessment? *Psychology, Public Policy, and Law, 4,* 50–72.

Hanson, R. K. (2002). Recidivism and age: Follow-up data on 4,673 sexual offenders. *Journal of Interpersonal Violence, 17,* 1046–1062.

Hanson, R. K. (2005). *The validity of Static-99 with older sexual offenders.* Retrieved August 8, 2005, from http://www.sppc-psepc.gc.ca/publications/Corrections/20050630

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. (2001). A structured approach to evaluating change among sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 13,* 105–122.

Hanson, R. K., & Morton-Bourgon, K. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology, 73,* 1154–1163.

Hanson, R. K., & Thornton, D. (1999). *Static-99: Improving actuarial risk assessments for sex offenders* [User Report No. 99–02]. Ottawa, Ontario, Canada: Department of the Solicitor General of Canada.

Hare, R. D. (2003). *Hare Psychopathy Checklist—Revised manual* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.

Hiday, V. A. (1990). Dangerousness of civil commitment candidates: A six-month follow-up. *Law and Human Behavior, 14,* 551–567.

Hildebrand, M., de Ruiter, C., & de Vogel, V. (2004). Psychopathy and sexual deviance in treated rapists: Association with sexual and nonsexual recidivism. *Sexual Abuse: A Journal of Research and Treatment, 16,* 1–24.

Hirschi, T., & Gottfredson, M. (1983). Age and the explanation of crime. *American Journal of Sociology, 89,* 552–584.

Hudson, S. K., Wales, D. S., Bakker, L., & Ward, T. (2002). Dynamic risk factors: The Kia Marama evaluation. *Sexual Abuse: A Journal of Research and Treatment,* 14, 103–120.

Hufford, M. R., Witkiewitz, K., Shields, A. L., Kodya, S., & Caruso, J. C. (2003). Relapse as a nonlinear dynamic system: Application to patients with alcohol use disorders. *Journal of Abnormal Psychology, 112,* 219–227.

In re Crane, 7 P.3d 285 (Kan. 2000).

Jackson, R. L., Rogers, R., & Shuman, D. W. (2004). The adequacy and accuracy of

sexually violent predator evaluations: Contextualized risk assessment in clinical practice. *International Journal of Forensic Mental Health, 3,* 115–129.

Janus, E. S. (1997). Toward a conceptual framework for assessing police power commitment legislation: A critique of Schopp's and Winnick's explications of legal mental illness. *Nebraska Law Review, 76,* 1–50.

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. (2001). Sex offender commitments and the "inability to control": Developing legal standards and a behavioral vocabulary for an elusive concept. In A. Schlank (Ed.), *The sexual predator: Legal issues, clinical issues, special situations.* Kingston, NJ: Civic Research Institute.

Janus, E. S. (2003). Treatment and the civil commitment of sex offenders. In B. J. Winick & J. Q. Lafond (Eds.), *Protecting society from dangerous offenders: Law, justice, and therapy* (pp. 119–130). Washington, DC: American Psychological Association.

Janus, E. S. (2004). Sexually violent predator laws: Psychology in service to a morally dubious enterprise. *The Lancet, 364,* 50–51.

Janus, E. S. (in press). *Failure to protect: America's sexual predator laws and the rise of the preventive state.* Ithaca, NY: Cornell University Press.

Janus, E. S., & Prentky, R. A. (2003). Forensic use of actuarial risk assessment with sex offenders: Accuracy, admissibility and accountability. *American Criminal Law and Review, 40,* 1443–1499.

Kaemingk, K. L., Koselka, M., Becker, J. V., & Kaplan, M. S. (1995). Age and adolescent sexual offender arousal. *Sexual Abuse: A Journal of Research and Treatment, 7,* 249–257.

Kafka, M. P. (2003). Sex offending and sexual desire: The clinical and theoretical relevance of hypersexual desire. *International Journal of Offender Therapy and Comparative Criminology, 47,* 439–451.

Kafka, M. P., & Hennen, J. (2002). A *DSM–IV* Axis I comorbidity study of males (n = 120) with paraphilias and paraphilia-related disorders. *Sexual Abuse: A Journal of Research and Treatment, 14,* 349–366.

Kafka, M. P., & Hennen, J. (2003). Hypersexual desire in males: Are males with paraphilias different from males with paraphilia-related disorders? *Sexual Abuse: A Journal of Research and Treatment, 15,* 307–323.

Kafka, M. P., & Prentky, R. A. (1998). Attention-deficit/hyperactivity disorder in males with paraphilias and paraphilia-related disorders: A comorbidity study. *Journal of Clinical Psychiatry, 59,* 388–396.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58,* 697–720.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

Kansas v. Crane, 534 U.S. 407 (2002).

Kansas v. Hendricks, 521 U.S. 346 (1997).

Kavoussi, R. J., Kaplan, M., & Becker, J. V. (1988). Psychiatric diagnosis in adolescent sex offenders. *Journal of American Academy of Child and Adolescent Psychiatry, 27,* 241–243.

Knight, R. A., & Sims-Knight, J. E. (2003). The developmental antecedents of sexual coercion against women: Testing alternative hypotheses with structural equation modeling. *Annals of the New York Academy of Sciences, 989,* 72–85.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Science, 19,* 1–53.

La Fond, J. Q. (2000). The future of involuntary civil commitment in the USA after *Kansas v. Hendricks. Behavioral Sciences and the Law, 18,* 153–167.

Långström, N., Sjöstedt, G., & Grann, M. (2004). Psychiatric disorders and recidivism in sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 16,* 139–150.

Levenson, J. S. (2004). Sexual predator civil commitment: A comparison of selected and released offenders. *International Journal of Offender Therapy and Comparative Criminology, 48,* 638–648.

Lieb, R., & Gookin, K. (2005, March). *Involuntary commitment of sexually violent predators: Comparing state laws.* Retrieved May 12, 2006, from www.wsipp.wa.gov/rptfiles/05–03–1101.pdf

Litwack, T. R. (2002). Some questions for the field of violence risk assessment and forensic mental health: Or, "back to basics" revisited. *International Journal of Forensic Mental Health, 1,* 171–178.

Malamuth, N. M. (2003). Criminal and noncriminal sexual aggressors: Integrating psychopathy in a hierarchical–mediational confluence model. *Annals of the New York Academy of Sciences, 989,* 33–58.

McElroy, S. L., Soutullo, C. A., Taylor, P., Nelson, E. B., Beckman, D. A., Brusman, A. L., et al. (1999). Psychiatric features of 35 men convicted of sexual offenses. *Journal of Clinical Psychiatry, 60,* 414–420.

Meehl, P. E. (1954). *Clinical vs. statistical prediction.* Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1996). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50,* 266–275.

Megargee, E. I. (1981). Methodological problems in the prediction of violence. In J. Hays, T. Roberts, & K. Solway (Eds.), *Violence and the violent individual* (pp. 179–191). New York: Spectrum.

Mercado, C. C., Schopp, R. F., & Bornstein, B. H. (2005). Evaluating sex offenders under sexually violent predator laws: How might mental health professionals conceptualize the notion of volitional impairment? *Aggression and Violent Behavior, 10,* 289–309.

Miller, H., Amenta, A. E., & Conroy, M. A. (2005). Sexually violent predator evaluations: Empirical evidence, strategies for professionals, and research directions. *Law and Human Behavior, 29,* 29–54.

Monahan, J. (1981). *Predicting violent behavior: An assessment of clinical techniques.* Newbury Park, CA: Sage Publications.

Monahan, J., & Steadman, H. (Eds.). (1994). *Violence and mental disorder: Developments in risk assessment.* Chicago: University of Chicago Press.

Monahan, J., Steadman, H., Silver, E., Applebaum, P. S., Clark-Robbins, P., Mulvey, E. P., et al. (2001). *Rethinking risk assessment: The MacArthur study of mental disorder and violence.* New York: Oxford Press.

Money, J. (1986). *Lovemaps.* New York: Irvington.

Money, J. (1995). Forensic sexology: Paraphilic serial rape (biastophilia) and lust murder (erotophonophilia). *Acta Sexologica, 1,* 47–62.

Morse, S. J. (1994). Culpability and control. *University of Pennsylvania Law Review, 142,* 1587–1660.

O'Connor, T. G., McGuire, S., Reiss, D., Hetherington, E. M., & Plomin, R. (1998). Co-occurrence of depressive symptoms and antisocial behavior in adolescence: A common genetic liability. *Journal of Abnormal Psychology, 107,* 27–37.

O'Donohue, W., Regev, L. G., & Hagstrom, A. (2000). Problems with the *DSM–IV* diagnosis of pedophilia. *Sexual Abuse: A Journal of Research and Treatment, 12,* 95–105.

Office of Juvenile Justice and Delinquency Prevention (2004, August). *Statistical briefing book.* Retrieved June 5, 2005, from http://ojjdp.ncjrs.org/ojstatbb/crime/qa05301.asp?qaDate=20040801

Parsons v. Alabama, 81 Ala. 577, 2 So. 854 (1886).

Pearson v. Probate Court of Ramsey County, 205 Minn. 545, 287 N.W. 297 (1939).

Perlin, M. L. (1991). Morality and pretexuality, psychiatry and law: Of "ordinary common sense," heuristic reasoning, and cognitive dissonance. *Bulletin of the American Academy of Psychiatry and the Law, 19,* 131–150.

Perlin, M. L. (1993). Pretexts and mental disability law: The case of competency. *University of Miami Law Review, 47,* 625–688.

Polaschek, D. L. L. (2003). The classification of sexual offenders. In T. Ward, D. R. Laws, & S. M. Hudson (Eds.), *Sexual deviance: Issues and controversies* (pp. 154–171). Thousand Oaks, CA: Sage Publications.

Prentky, R. A., Harris, B., Frizzell, L., & Righthand, S. (2000). An actuarial procedure for assessing risk in juvenile sex offenders. *Sexual Abuse: A Journal of Research and Treatment, 12,* 71–93.

Prentky, R. A., Knight, R. A., Lee, A. F. S., & Cerce, D. (1995). Predictive validity of lifestyle impulsivity for rapists. *Criminal Justice and Behavior, 22,* 106–128.

Prentky, R. A., Lee, A. F. S., Knight, R. A., & Cerce, D. (1997). Recidivism rates among child molesters and rapists: A methodological analysis. *Law and Human Behavior, 21,* 635–659.

Prentky, R. A., & Righthand, S. (2003). *Juvenile Sex Offender Assessment Protocol–II (J-SOAP-II) manual.* Justice Resource Institute, Bridgewater, MA.

Quinsey, V. L. (1996). Improving decision accuracy where base rates matter: The prediction of violent recidivism. *Behavioral and Brain Sciences, 19,* 37–38.

Quinsey, V. L., Chaplin, T. C., & Varney, G. (1981). A comparison of rapists' and non-sex offenders' sexual preferences for mutually consenting sex, rape, and sadistic acts. *Behavioral Assessment, 3,* 127–135.

Quinsey, V. L., Rice, M. E., & Harris, G. T. (1995). Actuarial prediction of sexual recidivism. *Journal of Interpersonal Violence, 10,* 85–105.

Raymond, N. C., Coleman, E., Ohlerking, E., Christenson, G. A., & Miner, M. (1999). Psychiatric comorbidity in pedophilic sex offenders. *American Journal of Psychiatry, 156,* 786–788.

Righthand, S., Prentky, R. A., Knight, R. A., Carpenter, E., Hecker, J. E., & Nangle, D. (2005). Factor structure and validation of the Juvenile Sex Offender Assessment Protocol (J-SOAP). *Sexual Abuse: A Journal of Research and Treatment, 17,* 13–30.

Rogers, R. (2000). The uncritical acceptance of risk assessment in forensic practice. *Law and Human Behavior, 24,* 595–605.

Rowland, D. L., Greenleaf, W. J., Dorfman, L. J., & Davidson, J. M. (1993). Aging and sexual function in men. *Archives of Sexual Behavior, 22,* 545–557.

Sampson, R. J., & Laub, J. (2003). Life-course desisters? Trajectories of crime among delinquent boys followed to age 70. *Criminology, 41,* 301–339.

Schopp, R. F., Scalora, M. J., & Pearce, M. (1999). Expert testimony and professional judgment: Psychological expertise and commitment as a sexual predator after *Hendricks*. *Psychology, Public Policy, and Law, 5,* 120–174.

Serin, R. C., Malcolm, P. B., Khanna, A., & Barbaree, H. E. (1994). Psychopathy and deviant sexual arousal in incarcerated sexual offenders. *Journal of Interpersonal Violence, 9,* 3–11.

Swets, J. A. (1992). The science of choosing the right decision threshold in high-stakes diagnostics. *American Psychologist, 47,* 522–532.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest, 1,* 1–26.

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

Thornton, D. (in press). Age and sexual recidivism: A variable connection. *Sexual Abuse: A Journal of Research and Treatment.*

Wolfgang, M. E., & Ferracuti, F. (1982). *The subculture of violence: Towards an integrated theory of criminology.* Beverly Hills, CA: Sage Publications.

Wollert, R. (2006). Low base rates limit expert certainty when current actuarials are used to identify sexually violent predators: An application of Bayes' theorem. *Psychology, Public Policy, and Law, 12,* 56–85.

Worling, J. R. (2004). The Estimate of Risk of Adolescent Sexual Offense Recidivism (ERASOR): Preliminary psychometric data. *Sexual Abuse: A Journal of Research and Treatment, 16,* 235–254.

Worling, J. R., & Curwen, T. (2001). *The "ERASOR": Estimate of Risk of Adolescent Sexual Offense Recidivism* (Version 2.0). Toronto, Ontario, Canada: Safe-T Program, Thistletown Regional Centre.

Zander, T. K. (2005). Civil commitment without psychosis: The law's reliance on the weakest link in psychodiagnosis. *Journal of Sexual Offender Civil Commitment: Science and the Law, 1,* 17–82.

Received December 23, 2005
Revision received July 9, 2006
Accepted July 9, 2006 ■