
**PUBLIC SAFETY, INDIVIDUAL LIBERTY, AND SUSPECT
SCIENCE: FUTURE DANGEROUSNESS ASSESSMENTS
AND SEX OFFENDER LAWS**

*Melissa Hamilton**

I.	INTRODUCTION	698
II.	SPECIAL TREATMENT FOR SEX OFFENDERS	700
	A. <i>SVP Laws</i>	703
	1. Civil Commitment Laws.....	703
	2. Registration Requirements.....	705
	3. Residency Restrictions.....	705
	B. <i>(Un)Intended Consequences of Sexual Predator Laws</i>	707
	1. Sexual Predators and Stranger Danger: Is the Concern Valid?.....	707
	2. Do Sexual Offender Laws Work?.....	709
	3. Rethinking the Problem	712
III.	EXPERT EVIDENCE AND SEX OFFENDERS	713
	A. <i>Expert Evidence Law</i>	713
	B. <i>Experts and Future Dangerousness Assessment</i>	717
IV.	ACTUARIAL TESTING OF FUTURE DANGEROUSNESS.....	720
	A. <i>The Sexual Recidivism Actuarial Tests</i>	721
	B. <i>Empirical Evaluation of Actuarial Evidence of Future Dangerousness</i>	725
	1. Testing and the Scientific Method	726
	2. Validity	727
	3. Reliability and Objectivity	731
	4. Training	733
	5. General Acceptance	733
V.	JUDICIAL PERSPECTIVES ON FUTURE DANGEROUSNESS EVIDENCE.....	735
	A. <i>Daubert/Frye Challenges</i>	735
	1. <i>Frye Challenges</i>	735
	2. <i>Daubert and Alternatives</i>	737
	B. <i>The Standard of Likelihood to Sexually Reoffend</i>	740
	C. <i>Battle of the Experts</i>	744
	D. <i>(Mis)Interpretation of Actuarial Prediction</i>	749
VI.	CONCLUSION	754

I. INTRODUCTION

Arguably one of the most prolonged and contentious debates in legal history centers on the balance between protecting basic civil liberties as guaranteed by the Constitution and protecting the public from harm. Certainly the shadow of terrorist threats after the 9/11 attacks reignited this debate and has fueled public interest in changing the balance between security and liberty in favor of introducing a more preventive state.¹ While tipping the scale in favor of using a preventive law model might initially serve to calm public fear, be it about terrorism or crime generally, such actions may ultimately reduce the freedoms of many citizens in the long-term.² This is particularly true in the highly emotional area of predatory sex crimes.³

In recent decades, federal, state, and local governments have become increasingly restrictive on the freedom and privacy of those labeled sexually violent predators (SVPs) for the purpose of preventing further sexual violence. The most commonly used tools to manage SVPs are involuntary commitments for mental treatment, sex offender registration, and residency restrictions (hereinafter “SVP laws”).⁴ In an effort to streamline the identification of sex offenders who pose a future danger and thereby might be subject to SVP laws, officials place substantive legal emphasis on psychosexual evaluations by individuals accepted as experts.⁵ These experts are generally mental health practitioners who offer opinion evidence about an individual’s likelihood of sexually reoffending, often using actuarial (statistical calculation of risk) assessments. While protecting the public from the damage that sexual violence causes is clearly an important goal, this Article critically analyzes whether future dangerousness assessments using actuarial tools are responsive to legal standards contained in SVP laws and whether courts, when confronted with such assessments, are adequately engaging in the gatekeeper role to accept only “good science”⁶ as per the

* Visiting Assistant Professor of Law, University of South Carolina School of Law. J.D., The University of Texas at Austin School of Law; Ph.D, Sociology, The University of Texas at Austin.

1. The term “preventive state,” coined by Professor Carol Steiker, refers to a state that “attempt[s] to identify and neutralize dangerous individuals before they commit crimes by restricting their liberty in a variety of ways.” Carol Steiker, *The Limits of the Preventive State*, 88 J. CRIM. L. & CRIMINOLOGY 771, 774 (1998). For a comparison of sex offender laws before and after September 11, 2011, see Danye Holley, *The Supreme Courts: Did September 11th Accelerate Their Sanctioning the Constitutionality of Criminalizing Suspicion?*, 7 PIERCE L. REV. 39, 60–63 (2008).

2. See Eric S. Janus & Robert A. Prentky, *Sexual Predator Laws: A Two-Decade Retrospective*, 21 FED. SENT’G REP. 90, 95 (2008) (noting that “[p]rofil[ing] designed to identify ‘dangerous people’ . . . destroy[ed] the lives of many innocent Americans” during Cold War and suggesting similar risks may result from resuming such practices against sex offenders and suspected terrorists).

3. See ERIC S. JANUS, *FAILURE TO PROTECT: AMERICA’S SEXUAL PREDATOR LAWS AND THE RISE OF THE PREVENTIVE STATE* 5 (2006) (arguing that “preventive detention” approach of sexual predator laws constitute “radical assault” on fundamental principles of U.S. criminal justice system).

4. Janus & Prentky, *supra* note 2, at 90.

5. Daniel F. Montaldi, *The Logic of Sexually Violent Predator Status in the United States of America*, 2 SEXUAL OFFENDER TREATMENT 1 (2007), <http://www.sexual-offender-treatment.org/index.php?id=57&type=123>.

6. “Good science” has been defined as “the faithful and rigorous adherence to the findings, the limitations, and the conclusions of published, peer-reviewed articles in scientific journals.” Robert A. Prentky et al., *Sexually Violent Predators in the Courtroom: Science on Trial*, 12 PSYCHOL. PUB. POL’Y & L. 357, 358

evidentiary benchmarks of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*⁷ and *Frye v. United States*.⁸

In Part II, the dominant SVP laws are outlined and a discussion follows about how and why politicians and the public are enamored with the assumed need for these specialized laws. Statistical evidence is then provided that challenges whether public fears about SVPs are realistic and whether SVP laws serve their preventive goals.

Part III provides a brief review of the current law on the admissibility criteria for expert evidence, with the concomitant summary of *Daubert*- and *Frye*-based standards. Additionally, the role of mental health practitioners in providing expert evidence in court when future dangerousness is at issue is discussed.

Part IV summarizes two of the most common actuarial risk assessment tools used to identify and label sex offenders who will be subject to SVP laws. An explanatory analysis is then provided of the empirical validity of actuarial tools in predicting sexual offense recidivism. Assessing whether actuarial-based risk prediction is good science is of extreme importance today. Not only is the implementation of SVP laws growing costlier and more intrusive in scope, new evidence has emerged to significantly undermine the validity of applying these actuarial tests to U.S. offenders.

Part V offers a review of case law involving the role of the two actuarial assessment tools in SVP status cases, including an assessment of how courts have responded to *Daubert*- and *Frye*-based challenges to the instruments. Considering that the vast majority of courts decline such challenges, this Part explores how the actuarial assessments instrumentally impact legal decisions on future dangerousness. In addition, Part V analyzes evidence indicating that instances of experts engaging in adversarial bias and presenting empirically incorrect representations of actuarial predictions are all too common.

Finally, the Conclusion offers cautionary notes and suggestions for reevaluating SVP laws and actuarial assessments of future dangerousness. Specifically, this Article concludes that because of uncritical reliance upon actuarial assessments of future dangerousness, legal professionals have largely failed to grasp the significant empirical and practical limitations of these tests.

Judges and lawyers participating in SVP cases must appreciate the serious challenges to the reliability and validity of actuarial assessments being presented as expert evidence. The potential that criminal justice practitioners use empirically questionable assessments to inform decisions on the SVP status of individual defendants poses significant negative impacts upon the public and upon defendants.⁹ SVP laws are very expensive and resource intensive for governmental institutions to implement, maintain, and enforce.¹⁰ Defendants also suffer a considerable infringement

(2006). By contrast, bad science has been defined as the “intentional or uninformed distortion, misinterpretation, or selective reporting of findings from scientific articles.” *Id.*

7. 509 U.S. 579 (1993).

8. 293 F. 1013 (D.C. Cir. 1923).

9. Technically, because the sex offender laws at issue in this Article are not criminal in nature, the individuals are not criminal *defendants*, but civil respondents in most cases. The term defendant is used herein to more easily identify the individual to which an SVP law is applied.

10. See *infra* notes 86–88 and accompanying text for a discussion of the costs of SVP laws.

of their constitutionally-protected interests in liberty and privacy through measures (i.e., incarceration, public registration, and residency restrictions) which can render them labeled as sexually violent predators. Hence, if pseudoscience greatly impacts these decisions, the significant risk of false positives (giving a sexual predator label to a defendant not likely to reoffend) and false negatives (not giving a sexual predator status to a defendant highly likely to reoffend) undermines the preventive goals of SVP laws while rendering significant costs to the public and to individual defendants.

II. SPECIAL TREATMENT FOR SEX OFFENDERS

If sex sells, and if violence sells, their combination is exponentially alluring,¹¹ and provides sensational headlines that suggest the United States is in the midst of a sex crimes wave.¹² Largely due to the media fomenting moral panic about dangerous sex offenders,¹³ the management of sex offenders is a top priority for legislative action today.¹⁴ The connection between the media and legislative efforts to invoke preventive measures on SVPs is clear, as the pressure to increase control and punishment of SVPs often comes from citizens after heinous sex acts become high profile cases in the news.¹⁵ However, the media's method of coverage has also magnified exceptional

11. See Joseph L. Lester, *Off to Elba! The Legitimacy of Sex Offender Residence and Employment Restrictions*, 40 AKRON L. REV. 339, 345 (2007) (“[S]ex and violence makes for a story line that sells.”).

12. This is not the first time that a crackdown on sex offenders through legislation has occurred. In 1947, J. Edgar Hoover, then director of the FBI, publicly declared that the “rapidly increasing” rate of sex crimes needed to “be placed under the spotlight and its evils disclosed so that something may be done to correct a situation that leaves maimed and murdered women lying in isolated areas, which leaves violated children in a state of hysteria, and which is a perpetual nightmare to the loved ones and friends of the victims.” Roxanne Lieb et al., *Sexual Predators and Social Policy*, 23 CRIME & JUST. 43, 53 (1998) (quoting J. Edgar Hoover, *How Safe Is Your Daughter?* 144 AM. MAG. 32, 32 (1947)). For a history of the origin of twentieth century laws to detain sexual psychopaths, see generally Tamara Rice Lave, *Only Yesterday: The Rise and Fall of Twentieth Century Sexual Psychopath Laws*, 69 LA. L. REV. 549 (2009). The prior generation of sexual psychopath laws differed from current civil commitment statutes in that the former required rehabilitative treatment in a mental health facility in lieu of punitive incarceration. Dawn J. Post, Comment, *Preventive Victimization: Assessing Future Dangerousness in Sexual Predators for Purposes of Indeterminate Civil Commitment*, 21 HAMLINE J. PUB. L. & POL’Y 177, 185 (1999). Most of the former laws were repealed in the 1970s when criminal justice policy initiatives shifted away from rehabilitation models. *Id.*

13. See Eric S. Janus & Emily A. Polachek, *A Crooked Picture: Re-framing the Problem of Child Sexual Abuse*, 36 WM. MITCHELL L. REV. 142, 147–48 (2009) (discussing connection between media coverage and public belief that there is “epidemic of sexual violence against children”); Meaghan Kelly, *Lock them Up—And Throw Away the Key: The Preventive Detention of Sex Offenders in the United States and Germany*, 39 GEO. J. INT’L L. 551, 553 (2008) (discussing “moral panic” surrounding sex offenders (quoting ANNE-MARIE MCALINDEN, *THE SHAMING OF SEX OFFENDERS: RISK, RETRIBUTION AND REINTEGRATION* 18 (2007))); Wayne A. Logan, *Criminal Justice Federalism and National Sex Offender Policy*, 6 OHIO ST. J. CRIM. L. 51, 60 (2008) (discussing how public outrage over sex offenders has prompted congressional action); Bela August Walker, Essay, *Deciphering Risk: Sex Offender Statutes and Moral Panic in a Risk Society*, 40 U. BALT. L. REV. 183, 184 (2010) (describing moral panic surrounding sex crimes as “perpetual moral anxiety”).

14. CTR. FOR SEX OFFENDER MGMT., *LEGISLATIVE TRENDS IN SEX OFFENDER MANAGEMENT 1* (2008), available at http://www.csom.org/pubs/legislative_trends.pdf. See generally Corey Rayburn Yung, *The Emerging Criminal War on Sex Offenders*, 45 HARV. C.R.-C.L. L. REV. 435 (2010) (analogizing war against sex offenders to war on drugs).

15. E.g., Laura Crimaldi, *Pols to Target Perv Law Loopholes: Bills Urge GPS, Parole for Life*, BOS. HERALD, Feb. 24, 2008, at 4 (identifying specific cases of harm caused by released sex offenders and

cases of sexual violence and incorrectly implied that sex crimes are mostly accomplished by fearsome strangers.¹⁶ Through repeated publication of their names and photographs, telegenic victims of sex crimes, particularly those young and cute, literally become the “poster children” for the moral panic and public demands for officials to do something to protect potential future victims.¹⁷ Likewise, the media has also hyped stories involving repeat sex offenders thereby leading to the iconic image of the recidivist sexual predator, i.e., the SVP.

Researchers have studied the cyclical nature of the moral panic to help explain its endurance. In one study, surveyed legislators admitted that their opinions on the need for restrictive measures for SVPs were informed greatly by media reports and constituent concerns.¹⁸ A terse comment by a legislator is telling: “You hear about these guys raping and killing kids all the time now. We have to do something. It’s gotta-stop.”¹⁹ In turn, with new SVP legislation, the media responds with further coverage reporting the government’s admission about the particular dangers imposed by the presence of SVPs in communities, thereby continuing the cycle and reinforcing the image of the menacing sexual predator. This rhetoric underlying the moral panic continues, despite contradictory evidence. Official statistics show that in the United States, rates of rape/sexual assault declined by approximately fifty-three percent between 1999 and 2008.²⁰ Similarly, the rate of child sexual abuse in the United States dropped by an estimated fifty-eight percent from 1992 to 2008.²¹ Further, sex offenders are not more likely than other types of

describing forty bills then pending before Massachusetts state legislature that would impose more stringent restrictions on sex offenders); Laura Mansnerus, *Stoking ‘Moral Panic’ Over Sex Offenders*, N.Y. TIMES, May 29, 2005, at 14NJ-2 (describing media hype over sex offenders using Medicare payments to purchase Viagra and how horror stories about pedophiles prompted legislative calls for stricter sex offender laws despite existence of other important pending issues).

16. CHRIS GREER, *SEX CRIME AND THE MEDIA: SEX OFFENDING AND THE PRESS IN A DIVIDED SOCIETY* 185–86 (2003); see also Carissa Byrne Hessick, *Disentangling Child Pornography from Child Sex Abuse*, 88 WASH. U. L. REV. 853, 887–88 (2011) (explaining that misconceptions about sex offenders stem from “stories of young children who were sexually assaulted and killed by strangers”).

17. See e.g., Adam Walsh Child Protection Act of 2006, Pub. L. No. 109-248, 120 Stat. 587 (codified as amended in scattered sections of 42 U.S.C. (Supp. 2010)); Megan’s Law, N.J. Stat. Ann. § 2C:7-1-7-11 (West 2005 & Supp. 2010); Jimmy Ryce Involuntary Civil Commitment for Sexually Violent Predators’ Treatment and Care Act, Fla. Stat. § 394.910 (2009); see also Catherine L. Carpenter, *Legislative Epidemics: A Cautionary Tale of Criminal Laws That Have Swept the Country*, 58 BUFF. L. REV. 1, 23 (2010) (contending that use of child victim’s name personalizes legislation and conveys sense of urgency); Wayne A. Logan, *Megan’s Laws as a Case Study in Political Stasis*, 61 SYRACUSE L. REV. 371, 373–80 (2011) (discussing high-profile cases that have influenced the enactment of sex offender legislation in the past three decades); Michael Vitiello, *Punishing Sex Offenders: When Good Intentions Go Bad*, 40 ARIZ. ST. L.J. 651, 676 (2008) (maintaining that sex offender “[p]olicies are crafted in fearful haste, often as symbolic gestures to honor the crime victims whose suffering has inspired them” (quoting FRANKLIN E. ZIMRING, *AN AMERICAN TRAVESTY: LEGAL RESPONSES TO ADOLESCENT SEXUAL OFFENDING* xiii (2004))).

18. Lisa L. Sample & Colleen Kadleck, *Sex Offender Laws: Legislators’ Accounts of the Need for Policy*, 19 CRIM. JUST. POL’Y REV. 40, 57 (2008).

19. *Id.* at 51 (quoting anonymous interviews of legislative officials).

20. MICHAEL R. RAND, U.S. DEP’T OF JUSTICE, *CRIMINAL VICTIMIZATION, 2008*, at 2 (2009), available at <http://bjs.ojp.usdoj.gov/content/pub/pdf/cv08.pdf>.

21. DAVID FINKELHOR ET AL., *CRIMES AGAINST CHILDREN RESEARCH CTR., UPDATED TRENDS IN CHILD MALTREATMENT, 2008*, at 1 fig.1 (2010), http://www.unh.edu/ccrc/pdf/CV203_Updated%20Trends%20in

offenders to either recidivate generally or to specialize in committing sex offenses.²²

So, in light of legislative convergence upon new types of preventive measures as applying only to sex offenders, one might reasonably consider: Why the special treatment? Many suggest that citizens view sexual deviance as qualitatively unique. While most offense-based laws regulate risky behavior, SVP laws focus on regulating risky persons, as if the sexual violence makes the sexual offender himself a unique characteristic.²³ Citizens commonly believe that sexual offenders are most likely to be strangers to the victim and more likely to reoffend than other types of criminals.²⁴ Experts indicate that the public tends to believe that sexualized violence has actually become more pervasive in recent years and that it is causing greater harm to society.²⁵ This is particularly true with the proliferation of technology as media portrayals fuel largely mythical fears regarding SVPs using the internet to lure victims.²⁶

As a result, legislators have articulated a number of responses to deter sexual violence through criminal laws, such as creating new criminal offenses and imposing longer sentences for sexual offenses, as well as other measures aimed at managing the sexual offender population.²⁷ In large part, complicity between public fear, media hype,

%20Child%20Maltreatment%202008_8-6-10.pdf (reporting statistics from the National Child Abuse and Neglect Data System, which aggregates data from state child protection agencies).

22. See Megan Magers et al., *An Exploration of the Sex Offender Specialization and Violence Nexus*, 6 SW. J. OF CRIM. JUST. 133, 141 (2009) (“While the general public and policymakers typically view sex offenders as persistent specialists and recidivists, the literature reviewed in this study and the findings from the present study suggest that sex offenders engage in a variety of criminal behaviors and are a heterogeneous group of offenders.”); Leonore M.J. Simon, *The Myth of Sex Offender Specialization: An Empirical Analysis*, 23 NEW ENG. J. ON CRIM. & CIV. CONFINEMENT 387, 401 (1997) (concluding that there is “no evidence that child molesters (or sex offenders in general) have higher recidivism rates or are in fact more dangerous than other types of offenders”).

23. See JANUS, *supra* note 3, at 4 (discussing how “predator archetype” hurts efforts to prevent sexual violence); see also Janus & Polachek, *supra* note 13, at 145 (contending that society condemns sexual predators for their internal characteristic that contaminates them).

24. Jill S. Levenson et al., *Public Perceptions About Sex Offenders and Community Protection Policies*, 7 ANALYSES OF SOC. ISSUES & PUB. POL’Y 137, 153–54 (2007); see also Ruth E. Mann et al., *Assessing Risk for Sexual Recidivism: Some Proposals on the Nature of Psychologically Meaningful Risk Factors*, 22 SEXUAL ABUSE 191, 192 (2010) (citing studies supporting contentions that “[t]he observed sexual recidivism rate of sexual offenders is less than commonly believed” and “the overall recidivism rate of sexual offenders is lower than the recidivism rate of other offender groups”).

25. Janus & Prentky, *supra* note 2, at 90; Erica Beecher-Monas & Edgar Garcia-Rill, *Genetic Predictions of Future Dangerousness: Is There a Blueprint for Violence?*, 69 LAW & CONTEMP. PROBS. 301, 317–18 (2006).

26. Compare 152 CONG. REC. H5723 (daily ed. July 25, 2006) (statement of Rep. Sensenbrenner) (referring to growing problem of online sexual predators) with Janis Wolak et al., *Online “Predators” and Their Victims: Myths, Realities, and Implications for Prevention and Treatment*, 63 AM. PSYCHOLOGIST 111, 112–13, 119 (2008) (reporting that public fears are erroneous since online predators are usually open about their sexual interest and are rarely violent, suggesting that prevention may be better served by teaching young people safe internet practices).

27. As mandated by the Adam Walsh Child Protection Act of 2006, the U.S. Department of Justice created the Office of Sex Offender Sentencing, Monitoring, Apprehending, Registering, and Tracking, whose mission is to “protect the public.” *About SMART*, OFFICE OF SEX OFFENDER SENTENCING, MONITORING, APPREHENDING, REGISTERING, AND TRACKING, <http://www.ojp.usdoj.gov/smart/about.htm> (last visited June 21, 2011).

and political pandering works efficiently because of the imbalance in the debate. Families of murdered victims of sex crimes act as informal lobbyists for increasing control of sex offenders in what is a uniquely bipartisan political environment on the issue. There is virtually no counter movement to represent the interests of those alleged to be SVPs, who are largely reviled by all.²⁸ Often, the legislative drive in imposing new preventive measures on SVPs is espoused in assertions about the prevalence and danger of sex offenders, often without any empirical support.²⁹

Because of the widespread fear of SVPs, concerns have focused on the likelihood of sexual recidivism after convicted sex offenders are released from imprisonment. The management of those convicted of sex-based crimes after release has been largely effectuated through the enforcement of civil (rather than criminal) laws. The use of civil regulations as management tools has the distinct benefit of permitting criminal justice officials to restrict the freedom and privacy of sex offenders without abiding by the stricter procedural requirements that would be constitutionally required in the criminal law arena. The civil law-based tools also permit state officials to monitor and supervise sex offenders beyond traditional parole and probation structures. The most common types of civil laws used today to manage the sex offender population are SVP civil commitment laws, registration requirements, and residency restrictions.³⁰

A. SVP Laws

1. Civil Commitment Laws

SVP civil commitment statutes are a distinct species of traditional civil commitment laws that permit a state to commit a person to a mental institution who is mentally ill and who poses a danger to others. The Kansas SVP law is typical. It

28. See John Douard, *Sex Offender as Scapegoat: The Monstrous Other Within*, 53 N.Y.L. SCH. L. REV. 31, 38–39 (2009) (noting that sex offenders are constructed as monsters and thus have few advocates); Lester, *supra* note 11, at 347–48 (stating that sex offenders are abhorred even by other criminals).

29. See, e.g., FLA. STAT. ANN. § 775.21(3)(A) (West 2010) (“Repeat sexual offenders, sexual offenders who use physical violence, and sexual offenders who prey on children are sexual predators who present an extreme threat to the public safety. Sexual offenders are extremely likely to use physical violence and to repeat their offenses, and most sexual offenders commit many offenses, have many more victims than are ever reported, and are prosecuted for only a fraction of their crimes. This makes the cost of sexual offender victimization to society at large, while incalculable, clearly exorbitant.”); 140 CONG. REC. H5612 (daily ed. July 13, 1994) (statement of Rep. Dunn) (campaigning on congressional floor for registration and notification bill and claiming that rate of recidivism of sex offenders is “astronomical. We know that.”); 152 CONG. REC. H5725 (daily ed. July 25, 2006) (statement of Rep. Foley) (referring to sexual predators as “growing and dangerous threat to our children”). Representative Mark Foley, shortly after making this public statement, was caught in a scandal for sending pornographic e-mails to teenage male pages and soon thereafter resigned. Ellen Goodman, *You Can't Make This Stuff Up: The Self-Proclaimed Party of Moral Values Can't Keep Its Own House in Order*, PITTSBURGH POST-GAZETTE, Oct. 6, 2006, at B-7.

30. Other, though less common, laws specially designed for sex offenders include: bans on certain employment, such as employment at day care facilities, IDAHO CODE ANN. § 18-8327 (2011); restrictions on working near specified locations, such as schools, ALA. CODE § 15-20-26(a) (LexisNexis 2011); GPS monitoring, GA. CODE ANN. § 42-1-14(e) (2009), LA. REV. STAT. ANN. § 560.4(A) (2011); and chemical castration, Matthew V. Daley, *A Flawed Solution to the Sex Offender Situation in the United States: The Legality of Chemical Castration for Sex Offenders*, 5 IND. HEALTH L. REV. 87 (2008) (listing statutes).

defines an SVP as a person who meets three main criteria: 1) the person has been convicted or charged with a sexually violent offense, and 2) the person “suffers from a mental abnormality or personality disorder” which 3) “makes the person likely to engage in repeat acts of sexual violence” if not confined in a secure facility.³¹ A person adjudged to be an SVP under this law is then committed for an indefinite period in a secure institution, despite not having committed a new offense. In a ground-breaking decision, the U.S. Supreme Court in the case of *Kansas v. Hendricks*³² upheld the Kansas SVP law in the face of constitutional challenges.³³ The Court dismissed Hendricks’ claims that the law violated the Double Jeopardy and Ex Post Facto Clauses.³⁴ The starting point for the analysis was that the Double Jeopardy and Ex Post Facto Clauses apply only to laws that are intended as punishment.³⁵ The majority ruled that civil commitment is civil in nature and since it is not intended for deterrent or retributive purposes, it does not constitute punishment.³⁶

Post-*Hendricks*, most SVP civil commitment statutes have been modeled after the Kansas statute.³⁷ Being civil in nature, defendants in civil commitment proceedings do not enjoy many of the benefits of the Fifth and Sixth Amendments, such as freedom from self-incrimination, jury trial, or confrontation of witnesses, but are generally provided some right to counsel.³⁸

In practice, most states use SVP civil commitment laws when a sex offender is about to be released from prison. The idea is to allow the offenders to serve their sentences under normal custodial arrangements but then to transfer them to other secure accommodations just before their scheduled release dates so that the offenders are actually never freed. Currently, twenty states and the federal government have implemented SVP civil commitment laws, and a recent estimate is that over 4,500 SVPs are held in civil commitment facilities under these laws.³⁹ Once committed, it is very difficult for defendants to ever be released.⁴⁰

31. KAN. STAT. ANN. § 59-29a02 (2011).

32. 521 U.S. 346 (1997).

33. *Hendricks*, 521 U.S. at 371.

34. *Id.*

35. *Id.* at 360–62.

36. *Id.* at 361–63.

37. John Matthew Fabian, *The Risky Business of Conducting Risk Assessments for Those Already Civilly Committed as Sexually Violent Predators*, 32 WM. MITCHELL L. REV. 81, 96 (2005).

38. *United States v. Abregana*, 574 F. Supp. 2d 1123, 1139–41 (D. Haw. 2008); *see also* Carpenter, *supra* note 17, at 52 (arguing that Supreme Court’s openness to describing civil commitment as civil in nature has led legislators to take advantage by “imposing increasingly harsh regiments for the convicted sex offender,” including registration and residency restrictions).

39. KATHY GOOKIN, WASH. STATE INST. FOR PUB. POLICY, COMPARISON OF STATE LAWS AUTHORIZING INVOLUNTARY COMMITMENT OF SEXUALLY VIOLENT PREDATORS: 2006 UPDATE, REVISED 1 (2007), available at <http://www.wsipp.wa.gov/rptfiles/07-08-1101.pdf>. *But see* Monica Davey & Abby Goodnough, *Doubts Rise as States Hold Sex Offenders After Prison*, N.Y. TIMES, Mar. 4, 2007, at 1-1 (reporting that approximately 2,700 sex offenders subject to civil commitment).

40. Eric. S. Janus & Robert A. Prentky, *Forensic Use of Actuarial Risk Assessment with Sex Offenders: Accuracy, Admissibility and Accountability*, 40 AM. CRIM. L. REV. 1443, 1447 (2003) (finding only five percent of committed offenders had been cleared for release).

2. Registration Requirements

Sex offender registration systems are public information devices. All fifty states, as well as the District of Columbia, have mandated sex offender registries.⁴¹ Much of the information is widely available to the public on freely accessible internet sites. The type of information that sex offenders are required to provide has expanded dramatically over the years. As an example, Georgia requires the registrant's name, date of birth, height, weight, fingerprints, photo, residence, employment information, and vehicle details.⁴² After initial registration, sex offenders are generally required to update the information from time to time.⁴³ Once a person is labeled an SVP for registration purposes, it may be quite difficult to challenge the designation, even on grounds that the label was erroneously imposed.⁴⁴

3. Residency Restrictions

States and local jurisdictions are also active in passing laws banning sex offenders from residing in ever larger swaths of areas. Unlike the other two SVP schemes that tend to be federal or statewide in scope, residency restrictions are more likely to be enacted at the local level by town or city councils.⁴⁵ The resulting multiplicity of regulations has resulted in a wide variety of restrictions that apply to different types of offenders. Some laws apply to all sex offenders⁴⁶ while others target just violent sex offenders or those who have committed sexual crimes against children.⁴⁷ In some states, residency restrictions apply even to offenders convicted of non-contact sex-based offenses, such as possession of pornography.⁴⁸ A common criterion is to delineate a certain area where the sex offender is not permitted to reside, such as 1,000 or 2,000 feet around specified locations like schools, parks, and bus stops.⁴⁹ These policies have resulted in effectively banning sex offenders from residing in whole

41. Amber Leigh Bagley, Comment, "An Era of Human Zoning": Banishing Sex Offenders from Communities Through Residence and Work Restrictions, 57 EMORY L.J. 1347, 1348 (2008).

42. GA. CODE ANN. § 42-1-12(a)(16) (2009).

43. Cassie Dallas, Comment, *Not in My Backyard: The Implications of Sex Offender Residency Ordinances in Texas and Beyond*, 41 TEX. TECH L. REV. 1235, 1243 (2009).

44. See Shawndra Jones, Note, *Setting Their Record Straight: Granting Wrongly Branded Individuals Relief from Sex Offender Registration*, 41 COLUM. J.L. & SOC. PROBS. 479, 496-97 (2008) (explaining that registries have deficient methods for ensuring accuracy).

45. Some courts, however, have overturned local regulations as preempted by state statute. See, e.g., *Fross v. Cnty. of Allegheny*, 612 F.Supp.2d 651, 660 (W.D. Pa. 2009) (overturning Allegheny residency restriction ordinance as preempted by state law).

46. E.g., ALA. CODE § 15-20-26(a) (LexisNexis 2011); GA. CODE ANN. § 42-1-15(a)-(b) (2009); KY. REV. STAT. ANN. § 17.545(1) (West 2008); OKLA. STAT. tit. 57 § 590(A) (2008).

47. E.g., FLA. STAT. ANN. § 947.1405(7)(a)(2) (LexisNexis 2011) (placing residency restrictions on certain offenders that committed specified crimes against children); LA. REV. STAT. ANN. §§ 14:91.2(A)(1)-(4), 15:538(D)(1)(a)-(c) (2011) (placing residency restrictions on convicted sex offenders if their victim was "under the age of thirteen years" and "whose offense involved a minor child").

48. Lester, *supra* note 11, at 344-45 (discussing Alabama statute).

49. *Id.* at 351-52 (noting that nineteen states have enacted these forms of residency restrictions).

towns, cities, and counties.⁵⁰ Further, lifetime residency restrictions are now common.⁵¹

A few items of note apply generally to both registration and residency laws. While many states require registration per se after a conviction for one of many enumerated sex offenses, about half of the state laws utilize a future dangerousness assessment model⁵² to individually tailor the exact nature of the registration requirements.⁵³ Similarly, while many residency restrictions apply per se to those convicted of specified sex offenses, many of the laws correspond to the registration system by invoking residency restrictions on registered sex offenders. A violation of these laws most often constitutes a new criminal offense and a new prison sentence as a consequence.⁵⁴ Registration requirements and residency restrictions have often been applied retroactively, meaning applying to offenders convicted prior to the laws' enactment.⁵⁵ Facing claims that the laws violate the Ex Post Facto Clause, the U.S. Supreme Court has ruled that registration is civil in nature and, thus, the clause is inapplicable.⁵⁶ While the question of the constitutionality of residency restrictions has

50. Asmara Tekle-Johnson, *In the Zone: Sex Offenders and the Ten-Percent Solutions*, 94 IOWA L. REV. 607, 621 (2009) (listing areas where sex offenders have been effectively banned from residing); Kari White, Note, *Where Will They Go? Sex Offender Residency Restrictions as Modern-Day Banishment*, 59 CASE W. RES. L. REV. 161, 168–69 (2008) (showing maps whereby officials highlighted few areas remaining where sex offenders could reside after factoring in residency restrictions).

51. See Lester, *supra* note 11, at 374–80 tbl.1 (listing statutes with information about geographic and durational scope of restrictions).

52. COUNCIL OF STATE GOV'TS, SEX OFFENDER MANAGEMENT POLICY IN THE STATES: STRENGTHENING POLICY & PRACTICE 6 (2010), available at <http://www.csg.org/policy/documents/SOMFinalReport-FINAL.pdf>.

53. The federal Sex Offender Registration and Notification Act of 2006 (SORNA) may change this. SORNA would require states, under penalty of losing certain funding, to follow a unitary system of sex offender registration based not on individual assessments of future risk, but on the type of offense or conviction. Generic titles of offenses are categorized into three levels and the requisite amount of time to be registered is based on the applicable level. Still, there is reason to believe that the SORNA requirements will not significantly overhaul the registration system and entirely replace future dangerousness assessments. SORNA's registration requirements were originally intended to be substantially implemented by states as of July 27, 2009. But, within weeks of the deadline, few states had acted to implement the regulations because of its burden and cost and, with pressure, the Attorney General delayed the deadline a year. News Release from Office of Sen. Patrick S. Leahy, *Sen. Leahy Applauds Extension for State Compliance with Sex Offender Registration and Notification Act*, U.S. FED NEWS, June 3, 2009. As of July 27, 2011, only fourteen states had "substantially implemented" SORNA's requirements. Press Release, Dep't of Justice, Justice Department Finds 24 Jurisdictions Have Substantially Implemented SORNA requirements (July 28, 2011), available at http://www.ojp.usdoj.gov/newsroom/pressreleases/2011/SMART_PR-072811.htm.

54. See Lester, *supra* note 11, at 381–84 tbl.2 (listing statutes with punishment information); GA. CODE ANN. § 42-1-15 (2008) (penalizing violation with up to thirty-year sentence).

55. Caleb Durling, Comment, *Never Going Home: Does It Make Us Safer? Does It Make Sense?: Sex Offenders, Residency Restrictions, and Reforming Risk Management Law*, 97 J. CRIM. L. & CRIMINOLOGY 317, 323 (2006).

56. *Smith v. Doe*, 538 U.S. 84, 105–06 (2003). Cf. Corey Rayburn Yung, *The Disappearing Ex Post Facto Clause: From Substantive Bulwark to Procedural Nuisance*, 61 SYRACUSE L. REV. 447, 454–58 (2011) (explaining how legislatures and judiciaries circumvent valid ex post facto challenges to sex offender laws).

not risen to the highest level, lower courts have generally denied such claims on the same basis.⁵⁷

Critics, meanwhile, continue to decry that SVP laws are punitive in nature since they serve the legislative and public thirst for retribution against sex offenders⁵⁸ and are highly restrictive measures used in addition to prison sentences. Moreover, future dangerousness assessments—which have “begun to colonize our theories of punishment”⁵⁹—signify a broader policy shift toward incapacitation as the key method for preventing crime, a trend which reflects

that we have given up on trying to reduce crime by investing in job opportunities, education, assistance to immigrants, drug rehabilitation programs, reentry programs, and the like. . . . In order to become evermore efficient, we develop actuarial methods to determine who should be exiled to prison and for how long.⁶⁰

B. (Un)Intended Consequences of Sexual Predator Laws

In conceptualizing how well SVP laws fit the preventive law model, two key issues emerge. One is whether the hype about the dangerous sexual predator is accurate. The second is whether these restrictions do more harm than good to those to whom they are applied and to the public welfare in general. While it is beyond the scope of this Article to fully explore these issues, it seems relevant to at least mention them here in order to highlight why the potential misuse of actuarial risk tools in SVP litigation is problematic to the notions of law and justice.

1. Sexual Predators and Stranger Danger: Is the Concern Valid?

Much empirical evidence indicates that the iconic image of the SVP is more mythical than real. A natural starting point is the relatively simplistic statistic of the base rate. The base rate for recidivism essentially means the overall rate of recidivism actually observed in a group of offenders.⁶¹ Before reviewing results from official

57. See e.g., *Doe v. Miller*, 405 F.3d 700, 719 (8th Cir. 2005) (starting with presumption that legislative intent was to create civil non-punitive regulatory scheme); *People v. LeRoy*, 828 N.E.2d 769, 778–82 (Ill. App. Ct. 2005) (finding residency restrictions at issue to be civil, not punitive, and thus not an ex post facto law).

58. See Lindsay A. Wagner, Comment, *Sex Offender Residency Restrictions: How Common Sense Places Children at Risk*, 1 DREXEL L. REV. 175, 176 (2009) (discussing trend toward retributive, punitive justifications for punishment); see also Douard, *supra* note 28, at 46–47 (noting that conditions in which sex offenders serve civil commitment are “extremely punitive” in nature); John A. Fennel, *Punishment by Another Name: The Inherent Overreaching in Sexually Dangerous Person Commitments*, 35 NEW ENG. J. ON CRIM. & CIV. CONFINEMENT 37, 61–62 (2009) (finding that sex offender civil commitment laws serve society’s desires for penance and just deserts).

59. BERNARD E. HARCOURT, *AGAINST PREDICTION: PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE* 188 (2007).

60. Yoav Sapir, *Against Prevention? A Response to Harcourt’s Against Prediction on Actuarial and Clinical Predictions and the Faults of Incapacitation*, 33 LAW & SOC. INQUIRY 253, 260–261 (2008) (book review).

61. CTR. FOR SEX OFFENDER MGMT., *RECIDIVISM OF SEX OFFENDERS 5* (2001) [hereinafter CSOM ON RECIDIVISM], available at <http://www.csom.org/pubs/recidsexof.pdf>.

studies, it is important to consider that base rates can vary by study. Three important methodological choices between studies can help explain the variation: the defining characteristic of recidivism (the operationalization of the outcome variable); the time period for observation; and the demographic characteristics of the group observed. For example, if the outcome variable is arrest then the base rate will naturally be higher than if the outcome requires a conviction. The longer period of time for follow-up will yield higher numbers since there is more opportunity to reoffend. Another consideration is that a group that is at higher risk of recidivism (e.g., younger offenders), who reoffend more often than older offenders, will yield a higher base rate because of the attribute (e.g., youth) that is correlated with higher risk.

In a prominent base-rate study sponsored by the U.S. Department of Justice, researchers conducted a three-year follow-up of over 9,600 male sex offenders released in 1994.⁶² They found that 5.3% of the released offenders had been rearrested for a sex crime within three years of their release.⁶³ This rate was much higher than the 1.3% rearrest statistic on sexual offenses for the non-sex offenders who were released the same year.⁶⁴ However, comparing the groups committing new non-sex offenses, non-sex offenders had a 68% rearrest rate overall, compared to the 43% rearrest rate for sexual offenders.⁶⁵ Another study of male offenders released from prisons in Massachusetts in 2002 also compared recidivism rates (there defined as any reincarceration or parole violation) of offenders who had been incarcerated for sex offenses with offenders who had been incarcerated for non-sex offenses.⁶⁶ The recidivism base rate among sex offenders for any offense was 22%, compared to base rates of 32% to 57% for other types of offenders (property, person, drug, and other).⁶⁷ If sex offenders are at minimal risk of sexually re-offending, and are at much lower risk of general recidivism than non-sex offenders, then specialized policies do not appear to be justified.⁶⁸

Other studies of sex offenders also show relatively small percentages of sex-crime recidivism after release. The rate for an Iowa study of a demographically mixed group

62. PATRICK A. LANGAN ET AL., U.S. DEP'T OF JUSTICE, *RECIDIVISM OF SEX OFFENDERS RELEASED FROM PRISON IN 1994*, at 1 (2003), <http://bjs.ojp.usdoj.gov/content/pub/pdf/rsorp94.pdf>.

63. *Id.* at 24.

64. *Id.*

65. *Id.* at 2.

66. RHIANA KOHL ET AL., URBAN INST., *MASSACHUSETTS RECIDIVISM STUDY: A CLOSER LOOK AT RELEASES AND RETURNS TO PRISON 14* (2008), available at http://www.urban.org/UploadedPDF/411657_massachusetts_recidivism.pdf.

67. *Id.*

68. See CSOM ON RECIDIVISM, *supra* note 61, at 1 (suggesting that if recidivism research shows that sex offenders do not specialize in committing sex offenses, the public response for released sex offenders should be no different than for released non-sex offenders); Tucker Culbertson, *After Comstock: Equal Protection Challenges to the Civil Commitment Provisions of the Adam Walsh Child Protection and Safety Act*, 61 SYRACUSE L. REV. 427, 445–46 (2011) (“Though the federal government’s interest in preventing sexual offenses is compelling, [sex offender civil commitment]—as a means by which to pursue that interest—is irrational because it is based on intense emotions, social stigma, and overbroad categorizations of ‘sexual’ and ‘non-sexual’ behavior, rather than proven evidence. As such, it fails even the lowest form of equal protection review.”).

(meaning children, adults, men, and women) was 3% after approximately four years.⁶⁹ Similarly, the rate was 3% in Washington after about five years,⁷⁰ and 8% in Ohio at ten-years.⁷¹ Nonetheless, the base rates reported in the foregoing studies are far below what the public seems to believe. According to one study, for example, the public placed the recidivism rate of sex offenders at 75%, a dramatic overestimation.⁷²

While base rates of sexual recidivism are relatively low in official studies, there are other reasons base rates may vary. Sex offenders are a heterogeneous grouping, and their variances can have substantial effect on the commission of sex offenses. For example, there is evidence of differing recidivism rates in subgroups based on factors such as age, gender, prior incarceration, whether the offender is receiving treatment, and type of sex offense.⁷³ Still, the foregoing studies involving mixed populations of predominantly male offenders together are strong evidence that the vast majority, up to 95% of released sex offenders in the national study, do not sexually reoffend, at least during the follow-up periods studied.⁷⁴ Empirical studies thus tend to show that the image of the reoffending sexual predator is, while not mythical, not as accurate as the media and politicians assert.

2. Do Sexual Offender Laws Work?

Even if it were true that sexual offenders pose the extreme risk suggested by the sexual predator symbolism, can and do these SVP laws work in preventing future

69. GENEVA ADKINS ET AL., IOWA DEP'T OF HUMAN RIGHTS, THE IOWA SEX OFFENDER REGISTRY AND RECIDIVISM 4, 10 (2000), available at http://www.iowa.gov/dhr/cjpp/images/pdf/01_pub/SexOffenderReport.pdf (defining recidivism as reconviction with sample of over 400 sex offenders on parole or probation).

70. ROBERT BARNOSKI, WASH. STATE INST. FOR PUB. POLICY, SEX OFFENDER SENTENCING IN WASHINGTON STATE: SEX OFFENDER RISK LEVEL CLASSIFICATION TOOL AND RECIDIVISM 2 (2006), available at <http://www.wsipp.wa.gov/rptfiles/06-01-1204.pdf> (defining recidivism as reconviction for a felony with a sample of almost 700 released sex offenders).

71. STATE OF OHIO OFFICE OF CRIMINAL JUSTICE SERVS., REPORT TO THE OHIO CRIMINAL SENTENCING COMMISSION: SEX OFFENDERS 13–14 (2006) [hereinafter OHIO COMMISSION], available at http://www.publicsafety.ohio.gov/links/ocjs_SexOffenderReport.pdf (defining recidivism as being reincarcerated after release with an otherwise unidentified sample).

72. Timothy Fortney et al., *Myths and Facts about Sexual Offenders: Implications for Treatment and Public Policy*, 2 SEXUAL OFFENDER TREATMENT 1, 4 (2007), <http://www.sexual-offender-treatment.org/index.php?id=55&type=123>; see also OHIO COMMISSION, *supra* note 71, at 13 (admitting “[i]t is a common misperception that sex offenders have a high recidivism rate”).

73. One recent study reported sexual recidivism rates, approximately three years after release, of 5% for rapists, 11% for incest offenders, 14% for extrafamilial child molesters, and 35% for non-contact offenders. James Vess & Alex Skelton, *Sexual and Violent Recidivism by Offender Type and Actuarial Risk: Reoffending Rates for Rapists, Child Molesters and Mixed-Victim Offenders*, 16 PSYCHOL. CRIME & L. 541, 542 (2010) (citing D.L. Bartosh, et al., *Differences in the Predictive Validity of Actuarial Risk Assessments in Relation to Sex Offender Type*, 47 INT'L J. OFFENDER THERAPY & COMP. CRIMINOLOGY 422 (2003)). As noted by Vess and Skelton, the composition of a sex-offender recidivism study, in terms of the subgroups of offenders included in the sample, is “perhaps the most salient factor contributing to the various recidivism rates” reported in the literature. *Id.*

74. See LANGAN ET AL., *supra* note 62, at 1 (“Released sex offenders with 1 prior arrest (the arrest for the sex crime for which they were imprisoned) had the lowest rearrest rate for a sex crime, about 3%; those with 2 or 3 prior arrests for some type of crime, 4%; 4 to 6 prior arrests, 6%; 7 to 10 prior arrests, 7%; and 11 to 15 prior arrests, 8%.”).

sexual violence? Presumably, civil commitment is effective, though at enormous cost to state coffers and to liberty and privacy interests.⁷⁵ The efficacy of registration and residency laws is far more questionable. An assumption of registration and residency laws is that dangerous predators are strangers and that the public can protect itself by being aware of who these people are and controlling where they live. Statistics belie this assumption. Of the rapes and sexual assaults reported in the National Crime Victimization Survey for 2008, 63% of female victims and 100% of male victims were attacked by non-strangers.⁷⁶ Another national study using police reports from a five-year period in the 1990s indicates that assaultive sex offenses committed against child victims (less than eighteen years old) are largely committed by family members (about 34%) and acquaintances (about 59%).⁷⁷ In children under six years, the statistic is more depressing, where almost half of the offenders were family members.⁷⁸ In other words, according to these studies, the majority of sexual assaults are perpetrated by persons already known to their victims.⁷⁹ The obvious supposition, then, is that SVP laws are not capable of protecting many victims of sexual violence as most are those with whom offenders are already familiar and may have continuing access.⁸⁰

Another basic theoretical problem regarding the registration system is that it naively presumes that the information is current,⁸¹ and that citizens access the information and act thereon to protect themselves.⁸² A recent study found that while a large proportion of survey respondents were aware of the existence of their state's public sex offender data file, most had never accessed it and, of those who did, few took any precautionary measures as a result.⁸³

Considering a more direct perspective on whether sex offender laws work, studies continue to show that neither sex offender registry implementation nor residency

75. See Jeslyn A. Miller, Comment, *Sex Offender Civil Commitment: The Treatment Paradox*, 98 CALIF. L. REV. 2093, 2108–22 (2010) (demonstrating that, while states purport to use civil commitment for treatment purposes, civil commitment in fact harms the offender).

76. RAND, *supra* note 20, at 5 tbl.6.

77. HOWARD N. SNYDER, U.S. DEP'T OF JUSTICE, SEXUAL ASSAULT OF YOUNG CHILDREN AS REPORTED TO LAW ENFORCEMENT: VICTIM, INCIDENT, AND OFFENDER CHARACTERISTICS 10 (2000), available at <http://bjs.ojp.usdoj.gov/content/pub/pdf/saycrle.pdf>.

78. *Id.* Strangers represent only about three percent of those who commit sexual assaults against victims under the age of six. *Id.*

79. See RAND, *supra* note 20, at 5 (reporting that, in 2008, sixty-three percent of reported rapes or sexual assaults against women were committed by non-strangers).

80. James Vess, *Sex Offender Risk Assessment: Consideration of Human Rights in Community Protection Legislation*, 13 LEGAL & CRIMINOLOGICAL PSYCHOL. 245, 246 (2008).

81. See, e.g., Jill S. Levenson & Leo P. Cotter, *The Effect of Megan's Law on Sex Offender Reintegration*, 21 J. CONTEMP. CRIM. JUST. 49, 54, 61 (2005) (reporting on study of 180+ registered sex offenders, half of whom reported that information about them in state public registry was incorrect).

82. Rachel Bandy, *Measuring the Impact of Sex Offender Notification on Community Adoption of Protective Behaviors*, 10 CRIMINOLOGY & PUB. POL'Y 237, 255–56 (2011) (“[T]his study found no statistically significant relationship between receiving notification about a high-risk sex offender and the adoption of self-protective behaviors . . .”).

83. Amy L. Anderson & Lisa L. Sample, *Public Awareness and Action Resulting from Sex Offender Community Notification Laws*, 19 CRIM. JUST. POL'Y REV. 371, 388 (2008).

restrictions actually reduce recidivism.⁸⁴ A likely explanation is that there appears little doubt that SVP laws conflict with everything that has been learned in the past few decades about successful reentry efforts. These laws interfere with those important factors that have been shown to mitigate reoffending, such as support from family, friends, and the community; maintaining a job; and having a healthy place to live.⁸⁵

For example, when the bans involve large areas, sex offenders have congregated in discrete areas, often socioeconomically and socially challenged neighborhoods, or become homeless.⁸⁶ The label of SVP and the collateral consequences of sexual predator laws also bring social scorn, loss of social support, reduction in employment possibilities, housing difficulties, and personal harassment.⁸⁷ Experts believe that the multiple pressures and lack of legitimate opportunities increase the risk of recidivism.⁸⁸

84. Jeffrey C. Sandler et al., *Does a Watched Pot Boil? A Time-Series Analysis of New York State's Sex Offender Registration and Notification Law*, 14 PSYCHOL. PUB. POL'Y & L. 284, 297 (2008); see also ADKINS ET AL., *supra* note 69, at 10 (finding no statistically significant difference between pre-registry and post-registry crime rates for sex or non-sex offenses among released sex offenders); BARNOSKI, *supra* note 70, at 2–3 (finding that level of risk classification of sex offenders and existence of notification program had no impact on recidivism rates); MINN. DEP'T OF CORR., RESIDENTIAL PROXIMITY & SEX OFFENSE RECIDIVISM IN MINNESOTA 24 (2007), available at <http://www.doc.state.mn.us/publications/documents/04-07SexOffenderReport-Proximity.pdf> (concluding after tracking 224 sex offenders released between 1990 and 2002 who were later reincarcerated for a sex crime that “not one of the 224 offenses would likely have been affected by residency restrictions”); JEFFERY T. WALKER ET AL., ARK. CRIME INFO. CTR., THE INFLUENCE OF SEX OFFENDER REGISTRATION AND NOTIFICATION LAWS IN THE UNITED STATES 15 (2004) (finding no “clear unidirectional conclusion as to whether sex offender notification laws prevent rapes”); Elizabeth J. Letourneau & Kevin S. Armstrong, *Recidivism Rates for Registered and Nonregistered Juvenile Sexual Offenders*, 20 SEXUAL ABUSE 393, 400 (2008) (reporting that in matched sample of 222 registered and nonregistered juvenile offenders, the only two sex offenses occurred among registered group); Richard Tewksbury & Wesley G. Jennings, *Assessing the Impact of Sex Offender Registration and Community Notification on Sex-Offending Trajectories*, 37 CRIM. JUST. & BEHAV. 570, 579 (2010) (finding no impact of registration and notification laws on recidivism rates among sample of released sex offenders); Bob Edward Vásquez et al., *The Influence of Sex Offender Registration and Notification Laws in the United States: A Time-Series Analysis*, 54 CRIME & DELINQ. 175, 188 (2008) (concluding that registration laws have had no overall measurable or consistent effect on reports of rape).

85. Ron Wilson, *Geographic Research Suggests Offender Residency Laws May Not Work*, GEOGRAPHY & PUB. SAFETY, May 2009, at 11, 11, available at <http://www.ojp.usdoj.gov/nij/maps/gps-bulletin-v2i1.pdf>.

86. Lester, *supra* note 11, at 359–60; see also Richard Tewksbury, *Exile at Home: The Unintended Collateral Consequences of Sex Offender Residency Restrictions*, 42 HARV. C.R.-C.L. L. REV. 531, 534 (2007) (“One of the most serious and far-reaching collateral consequences associated with sex offender registration is the difficulty [registered sex offenders] experience in locating and maintaining safe, affordable, and legal housing.”)

87. See Lester, *supra* note 11, at 359–62 (reviewing problems created and exacerbated by residency restrictions on sex offenders, their families, communities, and police resources); Elizabeth Ehrhardt Mustaine & Richard Tewksbury, *Residential Relegation of Registered Sex Offenders*, 36 AM. J. CRIM. JUST. 44, 54–55 (2011) (confirming that “registered sex offenders are relegated to socially disorganized, high crime communities” and that, consequently, residential regulations have been “concentrating known sex offenders in neighborhoods with known child sexual assault victims”); Tewksbury, *supra* note 86, at 532–34 (citing empirical studies concerning repercussions of registration and residency restrictions—including loss of friendship, employment, and harassment—on various groups of offenders).

88. 2 BARBARA K. SCHWARTZ & HENRY R. CELLINI, THE SEX OFFENDER: NEW INSIGHTS, TREATMENT INNOVATIONS AND LEGAL DEVELOPMENTS xvii (1997).

There are also ramifications to victims of sex crimes. Stigmatization of offenders and families may lead to reduced rates of reporting of sexual victimization, particularly when the perpetrators are non-strangers.⁸⁹ Regarding intrafamilial sex offending, the new laws may further discourage family victims from reporting or cooperating in criminal prosecutions because of the laws' impact on family members' freedom and the potential for public humiliation.⁹⁰ Though many states protect the secrecy of the victims' information in sex crimes, the publication of their offenders' information may permit others to extrapolate as to the victims' identities, particularly when the victims are family members.⁹¹

The sheer cost of these policies, in terms of cash outlays and governmental resources, is daunting. An informal survey recently indicated that the average cost of civil commitment of SVPs across the states was almost \$100,000 per person per year.⁹² State officials also are complaining about the soaring costs of monitoring sex offenders for compliance with registration and residency laws.⁹³ The estimated cost of implementing the national registry, for example, is \$1.5 billion.⁹⁴

3. Rethinking the Problem

Notably, this Article does not argue that sex offender laws are themselves inherently bad policy. The protection of the public from the horrendous damage that sexual violence causes is certainly a laudable goal of the government. And the strong and official condemnation of sexual violence voiced by the policies is commendable.⁹⁵ Employing bad science in future dangerousness assessments, however, does not advance any of these interests.

The Center for Sex Offender Management, a project of the U.S. Department of Justice, warns that effective management strategies for SVPs should not be politically based on reactionary public fears of the sex offender population.⁹⁶ To be fair, risk assessment is inherently difficult, and human behavior is largely unpredictable. Yet, because of the importance of invoking restrictive laws based on judgments of risk, criminal justice officials and the public should be realistic about the situation.

89. Christopher Uggen & Heather R. Hlavka, *Does Stigmatizing Sex Offenders Drive Down Reporting Rates? Perverse Effects and Unintended Consequences*, 35 N. KY. L. REV. 347, 363 (2008).

90. See Vitiello, *supra* note 17, at 685 (discussing how SVP laws may actually reduce willingness of family members to report their sexual-offender relatives).

91. See Amy Baron-Evans, *Still Time to Rethink the Misguided Approach of the Sex Offender Registration and Notification Act*, 20 FED. SENT'G REP. 357, 358 (2008) (arguing that offender identities should only be disclosed when disclosure will not implicitly identify victim).

92. GOOKIN, *supra* note 39, at 1.

93. Dan Gunderson, *Corrections Officials Critical of Expanded Sex Offender Monitoring*, MINN. PUBLIC RADIO (Feb. 22, 2006), <http://minnesota.publicradio.org/display/web/2006/02/16/gpstracking/> (reporting that Minnesota Department of Corrections' officials estimate that monitoring sex offenders "costs \$20 per day for each offender").

94. CONG. BUDGET OFFICE, COST ESTIMATE: H.R. 4472 CHILDREN'S SAFETY AND VIOLENT CRIME REDUCTION ACT OF 2005, at 2 (2006), available at <http://www.cbo.gov/ftpdocs/70xx/doc7061/hr4472.pdf>.

95. Cf. JANUS, *supra* note 3, at 146-47 (arguing that while SVP laws serve expressive message rejecting sexual violence, by highlighting most heinous sexual assaults by strangers they downplay more common types of sexual violence).

96. CSOM ON RECIDIVISM, *supra* note 61, at 4.

Proponents need to acknowledge that false positives and false negatives will commonly occur, and determine if continuing with future dangerousness assessments is justified legally, socially, and monetarily.⁹⁷ If the answer is yes, then at a minimum, we should demand the most empirically sound risk assessment procedures, learn to better interpret the results offered, and fully understand their advantages as well as their limitations.

III. EXPERT EVIDENCE AND SEX OFFENDERS

Specialized SVP laws remain an important focus in the justice system's response to sex offenders post-release. Because the preventive laws are beyond the traditional application of criminal laws, and, indeed, are directed toward potential future crimes, the application to individual offenders has been substantially based upon assessments of future dangerousness. Courts commonly call on mental health professionals to provide expert testimony to support these risk assessments. In turn, experts rely heavily on actuarial tools developed in an effort to standardize the assessment based on empirical principles. To make their assessments, experts may, but do not always, have access to a variety of data, from mental health records, criminal records (e.g., police reports, arrest reports, and probation/parole results), treatment reports, reports from other evaluators, actuarial assessments, diagnostic evaluations,⁹⁸ and interviews with the defendants. This Part begins with a brief overview of evidentiary law as applied to expert witnesses.

A. *Expert Evidence Law*

The introduction of expert witness testimony in criminal proceedings is not a new idea in the United States, but it is growing in its influence with the advent of forensics such as fingerprints and DNA. Studies show that a judge's admission of expert evidence is quite influential upon jurors' acceptance of the expert testimony as scientifically valid.⁹⁹ A recent trend is the growing use of experts to provide evidence based on knowledge gained through use of the "softer" social sciences.¹⁰⁰ Some of the

97. Some commentators have weighed the empirical flaws in actuarial predictions against the ability of expert testimony to assist judges in making future-dangerousness assessments and have concluded that the public health need to protect future victims justifies their use. Post, *supra* note 12, at 244–45.

98. Rebecca L. Jackson & Derek T. Hess, *Evaluation for Civil Commitment of Sex Offenders: A Survey of Experts*, 19 *SEXUAL ABUSE* 425, 431 (2007); see also *People v. Lopez*, No. H029248, 2006 Cal. App. Unpub. LEXIS 11573, at *5–6 (Cal. Ct. App. Dec. 21, 2006) (noting that expert at trial relied upon mental health, criminal, and hospital records, as well as reports by other evaluators and results of state-authorized instruments for predicting sex offenses). As a general rule, the evidentiary ban on hearsay evidence is vitiated when it comes to experts and the information they rely upon, at least to the extent that experts in their particular field reasonable rely on that type of evidence. *E.g.*, FED. R. EVID. 703.

99. N. J. Schweitzer & Michael J. Saks, *The Gatekeeper Effect: The Impact of Judges' Admissibility Decisions on the Persuasiveness of Expert Testimony*, 15 *PSYCHOL. PUB. POL'Y & L.* 1, 13 (2009) (describing result of authors' empirical study about influence of judge's decision to admit expert testimony on potential jurors' acceptance of such testimony and concluding that if *Daubert's* intent was to limit junk science, judges need to strongly embrace their gatekeeping role since study participants were more likely to value expert testimony if judge admitted it because they believed judge independently acted and competently verified quality of evidence provided by expert).

100. Social sciences are those that use scientific methods in the study of humans as social creatures, such as sociology, anthropology, psychology, criminology, political science, among others. Sarah H. Ramsey &

common areas in which expert evidence draws upon social science in the criminal context involve insanity, battered women syndrome, rape trauma, eyewitness identification, and future dangerousness.¹⁰¹ While much of the social science expert evidence assists the jury in assessing the facts as they previously happened, the future dangerousness issue uniquely requires the expert to provide the social context for determining *future* facts.¹⁰² While the model for reliability in the traditional hard sciences is based on experimentation, replication, and validation, the model that social scientists use replaces the experimentation component with observation.¹⁰³ This is true largely because maintaining a truly randomized, controlled experiment in the atmosphere of a laboratory-like setting when it involves people's social lives is often not practical since it is virtually impossible to control all potentially relevant variables, and in any event, may be unethical—particularly in the context of studying the occurrence of violence and sexual harm.¹⁰⁴

Before focusing on the expert rules as applied to future dangerousness testimony, a summary review of expert evidence law is appropriate. As a general matter, state courts follow either the *Frye* test for expert testimony or the more recent *Daubert*-led standards, though a few states maintain variations of these.¹⁰⁵ The *Frye* test, adopted by a federal appellate court in 1923, mandates that expert testimony involving new scientific evidence have proof of reliability. Reliability is shown if it has been generally accepted in the relevant scientific community.¹⁰⁶ This is known as an exercise in “counting heads.”¹⁰⁷

In 1975, Congress enacted the influential Federal Rules of Evidence, including Rule 702, which provides that expert testimony involving scientific, technical, or other specialized testimony is admissible in federal court if it “will assist the trier of fact to understand the evidence.”¹⁰⁸ Later, the U.S. Supreme Court addressed the issue of whether Rule 702 incorporated the *Frye* standard. In *Daubert v. Merrell Dow*

Robert F. Kelly, *Social Science Knowledge in Family Law Cases: Judicial Gate-Keeping in the Daubert Era*, 59 U. MIAMI L. REV. 1, 6 (2004).

101. See Laurens Walker & John Monahan, *Social Frameworks: A New Use of Social Science in Law*, 73 VA. L. REV. 559, 582–83 (1987) (offering that recent uses of social science in law is uniquely about permitting experts to use theory of human behavior to provide social framework to assist trier of fact in understanding relevant issues while disabusing them of any myths they may have).

102. See JOHN MONAHAN & LAURENS WALKER, *SOCIAL SCIENCE IN LAW: CASES AND MATERIALS* 398 (6th ed. 2006) (referring to social science as providing empirical context for estimations of likelihood of future criminality).

103. Mark S. Brodin, *Behavioral Science Evidence in the Age of Daubert: Reflections of a Skeptic*, 73 U. CIN. L. REV. 867, 868–69 (2005).

104. Andrew E. Taslitz, *Myself Alone: Individualizing Justice Through Psychological Character Evidence*, 52 MD. L. REV. 1, 75 (1993).

105. Alice B. Lustre, Annotation, *Post-Daubert Standards for Admissibility of Scientific and Other Expert Evidence in State Courts*, 90 A.L.R.5th 453, § 2, at 481 (2001).

106. *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923). The *Frye* expert offered evidence in the form of results from an early version of a lie detector test to show the defendant's innocence. *Id.* at 1013.

107. Andrew R. Stolfi, *Why Illinois Should Abandon Frye's General Acceptance Standard for the Admission of Novel Scientific Evidence*, 78 CHI.-KENT L. REV. 861, 876 n.110 (2003) (quoting *Harper v. State*, 292 S.E.2d 389, 395 (Ga. 1982)).

108. Act of Jan. 2, 1975, Pub. L. No. 93-595, art. VII, Rule 702, 88 Stat. 1926, 1937.

Pharmaceuticals, Inc.,¹⁰⁹ the Court ruled that, with Congress' "liberal thrust" for permitting opinion testimony, the Rule displaced *Frye* and its strict focus on the general acceptance test.¹¹⁰ Instead, the Court invoked the broader perspective that scientific evidence is admissible if it is valid and reliable, with general acceptance not the entire substitute for these criteria.¹¹¹ As the Court noted, validity asks "does the principle support what it purports to show?" while reliability asks "does application of the principle produce consistent results?"¹¹² Invoking the metaphor of the gatekeeper, the Court concluded that judges themselves should evaluate the scientific reliability and validity of the proposed evidence.¹¹³ The Court provided some "general observations" to guide judges in determining the validity and reliability of the offered expert evidence to the case at hand: testability; peer review and publication; methodological standards, including the error rate; and general acceptance.¹¹⁴ Hence, general acceptance remained, but only as one of several potentially applicable criteria. In a strong dissent, Chief Justice Rehnquist lamented that the majority's approach would cause judges to become "amateur scientists."¹¹⁵

Despite *Daubert* and its progeny, the Supreme Court left many other questions unanswered. For example, it is unclear whether the presence of any one of the general guideline criteria, even general acceptance, is sufficient if none of the other criteria are met.¹¹⁶ Since the Court made clear that its "general observations" were not a dispositive list of factors to consider,¹¹⁷ judges also struggle with what other factors may be relevant. Judges must also determine the degree of reliability required.¹¹⁸ In the case of SVP laws we might also add a query about whether the standard for reliability should differ depending on the type of deprivation involved, such as the greater infringement on liberty inherent in civil commitment statutes or the arguably less onerous burden imposed by registration or residency requirements.

As between *Daubert* and *Frye*, some commentators considered the new conceptualization from *Daubert* as potentially more generous than the *Frye* standard, thus substantively minimizing the role of the general acceptability standard in federal court.¹¹⁹ On the other hand, *Daubert* can also be more limiting. For instance, there may be general acceptance in the field of astrology about the methods and tools with which

109. 509 U.S. 579 (1993).

110. *Daubert*, 509 U.S. at 587–89 (quoting *Beech Aircraft Corp. v. Rainey*, 488 U.S. 153, 169 (1988)).

111. *Id.* Subsequently, in *Kumho Tire Co. v. Carmichael*, the Court clarified that the same *Daubert*-led standards applied to all expert testimony, including that based on skill or experience, not just of the scientific variety. 526 U.S. 137, 151 (1999).

112. *Daubert*, 509 U.S. at 590 n.9.

113. *Id.* at 597.

114. *Id.* at 593–94.

115. *Id.* at 601 (Rehnquist, C.J., dissenting).

116. Brodin, *supra* note 103, at 874–75.

117. *Daubert*, 509 U.S. at 592–94.

118. Janus & Prentky, *supra* note 40, at 1462.

119. See James Aaron George, Note, *Offender Profiling and Expert Testimony: Scientifically Valid or Glorified Results?*, 61 VAND. L. REV. 221, 233–35 (2008) (discussing how *Daubert* is broader than *Frye*, but but that courts have applied it strictly).

to predict future events based on planetary movement, but a *Daubert*-led court likely would exclude the evidence as specious (i.e., unreliable), and thus inadmissible.¹²⁰

Not all state courts follow *Daubert*, as some continue to follow *Frye* or some equivalent.¹²¹ Further, many of the *Frye* jurisdictions still distinguish between scientific evidence (to which *Frye* applies) and nonscientific evidence (to which *Frye* does not).¹²² In addition, for those states strictly following *Frye*, the reliability (via general acceptance) question is entertained only if an expert is offered to testify about a new scientific methodology.¹²³ In states retaining *Frye*, then, expert evidence that is either not new or not scientific is not subject to the *Frye* admissibility standard, though other general evidentiary rules regarding relevance and prejudice apply. Where the science is novel, a few courts, while purportedly holding onto the *Frye* test, require some additional measure of reliability.¹²⁴ For example, a New York court recognized that “general acceptance does not necessarily mean that a majority of the scientists involved subscribe to the conclusion. Rather it means that those espousing the theory or opinion have followed generally accepted scientific principles and methodology in evaluating clinical data to reach their conclusions.”¹²⁵

As a result of this body of law on expert evidence, two observations can be made, perhaps from a cynical perspective, of the goal-oriented actions that experts and judges appear to use to balance objectivity with efficiency. First, a consequence of relying upon prior rulings regarding general acceptance is that, once a novel science is admitted by judges, it maintains a veil of “good science” without meaningful review, despite advances in research and/or legitimate criticism. Second, experts “in a novel area sharing a common goal may develop a technique that furthers their professional

120. *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 151 (1999) (“[T]he presence of *Daubert*’s general acceptance factor [does not] help show that an expert’s testimony is reliable where the discipline itself lacks reliability, as, for example, do theories grounded in so-called generally accepted principles of astrology or necromancy.”); see also Erica Beecher-Monas & Edgar Garcia-Rill, *Danger at the Edge of Chaos: Predicting Violent Behavior in a Post-Daubert World*, 24 CARDOZO L. REV. 1845, 1854 (2003) (noting that inferences must be derived from scientific method to qualify as scientific knowledge, and thus, astrology cannot be scientific knowledge).

121. Lustre, *supra* note 105, at 481 (“Twenty-five states have affirmatively adopted the *Daubert* or similar test for use in their courts, or had previously abandoned *Frye* and had developed a similar test; fifteen states and the District of Columbia adhere to *Frye*; six states have not rejected *Frye* in toto but apply the *Daubert* factors; and four states developed their own tests.” (internal citations omitted) (capitalization omitted)).

122. *Id.* at 520–36.

123. See, e.g., *Donaldson v. Cent. Ill. Pub. Serv. Co.*, 767 N.E.2d 314, 324 (Ill. 2002) (“The trial judge applies the *Frye* test only if the scientific principle, technique or test offered by the expert to support his or her conclusion is ‘new’ or ‘novel.’”).

124. See, e.g., *People v. Kelly*, 549 P.2d 1240, 1248–50 (Cal. 1976) (requiring testimony of general acceptance from more than one source, and preferably from neutral party).

125. *Ratner v. McNeil-PPC*, 898 N.Y.S.2d 772, 773 (N.Y. App. Div. 2010) (citing *Zito v. Zabarsky*, 12 N.Y.S.2d 535 (N.Y. App. Div. 2006)). But see *Clemons v. State*, 896 A.2d 1059, 1078 (Md. Ct. App. 2006) (noting that controversy in relevant scientific community about reliability and validity may lead to exclusion of expert testimony).

aims and they may ‘generally accept’ it regardless of its scientific validity, sometimes despite strong scientific denial of its underlying premises.”¹²⁶

This Article suggests that these problems may help explain the state of admissibility decisions on expert evidence involving dangerousness assessment methodologies using actuarial tools in SVP cases. The result is tautological: if generally accepted, then it is reliable; actuarial risk assessments are generally accepted; therefore, actuarial risk assessments are reliable (and truthful).

B. *Experts and Future Dangerousness Assessment*

The legal and constitutional bases for permitting experts to testify as to an offender’s likelihood of recidivism derive from the U.S. Supreme Court in the capital punishment sentencing domain. In *Barefoot v. Estelle*,¹²⁷ the Supreme Court confronted a Texas death penalty statute that permitted a jury to sentence a capital defendant to death only after answering two questions affirmatively.¹²⁸ As pertinent here, one question to be answered, with a dichotomous yes/no response was whether there was “a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society.”¹²⁹ The issue the Court addressed in the case was the constitutionality (under the Eighth and Fourteenth Amendments) of mental health expert witnesses testifying before the jury on the question of the defendant’s future dangerousness.¹³⁰ The defendant in *Barefoot* argued that doing so would violate the Cruel and Unusual Punishment Clause applying to capital cases by contributing to arbitrary and capricious decisions in death sentencing.¹³¹

In support of the defendant’s position that such testimony could not be reliable, the American Psychiatric Association (APA) submitted an amicus brief in which the professional organization maintained that psychiatric predictions on recidivism were unreliable and that, in its estimate, two out of three predictions by psychiatrists of long-term future dangerousness were erroneous.¹³² Despite acknowledging the APA’s

126. ANDRE A. MOENSSENS ET AL., *SCIENTIFIC EVIDENCE IN CIVIL AND CRIMINAL CASES* 12 (5th ed. 2007).

127. 463 U.S. 880 (1983).

128. *Barefoot*, 463 U.S. at 883 n.1.

129. *Id.* at 884.

130. *Id.* at 884–85.

131. *Id.*

132. *Id.* at 920 (Blackmun, J., joined by Brennan and Marshall, JJ., dissenting). The psychiatrist who testified in the case was Dr. Grigson, nicknamed “Dr. Death,” who became infamous for the frequency of his testifying for the prosecution in Texas death penalty sentencing trials (testifying in over 150 death penalty cases before retiring), and for the strength and eloquence of his testimony (usually about the unlikelihood of murderers being rehabilitated). Hugh Aynesworth, *Texas Dr. Death Retires After 167 Capital Case Trials: Felt Murderers Would Kill Again*, WASH. TIMES, Dec. 21, 2003, at A02. For example, in the instant case, Dr. Grigson testified that there was a “one hundred percent and absolute” risk that Barefoot would commit violent crimes in the future. *Barefoot*, 463 U.S. at 919 (Blackmun, J., joined by Brennan and Marshall, JJ., dissenting). (quoting trial transcript). Despite not having ever examined Barefoot personally, Dr. Grigson declared—based on hypothetical information matching Barefoot’s case facts—that such a person would be a highly dangerous sociopath who could not be cured. *Id.* at 917–19. The jury sentenced Barefoot to death. *Id.* at 919. Dr. Grigson was later expelled by the APA for diagnosing mental illness without personally assessing the individual and for testifying that he could predict with 100% accuracy the likelihood of violent recidivism. Laura Beil,

position, the six-Justice majority nevertheless ruled against the defendant.¹³³ The majority explained that it was not convinced that expert testimony predicting future violence was “entirely unreliable,” and in any event, any “shortcomings” could be effectively minimized during the adversarial process.¹³⁴ According to the majority, even the APA did not assert that psychiatrists were always wrong and, even if many psychiatrists disagreed with the reliability of such predictions, others were willing to testify and to give their professional opinions about the defendant’s future risk of violence.¹³⁵ The three-Justice dissent provided a strong rejoinder, stating that “[o]ne can only wonder how juries are to separate valid from invalid expert opinions when the ‘experts’ themselves are so obviously unable to do so.”¹³⁶ Further, “when the Court knows full well that psychiatrists’ predictions of dangerousness are specious, there can be no excuse for imposing on the defendant, on pain of his life, the heavy burden of convincing a jury of laymen of the fraud.”¹³⁷

Further, the majority opinion was silent on methodology, and therefore did not expressly consider how the experts at trial derived their opinions that the defendant, Barefoot, was likely to pose a future danger. The Court merely noted that one of the state’s experts testified that he believed he could competently make a risk prediction about an individual defendant “if given enough background information.”¹³⁸

In the entirety of its opinion, the *Barefoot* majority did not substantively engage the question of evidentiary reliability standards as applied to violence risk assessment. Instead, the Court appeared more concerned with two collateral issues. First, the majority worried that if it categorically barred expert evidence on future dangerousness, such a ruling may present a slippery slope by also undermining the use of experts in other areas of law requiring risk assessments of danger. The Court specifically cited the use of future dangerousness predictions in bail, parole, and civil commitment proceedings.¹³⁹ Second, the Court noted that since the law requires juries to make this

Groups Expel Psychiatrist Known for Murder Cases: Witness Nicknamed ‘Dr. Death’ Says License to Practice Won’t Be Affected by Ethics Allegations, DALL. MORNING NEWS, July 26, 1995, at 21A.

133. *Barefoot*, 463 U.S. at 906.

134. *Id.* at 899. There is some evidence that the majority may be right here. In a mock-juror design, researchers concluded that jurors were much more likely to question an expert’s testimony on future dangerousness after it was subject to cross-examination or a competing expert’s testimony. Daniel A. Krauss & Bruce D. Sales, *The Effects of Clinical and Scientific Expert Testimony on Juror Decision Making in Capital Sentencing*, 7 PSYCHOL. PUB. POL’Y & L. 267, 302 (2001). Despite the cross-examination or competing expert’s testimony, however, the researchers still found that the mock jurors’ assessment of the purported defendant’s future dangerousness was significantly influenced by the initial expert’s testimony. *Id.* at 299–300.

135. *Barefoot*, 463 U.S. at 899–901.

136. *Id.* at 929.

137. *Id.* at 935–36; see also Erica Beecher-Monas, *The Epistemology of Prediction: Future Dangerousness Testimony and Intellectual Due Process*, 60 WASH. & LEE L. REV. 353, 361, 371 (2003) (discussing tension between *Barefoot* and *Daubert* and arguing that *Daubert*’s evidentiary standards should apply to expert testimony on future dangerousness); Mitzi Dorland & Daniel Krauss, *The Danger of Dangerousness in Capital Sentencing: Exacerbating the Problem of Arbitrary and Capricious Decision-Making*, 29 LAW & PSYCHOL. REV. 63, 102–03 (2005) (maintaining that expert opinions on future dangerousness in capital cases are so unreliable that they lead to arbitrary and capricious decisions).

138. *Barefoot*, 463 U.S. at 899 n.7.

139. *Id.* at 897–98.

type of factual assessment, jurors should at least get some external help.¹⁴⁰ In sum, the Court refused to constitutionally exclude “an entire category of expert testimony” about future dangerousness.¹⁴¹

Hence, with the Supreme Court’s approval of expert predictions of future violence in death penalty cases,¹⁴² and with the majority’s reference to expert assessments of the risk of violence in civil commitments, it seems reasonable to extrapolate *Barefoot*’s general conclusion to future dangerousness assessments of sex offenders. Some legal academics interpret the *Barefoot* decision as relaxing the evidentiary standard later adopted in *Daubert*, specifically for expert testimony on future dangerousness.¹⁴³ Still, the Court in *Barefoot* did not expressly do so as it did not confront the issue.

This Article maintains that courts should reengage in critically examining the reliability and validity of expert testimony in future dangerousness contexts. This should apply not only to SVP civil commitment proceedings but also to the arguably less restrictive regimes where individual assessments of future dangerousness are relevant to the application of sex offender registry requirements and residency restrictions. The legal conclusion from *Barefoot*—that the Constitution does not require expert predictions on future dangerousness to be categorically excluded—does not end the analysis. The Court certainly cannot have meant that just anyone could qualify as an expert witness or that any opinion the expert wishes to give should be admitted. In support of this contention, the Court, since *Barefoot*, has become a bit more wary about mental health expert testimony. In *Ford v. Wainwright*,¹⁴⁴ a case involving the insanity determination of death row inmates, the Court, citing *Barefoot*, warned against simply relying upon an expert’s testimony.¹⁴⁵ The *Ford* majority reasoned that focused questioning of the expert would serve the truth-finding function

by bringing to light the bases for each expert’s beliefs, the precise factors underlying those beliefs, any history of error or caprice of the examiner, any personal bias with respect to the issue of capital punishment, the expert’s degree of certainty about his or her own conclusions, and the precise meaning of ambiguous words used in the report. Without some questioning of the experts concerning their technical conclusions, a factfinder simply cannot be expected to evaluate the various opinions, particularly when they are themselves inconsistent.¹⁴⁶

140. *Id.* at 897.

141. *Id.* at 899.

142. The categorical approval of expert witness testimony on future dangerousness in capital cases was cited with approval in *Payne v. Tennessee*, 501 U.S. 808, 823 (1991) (approving prosecution’s use of victim impact statement in sentencing phase of capital trial). *Cf.* *Coble v. State*, 330 S.W.3d 253, 270–87 (Tex. Crim. App. 2010) (holding that while trial court in capital case erred by admitting insufficiently reliable expert testimony on future dangerousness, admission amounted to harmless error).

143. Beecher-Monas & Garcia-Rill, *supra* note 25, at 311–12.

144. 477 U.S. 399 (1986).

145. *Ford*, 477 U.S. at 415 (citing *Barefoot*, 463 U.S. at 899).

146. *Id.*

In dicta, the *Ford* Court also questioned the reliability of the experts' opinions in the underlying case, noting the "cursory nature of the underlying [joint] psychiatric examination" that appeared "dubious" at best.¹⁴⁷

In capital cases since *Barefoot*, the Court has expressed more concern about the reliability and truth-assisting nature of expert testimony on future dangerousness in capital cases. Though not overruling *Barefoot* on this point, the Court clarified the remedy to help ameliorate the deficiencies in expert predictions of future crime. For example, if the prosecution offers a psychiatrist to testify in the sentencing phase about future dangerousness, due process requires that the state provide the defendant with his own psychiatrist to rebut the prosecution's expert.¹⁴⁸ The significant consequence of an erroneous decision on future dangerousness based on just the prosecution's expert, the Court found, was unacceptable.¹⁴⁹

While the *Barefoot* case did not expressly address actuarial-based assessment of future dangerousness for purposes of death penalty sentencing, it suggests that the use of mental health experts in risk assessment will continue. This Article is thus interested in bringing *Frye*, *Daubert*, and the broader concepts of reliability and validity back into at least one aspect of expert testimony on future dangerousness: the use of actuarial assessment tools in SVP determinations.

IV. ACTUARIAL TESTING OF FUTURE DANGEROUSNESS

At the time of *Barefoot*, experts substantially based their predictions of future dangerousness on their clinical judgments,¹⁵⁰ while critics complained that clinical assessments were inherently unreliable and subject to bias.¹⁵¹ Practitioners thereby sought more empirically-based tools that could offer more reliable guides.¹⁵² The development of actuarial risk tools ensued.

Fundamentally, actuarial risk tools are about deriving statistics from groups. For example, automobile insurance companies assign policy rates to individuals based on predictive statistics derived from historical, group-based claims data. For car insurance, the common relevant factors include age, education, vehicle model, and driving history.¹⁵³ The general idea for actuarial ratings for any risk at issue is to identify those factors that are correlative to the potential occurrence of the future event at issue, and to effectively assign appropriate weights to each factor based on the observation that some factors have greater correlative abilities than others relating to the particular result. The

147. *Id.* at 415 n.3.

148. *Ake v. Oklahoma*, 470 U.S. 68, 84–85 (1985).

149. *Id.* at 84.

150. See Beecher-Monas & Garcia-Rill, *supra* note 120, at 1847 (noting that clinical testimony regarding future dangerousness is pervasive, and that "courts persist in circumventing any inquiry into the scientific validity of expert future dangerousness predictions").

151. *E.g.*, Prentky et al., *supra* note 6, at 371–72.

152. See, *e.g.*, John Monahan, *A Jurisprudence of Risk Assessment: Forecasting Harm Among Prisoners, Predators, and Patients*, 92 VA. L. REV. 391, 408–09 (2006) (claiming that, although actuarial risk assessment has been known to be superior to clinical risk assessment for half a century, only recently have tools for predicting violent behavior been developed and implemented).

153. Fennel, *supra* note 58, at 52.

theory is that a better model of prediction should be based not on any single risk factor, but an accumulation of relevant risk factors.¹⁵⁴ The developers of actuarial instruments, therefore, use existing data in an empirical way to create rules to combine the most relevant factors, provide the applicable weights, and create a final mechanistic score.¹⁵⁵ The assessor then compares the score against the experience tables which yield probabilities of the result observed from the reference group data. To support the empirical validity of the instruments, the scales are cross-validated by retesting with other samples.

Understanding the group-based nature of actuarial assessment tools is crucial. When determining the relative risk for an individual, the characteristics of the individual common to those in the population upon which the actuarial model is based are compared and ranked as to the outcome of interest.¹⁵⁶ To return to the automobile insurance example, the insurer's agent would input a prospective customer's data into the actuarial model to obtain a relative risk level (and corresponding premium) based on the experiential claims data from those in the historical sample with similar characteristics.

A. *The Sexual Recidivism Actuarial Tests*

Two of the most commonly used actuarial prediction tools in SVP determinations are the Rapid Risk Assessment for Sex Offence Recidivism (RRASOR) and STATIC-99.¹⁵⁷ The RRASOR is the more stream-lined of the two, assigning points on just four static factors: number of prior sex offenses (from 0 points = no convictions or charges to 3 points = 4 or more convictions or 6 or more charges), age at assessment (meant to be at release) (0 = more than 25 years; 1 = less than 25 years), victim gender (0 = only females; 1 = any male), and relationship to victim (0 = only related; 1 = any non-related).¹⁵⁸ The items are scored, and the sum of the scores is associated with a certain recidivism rate over a 5-year and 10-year period based on group statistics observed in developmental samples.¹⁵⁹ For example, a higher score will result for a subject who has a greater number of prior convictions or charges of sex offenses, age less than 25 years, any male victims, and at least one extra-familial victim.

The developer of RRASOR created the instrument based on preexisting studies of sex offender reoffending by identifying the factors that tended to be correlated with

154. See JOHN MONAHAN ET AL., *RETHINKING RISK ASSESSMENT: THE MACARTHUR STUDY OF MENTAL DISORDER AND VIOLENCE* 142 (2001) (making this conclusion with respect to propensity for violence).

155. See Kevin S. Douglas & Jennifer L. Skeem, *Violence Risk Assessment: Getting Specific About Being Dynamic*, 11 *PSYCHOL. PUB. POL'Y & L.* 347, 352 (2005) (demonstrating how actuarial instruments use mechanistic algorithm to combine heavily weighted static variables relevant to making ultimate determination of risk).

156. Fennel, *supra* note 58, at 52.

157. Fred S. Berlin et al., *The Use of Actuarials at Civil Commitment Hearings to Predict the Likelihood of Future Sexual Violence*, 15 *SEXUAL ABUSE* 377, 378 (2003). See generally *PARALLELING BEHAVIOUR: A CASE FORMULATION APPROACH TO OFFENDER ASSESSMENT AND INTERVENTION* (Michael Daffern et al. eds., 2010).

158. R. KARL HANSON, *THE DEVELOPMENT OF A BRIEF ACTUARIAL RISK SCALE FOR SEXUAL OFFENSE RECIDIVISM* 14 (1997).

159. *Id.* at 14–16.

sexual recidivism.¹⁶⁰ In reducing the factors to four, the developer's goal was to create a "brief, efficient actuarial tool that could be used to assess the risk for sexual offense recidivism."¹⁶¹ Because of the goal of efficiency, the developer acknowledged that RRASOR was not a comprehensive assessment tool.¹⁶²

To develop the experience table, the developer used the sexual recidivism rates observed in seven follow-up studies of released sex offenders in the United States, Canada, and England.¹⁶³ Sexual recidivism was variously defined in the studies as charges, convictions, or readmissions to inpatient psychiatric facilities.¹⁶⁴ The developer then created the final experience table that associates specific RRASOR scores (1–5)¹⁶⁵ with risk estimates for five- and ten-year periods. Instead of using the exact observed rates of sexual recidivism from the samples, the final experience table risk estimates were extrapolated by formula from the observed rates because the samples had varied follow-up periods (2 to 23 years).¹⁶⁶ In the end, the experience table's predictive rates range from a low of 4% sexual recidivism for 0 points at 5 years to a high of 73% for 5 points at 10 years.¹⁶⁷

STATIC-99 was developed by the RRASOR author with another researcher using four samples of male sex offenders, totaling just over 1,000, released from Canadian and English institutions.¹⁶⁸ It remains the most commonly used actuarial tool in the United States for SVP civil commitment hearings because of its resource efficiency.¹⁶⁹ The resulting tool is a combination of variables from two other instruments, including

160. *Id.* at 4.

161. *Id.*

162. *Id.* at 18.

163. *Id.* at 6.

164. *Id.*

165. Since the samples included no scores of 6, the experience table does not include separate risk estimates for it. *Id.* at 13.

166. *Id.* at 8, 10–11.

167. *Id.* at 16.

168. See generally R. Karl Hanson & David Thornton, *Improving Risk Assessments for Sex Offenders: A Comparison of Three Actuarial Scales*, 24 LAW & HUM. BEHAV. 119 (2000). The authors indicated they chose the name STATIC-99 to highlight that the tool was developed in 1999 and includes only static factors. *Id.* at 122. STATIC-2002 is a newer instrument, in which the authors altered some of the original ten factors and added factors involving juvenile sex offending, rate of sexual offending, young unrelated victims, community supervision violations, and years free prior to index crime. AMY PHENIX ET AL., CODING RULES FOR STATIC-2002, at 3–4 (2008), available at <http://www.static99.org/pdfdocs/static2002codingrules.pdf>. However, STATIC-99 remains the most used actuarial tool in sexual predator cases. Leslie Helmus et al., *Reporting Static-99 in Light of New Research on Recidivism Norms*, THE FORUM, Winter 2009, at 38, 38. Indeed, few cases to date provide any reference to STATIC-2002. Those cases that have referenced STATIC-2002 have generally reported STATIC-99 scores as well. *E.g.*, United States v. Hunt, 643 F. Supp. 2d 161, 172–73 (D. Mass. 2009); People v. Sims, 2010 Cal. App. Unpub. LEXIS 5820, at *4 (Ct. App. Cal. July 26, 2010); People v. Zavala, 2010 Cal. App. Unpub. LEXIS 5544, at *4–5 (Ct. App. Cal. July 15, 2010).

169. R. Karl Hanson et al., *Predicting Recidivism Amongst Sexual Offenders: A Multi-Site Study of Static-2002*, 34 LAW & HUM. BEHAV. 198, 198 (2010).

all four factors from RRASOR.¹⁷⁰ The STATIC-99 instrument includes 10 static factors¹⁷¹:

<u>Age at assessment:</u> 0 = 25 years or older 1 = between 18 and 25 years	<u>Number of prior sentencing dates:</u> 0 = 3 or less 1 = 4 or more
<u>Having lived with an age-appropriate intimate partner for at least 2 years:</u> 0 = yes; 1 = no	<u>Any convictions for a non-contact sexual offense:</u> 0 = no; 1 = yes
<u>Any convictions for an Index non-sexual violent offense:</u> 1 = yes; 0 = no	<u>Any nonfamilial victims:</u> 0 = no; 1 = yes
<u>Any convictions for non-sexual violence before the Index (most recent sexual offense) offense:</u> 1 = yes; 0 = no	<u>Any stranger victims:</u> 0 = no; 1 = yes
<u>Number of prior sex offenses:</u> 0 = none 1 = 1–2 charges or 1 conviction 2 = 3–5 charges or 2–3 convictions 3 = ≥ 6 charges or ≥ 4 convictions	<u>Any male victims:</u> 0 = no; 1 = yes

For STATIC-99, total scores range from 0 to 12, arranged within seven risk categories organized into four ordinal risk groups (from 0 = low risk to 6+ = high risk).¹⁷² The experience table provides 5-, 10-, and 15-year sexual recidivism rates for each total score from 0 through 6, with 6–12 points sharing the same experience rates.¹⁷³ Sexual recidivism was operationalized as reconviction in three of the samples and either charges or readmission for one sample, and the final experience rates simply aggregate them. A few examples from the experience table may be helpful. A total score of 3 provides estimates of sexual recidivism of 12% at 5 years, 14% at 10 years, and 19% at 15 years.¹⁷⁴ The greatest scores which were grouped into a 6+ range (6–12 points) yields the highest estimates in the experience table of 39% (5 years), 45% (10

170. ANDREW HARRIS ET AL., *STATIC-99 CODING RULES REVISED 3* (2003), available at http://www.publicsafety.gc.ca/res/cor/rep/_fl/2003-03-stc-cde-eng.pdf; Hanson & Thornton, *supra* note 168, at 121–22.

171. HARRIS ET AL., *supra* note 170, at 67.

172. *Id.*

173. *Id.* at 69. “Experience rate” is a term the developers use for the recidivism statistics observed in the developmental samples.

174. *Id.*

years), and 52% (15 years).¹⁷⁵ Even a score of 0 yields positive estimates of 5%, 11%, and 13% at 5-, 10-, and 15-year intervals, respectively.¹⁷⁶

The STATIC-99 coding rules indicate that data can be derived from some combination of self-reports, formal records, and informal records.¹⁷⁷ The test was designed to be used only on adult males who have been charged with or convicted of a contact sex crime involving a child or a nonconsenting adult.¹⁷⁸ The authors specifically warn that it is not recommended for females, juveniles, individuals with no prior sex crime, possessors of pornography, or individuals who have engaged in certain consensual sex activities that are otherwise considered a crime, such as prostitution, sex in public, and statutory rape.¹⁷⁹

The application of the actuarial predictions to an individual (the defendant) is basically an exercise in inductive logic: those in the development samples who were like the defendant reoffended X percent of time; ergo the risk that the defendant will reoffend is similar to X percent.¹⁸⁰ Among its highest risk group (score of 6 or more) and at the longest follow-up period (15 years), the STATIC-99 indicates a 52% chance of sex reoffending for those in the group with a score of 6+. So, even for this extreme, many (almost half) did not reoffend.¹⁸¹ Further, other studies have failed to replicate the over 50% reoffense rate, even for STATIC-99's high risk group.¹⁸²

Neither RRASOR nor STATIC-99 measures the specific type of sexual reoffense (such as rape or child molestation), or the "severity, imminence, duration, or frequency" of future sexual misconduct.¹⁸³ Nor do they limit their recidivism statistics to predatory sexual violence (which civil commitment laws require without clearly clarifying the term "predatory"), serious sexual violence, or to contact sexual violence (which registration and residency laws would seem to target). Both tests consider static factors, that is, factors that are preexisting characteristics at the time of assessment.

175. *Id.*

176. *Id.* These numbers are from the original STATIC-99 forms. The authors recently (October 2008) updated the tables because of their recognition that recidivism numbers have generally dropped. However, the newer tables have been criticized for not being cross-validated or peer reviewed. See generally Brian R. Abbott, *Applicability of the New Static-99 Experience Tables in Sexually Violent Predator Risk Assessments*, 4 SEXUAL OFFENDER TREATMENT 1 (2009), <http://www.sexual-offender-treatment.org/index.php?id=73&type=123>.

177. See HARRIS ET AL., *supra* note 170, at 4–5 (demonstrating how data can be collected through official reports, collateral sources, or self-reporting depending on nature of data being collected).

178. *Id.* at 5.

179. *Id.*

180. Stephen D. Hart, *Actuarial Risk Assessment: Commentary on Berlin et al.*, 15 SEXUAL ABUSE 383, 385 (2003) (describing this as argument by analogy).

181. Berlin et al., *supra* note 157, at 380–81.

182. See generally Reinhard Eher et al., *Failure of Static-99 and SORAG to Predict Relevant Reoffense Categories in Relevant Sexual Offender Subtypes: A Prospective Study*, 3 SEXUAL OFFENDER TREATMENT 1 (2008), <http://www.sexual-offender-treatment.org/index.php?id=65&type=123>; Grant T. Harris et al., *A Multisite Comparison of Actuarial Risk Instruments for Sex Offenders*, 15 PSYCHOL. ASSESSMENT 413 (2003); Jan Looman, *Comparison of Two Risk Assessment Instruments for Sexual Offenders*, 18 SEXUAL ABUSE 193 (2006).

183. Stephen D. Hart et al., *Precision of Actuarial Risk Assessment Instruments: Evaluating the 'Margins of Error' of Group v. Individual Predictions of Violence*, 190 BRIT. J. PSYCHIATRY (supp. 49) s60, s60 (2007) (discussing flaws with actuarial risk assessment instruments in general).

Dynamic factors, factors that can change over time and alter an individual's risk for reoffense, are highly relevant to an assessment of risk, but are generally ignored.¹⁸⁴ Examples of dynamic risk factors include treatment, impulsiveness, anger, substance use, and interpersonal relationships.¹⁸⁵ The lack of dynamic factors is a common lament about the usefulness of actuarial models since, considering the variability of human behavior in general, risk states ebb and flow over time.¹⁸⁶

In the years since *Barefoot* and *Hendricks*, the use of actuarial tools in risk assessment has increased significantly.¹⁸⁷ Many mental health professionals who work with risk assessments of future violence claim that actuarial risk assessments are better and more objective tools than merely using clinical assessments. Some even argue that actuarial tools represent the best practice in the field.¹⁸⁸ Indeed, there is evidence from sex offender researchers that a vast majority of mental health evaluators testifying in sexual predator civil commitment hearings use one or more actuarial instruments.¹⁸⁹ Several state laws outlining the sexual predator classifications require reliance on a specified actuarial test.¹⁹⁰ Still, even purported best practices should not be admissible as evidence in court if they do not meet the legal standards of validity and reliability for expert evidence.

B. *Empirical Evaluation of Actuarial Evidence of Future Dangerousness*

The ability of lawyers and judges to challenge and evaluate expert testimony on future dangerousness is critically important to the legal legitimacy of SVP litigation. Still, several practical barriers to effective evaluation exist. Humans are simply hard to predict, making assessments of future behavior impractical.¹⁹¹ As a result, the

184. See LEAM A. CRAIG ET AL., *ASSESSING RISK IN SEX OFFENDERS: A PRACTITIONER'S GUIDE* 55 (2008) (indicating that most predictive algorithms use only static factors).

185. See Douglas & Skeem, *supra* note 155, at 357–59 (discussing these dynamic risk factors with respect to risk of future violence).

186. *Id.* at 348; Andrew John Rawson Harris & R. Karl Hanson, *Clinical, Actuarial and Dynamic Risk Assessment of Sexual Offenders: Why Do Things Keep Changing?*, 16 J. SEXUAL AGGRESSION 296, 300–01 (2010); John Matthew Fabian, *To Catch a Predator, and Then Commit Him for Life: Sexual Offender Risk Assessment—Part Two*, CHAMPION, Mar. 2009, at 32, 37.

187. See Jackson & Hess, *supra* note 98, at 434 (noting that 95.1% of responding experts reported using actuarial risk assessments “always or most of the time”).

188. Janus & Prentky, *supra* note 40, at 1497.

189. Daniel C. Murrie et al., *Rater (Dis)Agreement on Risk Assessment Measures in Sexually Violent Predator Proceedings: Evidence of Adversarial Allegiance in Forensic Evaluation?*, 15 PSYCHOL. PUB. POL'Y & L. 19, 21 (2009).

190. CAL. PENAL CODE § 290.04(b) (West 2010) (requiring STATIC-99 for risk assessment of male offenders unless and until designated committee chooses alternative actuarial tool); VA. CODE ANN. § 37.2-903(B)–(C) (2009) (requiring STATIC-99 score of five, but four if offense involved minor under age of thirteen; alternatively permitting clinical assessment if state officials believe no scientific actuarial instrument is available). Two Virginia Supreme Court justices warned that this requirement will “encourage a battle between expert witnesses with regard to whether an inmate received ‘a correctly computed score.’” Miles v. Commonwealth, 645 S.E.2d 924, 925 (Va. 2007) (Kinser and Lemon, J.J., concurring).

191. See Steve Erickson, *The Ethics of SVP Evaluations*, CRIME & CONSEQUENCE, Oct. 7, 2010, available at 2010 WLNR 20072402 (noting argument of many psychologists that “we don't know that much about sexual recidivism with the level of precision necessary to justify civil commitment”); Prentky et al., *supra* note 6, at 371 (noting that future is inherently unknowable).

politically charged atmosphere surrounding the post-release management of sex offenders may lead participants in the process to err on the side of confirming SVP status rather than risk the consequences of not applying SVP restrictions to those who eventually reoffend.¹⁹² Expert witnesses admit feeling pressure in the adversarial process to provide positive assessments of risk without adequately explaining contrary research, and even distorting the limitations of the actuarial tools.¹⁹³ Further, the situation is somewhat unique for the treatment field. The mental health professional who has previously worked directly with the individual to be assessed (e.g., a therapist or counselor), and who thereby has greater insight into the individual's likelihood of relapse outside of the limited static factors covered by the tests, is exactly the one who cannot give predictions in criminal proceedings. Ethics guidelines generally prohibit a professional from being both the treating professional and testifying expert because the dual roles often conflict.¹⁹⁴ Finally, the lack of clarity in the complicated field of actuarial risk assessment for sex offenders undermines efforts to logically assess it as a science. With these challenges in mind, the goal of this section is to provide an empirical assessment of the actuarial tools in a way that is accessible to legal professionals working in SVP litigation.

1. Testing and the Scientific Method¹⁹⁵

As applicable to sex offenders in the United States, the creators of the actuarial tools took some liberties with pure scientific method, in particular, the scientific principle that developmental samples underlying actuarial tools designed to be normative should be representative of the larger population for which the tools are intended.¹⁹⁶ This could include weighted random sampling to match the population at issue on relevant variables, such as age, gender, geographic location, treatment, type of sexual offense, etc. The RRASOR and STATIC-99 developmental samples were derived from a limited number of small, nonrandom samples from mostly Canadian and English institutions, with one U.S. sample included in RRASOR.¹⁹⁷ Moreover, the

192. *Id.* at 360.

193. *Id.*

194. AM. ACAD. OF PSYCHIATRY & THE LAW, ETHICS GUIDELINES FOR THE PRACTICE OF FORENSIC PSYCHIATRY 3 (2005); AM. PSYCHIATRIC ASS'N, ETHICS PRIMER 69 (2001).

195. This Article's focus upon a social science does not mean that there are not similar issues with experts in the hard sciences. For instance, in the 2008 Annual Review of Law and Social Science, experts who have worked with issues on science in the law for years summarized:

It has been suggested that over the decades forensic examiners have been pressured to make statements as extreme as they can get—100% certainty, zero error rates, identification to the exclusion of all others in the world—not because such statements grow out of their science, but because they serve the needs of those who use their work.

Michael J. Saks & David L. Faigman, *Failed Forensics: How Forensic Science Lost Its Way and How It Might Yet Find It*, 4 ANN. REV. L. & SOC. SCI. 149, 159 (2008). Further, these experts observe that judges tend to be inclined to permit the evidence despite more current knowledge raising doubts about its validity and reliability because of tradition and deference to prior appellate decisions. *Id.* at 153.

196. Abbott, *supra* note 176, at 10.

197. HANSON, *supra* note 158, at 7–10; R. Karl Hanson & David Thornton, *STATIC-99: Improving Actuarial Risk Assessments for Sex Offenders* 6–8 (1999), <http://www.courtdiagnostic.com/Static%2099-02.pdf>.

majority of the sex offenders in the samples were released from maximum security prisons or mental health institutions, and thus may represent higher risk groups than typical sex offenders.¹⁹⁸ While intending the tools to be applicable on an international scale, there is no sign the developers made any attempt to conduct truly representative sampling to satisfy scientific principles for a more global application.¹⁹⁹ Another issue is the fact that the developmental samples included inconsistent definitions of the outcome variable of recidivism, including charges, readmissions, and/or reconvictions, and used widely varying time frames for follow-up.

The instruments also obscure common scientific standards for determining the reliability of the scoring system by failing to provide error rates,²⁰⁰ as the *Daubert* court mentioned.²⁰¹ For the purpose of risk prediction research, the error rate is normally reported as a 95% confidence interval.²⁰² Other researchers recently attempted to fill this gap by extrapolating confidence intervals from the STATIC-99 data.²⁰³ They found that at the highest risk of the original STATIC-99 experience table, of 52% for a 15-year period for the 6+ score, the group confidence interval was 43–60%, while the individual confidence interval was 6–95%.²⁰⁴ Confidence interval information, a scientifically useful statistic in social science studies, could be important where a factfinder may lean toward the lower or upper bound of the confidence interval in determining whether the defendant meets the relevant legal threshold of future dangerousness.²⁰⁵

2. Validity

Measurement of the predictive validity of actuarial tools is useful to appreciating their abilities. “Evidence demonstrating the predictive validity of any instrument or assessment procedure is of paramount importance when the goal of the clinician is to draw inferences or conclusions about an individual’s likely conduct in the future.”²⁰⁶ Two common statistical measures of the validity of actuarial tests include the

198. See HANSON, *supra* note 158, at 7–10 (indicating over half of sex offenders in the samples for RRASOR were released from maximum security facilities, psychiatric inpatient treatment, or were referred for treatment at maximum-security mental-health facilities); Hanson & Thornton, *supra* note 197, at 6–8 (noting that a majority of STATIC-99 developmental samples were derived from populations released from maximum security and psychiatric treatment facilities in England and Canada).

199. *Id.* at 10.

200. *Id.* at 13.

201. *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 594 (1993) (stating that courts should consider “known or potential rate[s] of error”).

202. Abbott, *supra* note 176, at 9.

203. Hart et al., *supra* note 183, at s62.

204. *Id.*

205. Cf. R. Karl Hanson & Phillip D. Howard, Commentary, *Individual Confidence Intervals Do Not Inform Decision-Makers About the Accuracy of Risk Assessment Evaluations*, 34 LAW & HUM. BEHAV. 275, 275 (2010) (“[I]ndividual confidence intervals provide little information concerning the accuracy of a risk assessment. When the outcome is dichotomous, the confidence intervals for recidivism prediction will almost always range from zero to one (i.e., be uninformative). Consequently, other indicators of predictive accuracy are needed, many of which are non-quantitative.”).

206. David DeMatteo & John F. Edens, *The Role and Relevance of the Psychopathy Checklist-Revised in Court: A Case Law Survey of U.S. Courts (1991–2004)*, 12 PSYCHOL. PUB. POL’Y & L. 214, 215 (2006).

correlation coefficient and the Receiver Operating Characteristic (ROC). This section will review research results using these statistical measures and consider them from practical and critical perspectives.

The correlation coefficient is a statistic ranging from -1.0 to 1.0 that indicates the direction (positive or negative) and strength of the linear relationship between two variables. For example, height is strongly and positively correlated with weight, such that the taller a person is the more he or she is likely to weigh. A correlation coefficient of 0 means no correlation while -1.0 or 1.0 indicates perfect correlation. For our purposes here, we are concerned with how strong the actuarial assessment is positively associated with sex offense recidivism.

In meta-analyses of international samples, researchers have observed correlation coefficients of .27²⁰⁷ and .28²⁰⁸ for RRASOR and .33 for STATIC-99.²⁰⁹ In the social sciences, the strength of a correlation coefficient of .30 is considered to be only moderately predictive.²¹⁰ Yet, since positive correlation coefficients range from 0 (no correlation) to 1.0, the results are not very strong as a practical matter, particularly when considering the imposition on liberties and privacy at issue.

Another statistic supports this claim that these tests are rather weak statistical tools. The correlation coefficient leads to a percentage of the variance statistic (r^2) that permits a better understanding of what the instrument can actually account for in terms of recidivism.²¹¹ Taking the higher correlation listed for STATIC-99, .33, the variance is its square (.33²) which equals .1, or 10%. Thus, 10% of the variance in sexual recidivism can be explained by the STATIC-99 factors. Alternatively, this means that 90% of what helps influence sex-offense recidivism is based on other factors. This further suggests that STATIC-99 has little practical significance even if there is a statistically significant correlation with sexual reoffense. Commentators describe this result as meaning STATIC-99's "high risk" label is a misnomer, even for its highest risk category of 6+, and that STATIC-99's performance is not much "better than a coin flip."²¹²

A more recently adopted statistical measure of predictive accuracy is the Receiver Operating Characteristic (ROC), which is derived from a plotting of true positives and false positives.²¹³ Proponents of this statistic indicate that, unlike a correlation coefficient, the ROC is not reliant upon the base rate of the sample and is therefore useful in order to compare the accuracy of different instruments on samples with differing base rates.²¹⁴ ROC statistics range from 0 to 1.0. A ROC of 0 means the

207. Hanson & Thornton, *supra* note 168, at 121.

208. *Id.* at 126 tbl.4.

209. *Id.*

210. See JACOB COHEN, STATISTICAL POWER ANALYSIS FOR THE BEHAVIORAL SCIENCES 79–80 (2d ed. 1988) (characterizing $r=.10$ as low, $r=.30$ as medium, and $r=.50$ as high).

211. See Berlin et al., *supra* note 157, at 379 (explaining that correlation coefficient of .13 between RRASOR score and recidivism rate means that only two percent of variance in recidivism rate can be attributed to factors measured by RRASOR).

212. *Id.* at 381.

213. CHRISTOPHER SLOBOGIN, PROVING THE UNPROVABLE: THE ROLE OF LAW, SCIENCE, AND SPECULATION IN ADJUDICATING CULPABILITY AND DANGEROUSNESS 107 (2007).

214. Harris et al., *supra* 182, at 413.

instrument is completely inaccurate in its predictive ability, whereas a ROC of 1.0 means the instrument is completely accurate; a ROC value of .50 means that the predictive ability of the instrument is no better than chance, much like the proverbial coin flip.²¹⁵ Of course, various studies may report different ROC values where sample characteristics vary (such as region, type of sex offender, or treatment success) or the study methods differ.

STATIC-99 has been tested by one of its developers using international samples with ROC scores of between .63²¹⁶ and .70.²¹⁷ The developer of RRASOR has also observed ROC scores in various meta-analyses from .59²¹⁸ to .68.²¹⁹ In another small study by others, researchers directly compared the ROC rates on the same group of offenders, yielding comparative ROC scores of .68 for STATIC-99 and .73 for RRASOR.²²⁰ Interestingly, STATIC-99 did not fare better in its accuracy than RRASOR—it fared slightly worse in some studies—despite the addition of six factors to the four from RRASOR.

Correctly interpreting the ROC scores is important. The ROC score of .63 for STATIC-99 means that the test yields a 63% chance that a recidivist will receive a higher risk ranking than a nonrecidivist. This does not mean that 63% of the group with those characteristics will reoffend (since the statistic is not dichotomous). It also does not mean that there is a 63% chance the individual being scored will recidivate (as the model is based on group statistics).²²¹ Rather, the ROC statistic is about the accuracy of the relative rankings of the test.²²² The value represents the “probability that a randomly selected recidivist would have a more deviant score than a randomly selected nonrecidivist.”²²³

Since the ROC statistic obscures base rate differences between groups, some critics assert that it leads to an overestimation of risk predictions when base rates change over time and when applied to other groups with lower base rates than the developmental sample.²²⁴ The authors of STATIC-99 now seem to agree.²²⁵ Recently, the STATIC-99 developers admitted that their original recidivism experience tables overestimated recidivism risk in light of reductions in base rates of sexual recidivism.

215. SLOBOGIN, *supra* note 213, at 107.

216. R. KARL HANSON & KELLY MORTON-BOURGON, PREDICTORS OF SEXUAL RECIDIVISM: AN UPDATED META-ANALYSIS 32 (2004).

217. R. KARL HANSON & KELLY E. MORTON-BOURGON, THE ACCURACY OF RECIDIVISM RISK ASSESSMENTS FOR SEXUAL OFFENDERS: A META-ANALYSIS 31 (2007).

218. *Id.*

219. Hanson & Thornton, *supra* note 168, at 126.

220. Howard E. Barbaree et al., *Evaluating the Predictive Accuracy of Six Risk Assessment Instruments for Adult Sex Offenders*, 28 CRIM. JUST. & BEHAV. 490, 510 (2001).

221. Fennel, *supra* note 58, at 54.

222. *Id.*

223. Hanson & Thornton, *supra* note 168, at 125.

224. Fennel, *supra* note 58, at 59–61.

225. See R. Karl Hanson et al., *Predicting Recidivism Amongst Sexual Offenders: A Multi-site Study of STATIC-2002*, 34 LAW & HUM. BEHAV. 198, 198–99 (2010) (“More complex risk assessment systems are available that may be more accurate than STATIC-99 for predicting sexual recidivism and violent recidivism . . .”).

Indeed, they no longer recommend that experts use the original estimates.²²⁶ In a meta-analysis of international samples released in 2009, the developers found that the average base rate of recidivism was 11.6% from over 100 studies with follow-up periods ranging from 6 to 276 months.²²⁷ Compared to the original STATIC-99 data, the 11.6% base rate in the more recent samples was two-thirds of what was found in the original samples.²²⁸ The developers issued new recidivism estimates to replace the original estimates, along with new criteria for the use of the replacement-experience table.²²⁹ However, the new table and criteria have not yet been subject to cross-validation, and it will likely take some time for the field to determine how to consider the new evidence. To maintain its statistical validity, therefore, potential fluctuations in base rates over time require that developers re-estimate the tool's risk scores even for the reference group to which it should apply.

It is also important to recognize that one of the most important limitations of actuarial assessments as a rule is the problem of overgeneralization or, more empirically, external validity. One overgeneralizes results of research by presuming the results derived from one population (the reference group) are reliable when applied to a second population. If the second population differs in any risk-relevant way from the reference group, then the predictive result is invalid.²³⁰ As actuarial testing of future dangerousness for sex offenders is a relatively recent phenomenon and almost exclusively accomplished on adult male offenders released from prison or mental institutions, the reference group is notably limited in several risk-relevant ways. One engages in overgeneralization by applying the same actuarial estimates on sexual recidivism to groups bearing any risk-relevant attribute that significantly differs from the reference group, such as women, juveniles, incest offenders, older offenders, first time offenders, and those who were not incarcerated.²³¹ A number of studies have highlighted this overgeneralization problem, by showing that the instruments vary, sometimes dramatically, when trying to predict sexually violent recidivism in subgroups.²³²

226. Helmus et al., *supra* note 168, at 39.

227. R. Karl Hanson & Kelly E. Morton-Bourgon, *The Accuracy of Recidivism Risk Assessments for Sexual Offenders: A Meta-Analysis of 118 Prediction Studies*, 21 *PSYCHOL. ASSESSMENT* 1, 4, 6 (2009).

228. Helmus et al., *supra* note 168, at 2; *see also* Leslie Helmus, *Re-Norming Static-99 Recidivism Estimates: Exploring Base Rate Variability Across Sex Offender Samples 67* (Sept. 2009) (unpublished M.A. thesis, Carleton University), *available at* <http://www.static99.org/pdfdocs/helmus2009-09static-99normsmathesis.pdf> (combining multiple studies and finding base rate of 60% of original STATIC-99 developmental samples).

229. Abbott et al., *supra* note 176, at 1; *see also* David Thornton et al., *Moving Beyond the Standard Model for Actuarial Assessment for Sexual Offenders*, *THE FORUM*, Summer 2010, at 17, 18–19 (acknowledging that previous tables were unreliable).

230. Prentky et al., *supra* note 6, at 373.

231. *Id.* at 373–74.

232. *See* Eher et al., *supra* note 182, at 8 (testing STATIC-99 on child sexual abusers and rapists); Looman, *supra* note 182, at 193 (finding variances of STATIC-99 on treated high-risk offenders); *see also* Abbott, *supra* note 176, at 6 (citing additional studies showing subgroup differences); R. Karl Hanson, *Does STATIC-99 Predict Recidivism in Older Sexual Offenders?*, 18 *SEXUAL ABUSE* 343, at 344 (2006) (finding base rate of 2% in 5 years for over age 60, compared to 15% for under age 40).

It is also prudent to be cognizant of the potential differences in reoffending by geographic and cultural region. For instance, FBI statistics indicate that criminal offending can vary, sometimes significantly, by state and geographic region of the United States, including for sexual offenses.²³³ An author of STATIC-99 conducted a 2009 meta-analysis of international samples that underscored the geographic disparity in predictive ability: the average STATIC-99 ROC was .90 for the United Kingdom, but only .60 for the United States and .58 in Canada.²³⁴ Hence, using recidivism statistics based largely on incarcerated populations in England and Canada “should be a great cause of concern for making recidivism predictions in the United States”²³⁵ as a whole, much less to any particular region of the United States where recidivism risk may vary.

Another conceptualization of the practical significance of the ROC scores concerns the actuarial tool’s positive predictive accuracy, or the accuracy of predicting reoffending.²³⁶ If we borrow the Department of Justice-sexual recidivism rate of 5.3%, rounding up to 6% for easy interpretation, and then apply a ROC score of .70, the positive predictive accuracy measures indicate that the actuarial tool will be wrong 9 times out of 10.²³⁷ This is a common problem as the prediction of relatively rare events is inherently unreliable.²³⁸ Others tend to agree that because of the high incidence of false positives with these actuarial tests, the uncritical use of them produces systematic overestimation of risk.²³⁹

3. Reliability and Objectivity

Actuarial assessments of risk carry an aura of science and objectivity.²⁴⁰ Perhaps this is because the use of numerical percentages and rankings of bounded tiers imbues the predictions with a connotation of mathematical precision.²⁴¹ In addition, the

233. FBI, *Crime in the United States: By Region, Geographic Division, and State, 2006–2007*, FBI.GOV, http://www2.fbi.gov/ucr/cius2007/data/table_04_dd.html (last visited Aug. 24, 2011).

234. Hanson & Morton-Bourgon, *supra* note 227, at 7; *see also* Hanson et al., *supra* note 169, at 203 (comparing samples and finding U.S.-based sample had lowest ROC of .61 compared to samples in Canada and United Kingdom).

235. Fennel, *supra* note 58, at 59 (maintaining that Canada’s sexual assault rates are more than twice that of United States).

236. Leam A. Craig et al., *Limitations in Actuarial Risk Assessment of Sexual Offenders: A Methodological Note*, BRIT. J. FORENSIC PRAC., Feb. 2004, at 16, 18.

237. CRAIG ET AL., *supra* note 184, at 41 (providing formula for result).

238. Prentky et al., *supra* note 6, at 374.

239. Eher et al., *supra* note 182, at 41, 42. *See generally* Tamara Rice Lave, *Controlling Sexually Violent Predators: Continued Incarceration at What Cost?*, 14 NEW CRIM. L. REV. 213, 236–49 (2011) (demonstrating that significant amount of false positives result from use of STATIC-99).

240. Peter R. Jones et al., *Risk Classification and Juvenile Dispositions: What Is the State of the Art?*, 79 TEMP. L. REV. 461, 495 (2006).

241. *See* ERICA BEECHER-MONAS, *EVALUATING SCIENTIFIC EVIDENCE: AN INTERDISCIPLINARY FRAMEWORK FOR INTELLECTUAL DUE PROCESS* 38–39 (2007) (“Regardless of whether the scientific method exists in practice, people—scientists and nonscientists alike—frame their arguments as though it does.”); TOM JAGTENBERG, *THE SOCIAL CONSTRUCTION OF SCIENCE: A COMPARATIVE STUDY OF GOAL DIRECTION, RESEARCH EVOLUTION AND LEGITIMATION* 12–40 (1983) (arguing that science is socially constructed through goal-oriented behavior of researchers and that its accreditation is negotiated).

consistent use of the same factors and scoring methodology across cases reduces the appearance of bias.²⁴²

A relevant measure of reliability involves interrater reliability, which measures how consistent observers are in rating the same variable with the same value. In small studies, study observers rate quite high at assigning consistent results with the same actuarial tools; the correlation coefficients (1.0 meaning perfect) for interrater reliability has been found to be .95 for RRASOR and .87 for STATIC-99.²⁴³ Still, these scores provide estimates of raters in supervised studies, presumably with guidelines to assist consistency scores. Outside the structured study environment, there may be less consistency in scoring the instruments, particularly when adversarial allegiance exists. A study on rater agreement using STATIC-99 and another actuarial tool in adversarial civil commitment hearings found that while interrater reliability was high, there was greater variation in ratings by evaluators on opposing sides in SVP hearings than between experts on the same side.²⁴⁴ The same study also found that, on average, state experts reported higher risk-score computations than experts retained by the defendants.²⁴⁵

Two other issues pertaining to objectivity are concerning. Another type of allegiance effect occurs in which developers of several of the actuarial tools may have professional incentives to conduct further studies that report results supporting the validity of their own tools.²⁴⁶ One study found evidence of allegiance in peer-reviewed validation studies in which the instrument's author(s) participated.²⁴⁷ This study's researchers compared the validity coefficients between studies conducted by the authors of three actuarial instruments for sexual recidivism, including STATIC-99, and found that on average the instrument authors reported significantly larger correlation coefficients (average $r=.37$) compared to non-authors (average $r=.28$).²⁴⁸ A recent meta-analysis found that the average ROC for STATIC-99 varied dramatically depending on whether the studies were published (ROC .80; $n = 21$) or not published (ROC .60; $n = 42$).²⁴⁹ This suggests that researchers are more likely to reveal studies with statistically stronger results.

It is also noted that while the RRASOR and STATIC-99 developers intended their instruments to be easily and objectively scored, there is some room for error based on the availability and veracity of the data. With developers encouraging raters to access a

242. Richard Wollert, *Low Base Rates Limit Expert Certainty When Current Actuarials Are Used to Identify Sexually Violent Predators: An Application of Bayes's Theorem*, 12 PSYCHOL. PUB. POL'Y & L. 56, 58 (2006).

243. Harris et al., *supra* note 182, at 416; *see also* Barbaree et al., *supra* note 220, at 499 (reporting correlation coefficients for interrater reliability of .94 for RRASOR and .90 for STATIC-99).

244. Murrie et al., *supra* note 189, at 39.

245. *Id.* at 40 (cautioning that these results are not definitive to show litigant bias).

246. Douglas P. Boer, *Ethical and Practical Concerns Regarding the Current Status of Sex Offender Risk Assessment*, 3 SEXUAL OFFENDER TREATMENT 1, 3 (2008), <http://www.sexual-offender-treatment.org/index.php?id=64&type=123>.

247. Pamela R. Blair et al., *Is There an Allegiance Effect for Assessment Instruments? Actuarial Risk Assessment as an Exemplar*, 15 CLINICAL PSYCHOL.: SCI. & PRAC. 346, 354 (2008).

248. *Id.* at 353.

249. Hanson & Morton-Bourgon, *supra* note 227, at 7.

broad spectrum of data, including self-reporting, the potential for error is real. Accurate scoring relies upon good data. Similarly, the probability of missing data may also skew results. When the factors involving arrest are considered, another source of error occurs since arrests may be overinclusive due to the fact that arrests are not legally sufficient proof of guilt. Accurate scoring also relies upon adequate training.

4. Training

The authors of the STATIC-99 write in their coding rules that they “strongly recommend training in the use of the STATIC-99 before attempting risk assessments that may affect human lives.”²⁵⁰ There are no criteria, however, for the scope, time, or regimen for training or otherwise certifying potential assessors on the actuarial instruments. There are also no formal or published coding rules or training manuals. Mostly, information is vicariously available on the internet and through occasional training classes.

5. General Acceptance

The use of actuarial tools for sex offenders in regular treatment is one matter, and there is no reason here to challenge their general acceptance and use in the treatment setting. Their use in court, where the stakes are qualitatively higher and the professional standards different, is another matter. Many mental health experts who work in sex offender treatment believe the actuarial instruments are currently best practice and are willing to use them and testify as to their conclusions.²⁵¹ There are reasons to believe, however, that the tide of approval of actuarial tools—if there indeed was ever general acceptance—is turning in recent years. Mental health professionals are starting to realize that declining base rates of sexual recidivism in the United States and the variability of base rates among different sex offender populations undermines the continued viability of the experience tables and ROC scores.²⁵² In an amicus brief in a death penalty case, the American Psychological Association recently contended that any prediction of dangerousness is unreliable in court if it does not consider the base rate of the specific population for a set period of time.²⁵³ The American Psychiatric Association’s recent stance has not addressed the actuarial tools directly,

250. HARRIS ET AL., *supra* note 170, at 3.

251. See *infra* Part V for a discussion of cases addressing the use of actuarial instruments in court.

252. Douglas Mossman, *Another Look at Interpreting Risk Categories*, 18 *SEXUAL ABUSE* 41, 60–61 (2006); Shoba Sreenivasan et al., *Predicting the Likelihood of Future Sexual Recidivism: Pilot Study Findings from a California Sex Offender Risk Project and Cross-Validation of the Static-99*, 35 *J. AM. ACAD. PSYCHIATRY & L.* 454, 466 (2007); see also Shoba Sreenivasan et al., *Alice in Actuarial-Land: Through the Looking Glass of Changing Static-99 Norms*, 38 *J. AM. ACAD. PSYCHIATRY & L.* 400, 405 (2010) (“Although they purport to be empirically based, the current Static-99 . . . violate[s] the basic tenets of evidence-based medicine that require reasoned, not mechanical, application of group findings to the individual. Two core elements must be present to apply an actuarial risk model to a specific individual: sample representativeness and uniform measurement of outcome. Both of these elements are lacking in Static-99 . . . research reviews. Thus, a call for caution must be sounded when using these tools to make weighty decisions involving an individual’s liberty and the protection of public safety.”).

253. Brief of Amicus Curiae American Psychological Ass’n in Support of Defendant-Appellant at 12–13, *United States v. Fields*, 483 F.3d 313 (5th Cir. 2007) (No. 04-50393).

but in a position statement, it asserts: “Although psychiatrists cannot predict dangerousness with definitive accuracy, they can often identify risk factors associated with an increased likelihood of violent behavior.”²⁵⁴

Thus, many legal and mental health practitioners and researchers who work in the sex offender area, and who feel strongly and justifiably about it enough to publish their professional opinions in peer-reviewed journals, warn against the use of actuarial tests in legal settings because significant limitations with the tests make their use questionable in light of the significant deprivation of liberty they may facilitate.²⁵⁵ Putting the use of actuarial evidence in context, a mental health practitioner extrapolated from the Department of Justice’s sex-offender recidivism study to conclude that using STATIC-99 would have averted only three percent of sexual offenses committed by released offenders (sex offender and non-sex offenders) while hundreds of nonrecidivists would have been unnecessarily detained.²⁵⁶ Further, some suggest it may be professionally unethical for mental health practitioners to testify in court about the likelihood of an individual reoffending,²⁵⁷ at least without being absolutely clear about all of the substantive limitations in making such predictions.²⁵⁸ Others argue that the group-based model of tools means that even if using the actuarial assessments may be appropriate at initial assessments to consider pursuing civil commitment, mental health professionals should decline to use them in actual court hearings about individual risk predictions.²⁵⁹ In sum, even if there had been general acceptance of RRASOR and STATIC-99 near the time of their inception, such general acceptance was likely only about their ability for treatment purposes. It is highly questionable whether there ever was—and even more questionable whether there is

254. AM. PSYCHIATRIC ASS’N, ASSESSING THE RISK FOR VIOLENCE: POSITION STATEMENT (Dec. 2001), available at <http://www.psych.org/Departments/EDU/Library/APAOfficialDocumentsandRelated/PositionStatements/200109.aspx>.

255. E.g., Hart, *supra* note 180, at 385–86; Prentky et al., *supra* note 6, at 360; Scott I. Vrieze & William M. Grove, *Predicting Sex Offender Recidivism. I. Correcting for Item Overselection and Accuracy Overestimation in Scale Development. II. Sampling Error-Induced Attenuation of Predictive Validity Over Base Rate Information*, 32 L. & HUM. BEHAV. 266, 276 (2008); Wollert, *supra* note 242, at 72–73. See generally Boer, *supra* note 246, at 2 (arguing that only thing risk researchers agree upon is general description of types of risk assessments, such as clinical, actuarial, and structured professional judgments).

256. Wollert, *supra* note 242, at 78–80.

257. JOHN MONAHAN, THE CLINICAL PREDICTION OF VIOLENT BEHAVIOR 13–16 (1981) (contending that extreme deprivation of liberty that civil commitment causes raises spectre of ethical concerns for psychologists and psychiatrists participating in legal system); Boer, *supra* note 246, at 3–4; Charles P. Ewing, “*Dr. Death*” and the Case for an Ethical Ban on Psychiatric and Psychological Predictions of Dangerousness in Capital Sentencing Proceedings, 8 AM. J. L. & MED. 407, 418 (1983); Hart et al., *supra* note 183, at 64; Prentky et al., *supra* note 6, at 383.

258. Theodore Donaldson & Richard Wollert, *A Mathematical Proof and Example That Bayes’s Theorem Is Fundamental to Actuarial Estimates of Sexual Recidivism Risk*, 20 SEXUAL ABUSE 206, 213–15 (2008).

259. Berlin et al., *supra* note 157, at 382; see also Eher et al., *supra* note 182, at 10–12 (stating that even if instruments are statically confirmed as better than chance, their practical value in legal settings should be questioned).

today—a general acceptance in the mental health field about the validity of using actuarial risk assessments in SVP legal determinations.²⁶⁰

V. JUDICIAL PERSPECTIVES ON FUTURE DANGEROUSNESS EVIDENCE

Since the Supreme Court approved mental health testimony about future dangerousness and found civil commitment of sexual predators and registration laws to be constitutional, the introduction of actuarial risk assessments through expert testimony has become common practice in SVP determinations. Empirical observations about actuarial predictions of future dangerousness, outlined in Part IV.B above, have not gone entirely unnoticed. In a general, retrospective critique, Janus and Prentky summarized the development of conflict between science and law in SVP litigation as follows:

Suddenly, courts were confronted with a number of potentially embarrassing facts: the group-based nature of risk assessment, the tension of applying probabilistic estimates from life tables to defendants who departed significantly from the membership of the reference groups used to derive the estimates, the difficulty of evaluating and incorporating dynamic risk factors, and the problem of translating statutory language into scientifically meaningful terms all became quite clear.²⁶¹

Some observers assert that the common judicial reaction to these issues has been to permit expert testimony without any meaningful inquiry into the scientific validity and reliability of actuarial assessments.²⁶² This Part examines the evidence that supports this observation by reviewing the main themes that emerge from a comprehensive review of case law in which either of the actuarial tools addressed in this Article was mentioned.

A. Daubert/Frye Challenges

Overall, courts have not been inclined to find challenges to the reliability of actuarial tests to be dispositive about the admissibility of actuarial evidence. Indeed, relatively few courts considering the admissibility of RRASOR or STATIC-99 have conducted any type of reliability analysis, whether *Daubert*, *Frye*, or a variant thereof. Most of the discussion in the cases has been on other issues, such as sufficiency of the evidence, due process, or ineffective assistance of counsel. The courts have employed common strategies for averting the reliability issue.

1. Frye Challenges

Of those cases referring to *Frye*, there has been a split as to whether *Frye* applied at all to the actuarial tests. First, most *Frye*-based courts have declined to hold *Frye*

260. See David J. Cooke & Christine Michie, *Limitations of Diagnostic Precision and Predictive Utility in the Individual Case: A Challenge for Forensic Practice*, 34 LAW & HUM. BEHAV. 259, 269–72 (2010) (discussing errors that may arise when extrapolating from group statistics for purpose of assessing a particular individual's future behavior).

261. Janus & Prentky, *supra* note 2, at 92.

262. Beecher-Monas & Garcia-Rill, *supra* note 25, at 306.

hearings to determine the admissibility of actuarial-based evidence based on these courts' determination that the actuarial tool is not a scientific test, making the *Frye* general acceptance test inapplicable.²⁶³ Most of these courts ended the matter there, without further explication. Still, a few opinions have referred to precedent holding that medical testimony does not constitute scientific evidence for purposes of *Frye*. According to these courts, since a mental health professional's assessment of future dangerousness is medical testimony, it cannot be scientific evidence.²⁶⁴ A few judges have also explained that because actuarial tools on future dangerousness have a predictive value "far less than 100%," they cannot "have an aura of scientific infallibility" with which *Frye* is concerned.²⁶⁵ This was true despite experts in other cases referring to the actuarial tools in scientific terms.²⁶⁶

The consequence of the no-science, no-*Frye* ruling was made quite clear by a California appellate court. In overruling the trial court's decision that STATIC-99 was unreliable, the appellate court stated that "while the accuracy rate of 71 percent may not meet the certainty requirements applicable to new scientific evidence, such requirements have no application to expert psychological opinion testimony based in part on actuarial instruments."²⁶⁷ A reasonable implication, then, was that if *Frye* applied, the appellate panel thought that the actuarial tool would be inadmissible. But, as shall be seen, this implication has not been adopted by other courts.

The alternative *Frye* path is evident among courts that assume, with little or no discussion, that actuarial tests of future dangerousness constitute scientific evidence for which a *Frye* hearing is appropriate. Almost all of these courts have found the tests to be reliable on the asserted basis that they are generally accepted and thereby require no further validation.²⁶⁸ At least one *Frye* court, however, has ruled that STATIC-99 is "not scientifically accepted" for the purpose of determining the requisite "mental abnormality" under the state's civil commitment statute (where recidivism risk was not at issue).²⁶⁹

263. *E.g.*, *State ex rel. Romley v. Fields*, 35 P.3d 82, 89 (Ariz. Ct. App. 2001); *People v. Therrian*, 6 Cal. Rptr. 3d 415, 420 (Ct. App. 2003); *Garcetti v. Superior Court*, 102 Cal. Rptr. 2d 214, 238 (Ct. App. 2000); *People v. Langhorne*, No. H027495, 2005 Cal. App. Unpub. LEXIS 11150, at *48 (Dec. 1, 2005); *In re Girard*, No. 103,404, 2011 Kan. App. LEXIS 103, at *3 (June 24, 2011).

264. *Fields*, 35 P.3d at 88; *Wilson v. Phillips*, 86 Cal. Rptr. 2d 204, 207–08 (Ct. App. 1999).

265. *Fields*, 35 P.3d at 89; *see also* *People v. Miller*, No. E035921, 2005 Cal. App. Unpub. LEXIS 3130, at *15 (Apr. 6, 2005) (stating that expert's testimony about using RRASOR and STATIC-99 "did not carry a misleading aura of infallibility").

266. *See, e.g.*, *People v. Castillo*, 89 Cal. Rptr. 3d 71, 77 (Ct. App. 2009) (describing state expert as "applying the scientifically validated 'Static-99' analysis to predict the likelihood that Castillo would commit future sex offenses"); *In re Civil Commitment of J.Z.M.*, No. A-0940-07T2, 2008 N.J. Super. Unpub. LEXIS 1343, at *8 (Mar. 6, 2008) (describing STATIC-99 as "well-validated actuarial instrument").

267. *Garcetti*, 102 Cal. Rptr. 2d at 240.

268. *Ortega-Mantilla v. Florida*, 898 So. 2d 1164, 1168 (Fla. 2005); *Lee v. State*, 854 So. 2d 709, 712 (Fla. Dist. Ct. App. 2003); *In re Commitment of Simons*, 821 N.E.2d 1184, 1196 (Ill. 2004); *In re Commitment of R.S.*, 801 A.2d 219, 220 (N.J. 2002); *In re Det. of Rudolph*, No. 48744-2-I, 2004 Wash. App. LEXIS 1203, at *4 n.2 (June 14, 2004); *see also In re Det. of Holtz*, 653 N.W.2d 613, 619 (Iowa Ct. App. 2002).

269. *State v. Rosado*, 889 N.Y.S.2d 369, 397 (N.Y. App. Div. 2009); *see also In re Commitment of J.P.*, 772 A.2d 54, 62 (N.J. Super. Ct. App. Div. 2001) (holding actuarial test invalid as applied to juveniles).

In the two primary applications of *Frye*, the impact of case law precedent has been of utmost importance in the courts' decisions. Most of the decisions have not sought to justify either the *Frye*-inapplicable ruling or the general acceptance determination with much of an independent analysis. Instead, almost all of the cases have relied expressly upon case precedent.²⁷⁰ As an example, the Illinois Supreme Court found it important that "at least 19 other states rely upon actuarial risk assessment in forming their opinions on sex offenders' risks of recidivism" and "eight of these states have directly addressed the *Frye* question."²⁷¹ The court recognized that reliance upon case law precedent may be a "hollow ritual," but justified it by arguing that the issue of general acceptance had been "thoroughly litigated" already in several other states.²⁷² The Illinois high court specifically pointed to a Florida case which held a *Frye* hearing and found general acceptance based on the affirming testimony of the experts in the underlying case and a list of academic papers.²⁷³ Notably, as the Florida court had acknowledged, many of the academic articles relied upon were authored by "those who have developed the actuarial tools."²⁷⁴

2. *Daubert* and Alternatives

Daubert analyses have been virtually absent in the SVP case law to date. The few courts to analyze actuarial risk assessments under *Daubert* have found them admissible. In *United States v. Shields*,²⁷⁵ the district judge summarily admitted actuarial-based predictions, concisely concluding the standards of general acceptance and peer review had been met, without further discussion.²⁷⁶ Another court, finding that STATIC-99 was scientific evidence, also summarily ruled—based upon the state expert's assertions—that the tool satisfied *Daubert*.²⁷⁷

A court in a state that has rejected *Frye*'s general acceptance standard concluded that actuarial evidence is sufficiently reliable to be admissible, based in part on finding no other courts had excluded it.²⁷⁸ A few cases have eschewed the reliability question by referring, in summary terms, to the influence of *Barefoot* on future dangerousness testimony.²⁷⁹ One court acknowledged the seeming inconsistency between *Daubert* and

270. *E.g.*, *Ortega-Mantilla*, 898 So. 2d at 1168; *Simons*, 821 N.E.2d at 1193.

271. *Simons*, 821 N.E.2d at 1192.

272. *Id.* at 1193.

273. *Id.* (citing *Roeling v. State*, 880 So. 2d 1234 (Fla. Dist. Ct. App. 2004)).

274. *See Roeling*, 880 So.2d at 1239 (citing, for example, four articles by Karl Hanson, the developer of STATIC-99).

275. No. 07-12056-PBS, 2008 U.S. Dist. LEXIS 13837 (D. Mass. Feb. 26, 2008).

276. *Shields*, 2008 U.S. Dist. LEXIS 13837, at *1-3 (basing decision solely upon precedent of *Simons*).

277. *In re Goddard*, 144 S.W.3d 848, 853 (Mo. Ct. App. 2004).

278. *In re Det. of Holtz*, 653 N.W.2d 613, 616 (Iowa Ct. App. 2002) (noting that, based on its review of case law from other jurisdictions, it concurred with assessment of New Jersey court that "no state appellate court decision . . . has found actuarial instruments inadmissible at SVP proceedings" (quoting *In re Commitment of R.S.*, 773 A.2d 72, 96 (N.J. Super. Ct. App.Div. 2001) (internal quotation mark omitted)), cited in *In re Det. of Shearer*, No. 5-858/05-0048, 2006 Iowa App. LEXIS 24, at *9-10 (Jan. 19, 2006).

279. *See, e.g.*, *R.S.*, 773 A.2d at 90 (concluding that since *Barefoot* accepted reliability of clinical judgment as to future dangerousness, then actuarial evidence must also be admissible).

Barefoot on future dangerousness testimony, but dismissed it.²⁸⁰ A few others have pointed to *Barefoot* as holding that the question of reliability in actuarial-assessment testimony goes to the weight of the evidence, not its admissibility.²⁸¹

Another strategy courts have employed to avoid directly addressing the reliability issue has been to contextualize the actuarial assessment as not being the sole basis for the expert's opinion.²⁸² One court, for instance, ruled that even if the experts used STATIC-99 heavily in their testimony, the fact that the results were merely a part of their overall assessments meant the defendant's "quibbles with their methodology in employing [STATIC-99] are irrelevant."²⁸³ Another court expressly declined to "second-guess" the experts in their use of the actuarial tests, contending it was up to the defense to have an expert challenge the tests rather than look to the courts to do so.²⁸⁴

Overall, there has been little substantive critique of the empirical quality of the actuarial-based assessments. Judges who have paid heed to the issue have done so mostly in lone dissents or concurrences.²⁸⁵ One concurring judge, for example, complained that RRASOR used "only" four simple factors in scoring.²⁸⁶ Another judge put it in simple terms:

Intuitively, I find it hard to believe that the knowledge that an 18-year-old man has one conviction for lewd behavior involving an unrelated boy is sufficient information to conclude that there is a 48.6% probability that the man will commit a violent sexual crime during the next decade. It also troubles me that [the defendant] can successfully complete a full course of rehabilitation and the RRASOR will not have changed its assessment of him.

280. *State ex rel. Romley v. Fields*, 35 P.3d 82, 87–89 (Ariz. Ct. App. 2001) (observing inconsistency between *Barefoot* and *Daubert/Kumho*, but finding actuarial evidence admissible on other grounds).

281. *People v. Litmon*, No. H021538, 2002 Cal. App. LEXIS 8195, at *56–57 (August 6, 2002); *In re Det. of Hauge*, 812 N.E.2d 571, 573 (Ill. App. Ct. 2004); *In re Girard*, No. 103505, 2011 Kan. App. LEXIS 103, at *7 (June 24, 2011); *In re Det. of Davenport*, No. 23545-9-III, 2010 Wash. App. LEXIS 1799, at *32 (Aug. 5, 2010); *In re Det. of Thorell*, 72 P.3d 708, 725 (Wash. Ct. App. 2003).

282. *Orozco v. Ahlin*, No. CV 08-5504 AHM, 2008 U.S. Dist. LEXIS 108797, at *51–52 (C.D. Cal. Dec. 3, 2008); *In re Commitment of Field*, 813 N.E.2d 319, 327 (Ill. App. Ct. 2004).

283. *People v. Clotfelter*, No. C038351, 2002 Cal. App. LEXIS 8921, at *33 (Sept. 25, 2002).

284. *In re P.F.*, 712 N.W.2d 610, 616 (N.D. 2006).

285. *See, e.g., In re Commitment of Burton*, 884 So. 2d 1112, 1118 (Fla. Dist. Ct. App. 2004) (Altenbernd, C.J., concurring) (arguing that predicting future human behavior is likely novel and that relevant scientific community for assessing general acceptance of actuarial tests should be comprised of more than small set of professionals who profit from them); *Lee v. State*, 854 So. 2d 709, 718 (Fla. Dist. Ct. App. 2003) (Casanueva, J., concurring) ("Unfortunately, the existing actuarial tools do not seem to address all relevant static risk factors."); *Girard*, 2011 Kan. App. LEXIS 103, at *7 (Malone, J., concurring) (noting that actuarial test should be subject to *Frye* testing but that defendant had not challenged on appeal); *In re G.R.H.*, 2011 N.D. 21, 25 (Kapsner, J., dissenting) (commenting that actuarial instruments cannot reliably predict an individual's risk); *In re Anderson*, 730 N.W.2d 570, 583–90 (N.D. 2007) (Kapsner, J., dissenting) (analyzing empirical literature about STATIC-99 and criticizing use of actuarial assessments). *Cf. United States v. C.R.*, No. 09–CR–155, 2011 WL 1901645, at *84 (E.D.N.Y. May 16, 2011) ("[U]se of these tests [including STATIC-99] on an individual who has never been charged or convicted of a hand-on sexual offense is highly controversial; the court finds their use unacceptable because it is contrary to acceptable scientific usage in the present case.").

286. *Lee*, 854 So. 2d at 717 (Casanueva, J., concurring).

With or without successful treatment, he has a 48.6% chance of doing a bad act in the future according to this test.²⁸⁷

A dissenting judge in another case also criticized the use of STATIC-99 due to its failure to incorporate dynamic factors that could reduce the individual's risk and warned: "Does anyone remember the Soviets' misuse of their mental health system for incarcerating enemies of the state? Does this seem at all similar?"²⁸⁸ Notably, in contrast to the majority opinions, the critical opinions have gone beyond legal precedent and have cited extensively to empirical publications challenging actuarial assessments.²⁸⁹

The eminent Judge Posner of the U.S. Court of Appeals for the Seventh Circuit has done perhaps the best job of trying to understand the limits of actuarial assessment, albeit in a sentencing guidelines case. In *United States v. McIlrath*,²⁹⁰ the defendant called a forensic psychologist to testify at the defendant's sentencing hearing on a charge involving internet predation of a minor. The doctor testified he used STATIC-99 and derived an estimate of a nine to thirteen percent sexual recidivism risk.²⁹¹ McIlrath appealed on the claim that his sentence was too severe. While Judge Posner affirmed the sentence, he pointedly questioned the expert's use of the actuarial tool. Posner expressed frustration that neither party demonstrated whether STATIC-99 had been validated by generally accepted methods or whether the test would pass the *Daubert* admissibility standard.²⁹² Posner then, seemingly *sua sponte*, raised the empirical issues that would support the devaluation of the expert's risk assessment, though without making any conclusions.²⁹³ The court indicated that even the advocates of STATIC-99 admitted its moderate predictive accuracy and that even though "[i]t may be more accurate than clinical assessments, . . . that might not be saying much."²⁹⁴

Evidence abounds that glossing over of reliability issue may be pragmatic. This can be implied from some of the unnecessary dicta that judges have provided in cases where the relevant *Daubert* or *Frye* standard does not exclude the actuarial evidence. In support of its approval of actuarial evidence, for example, a state court pointed to the fact that "in several jurisdictions actuarial risk assessment is mandated by either statute or regulation."²⁹⁵ In a more dramatic ceding of this question, one opinion involving a challenge to STATIC-99 evidence stated that courts must "respect [the] policy of [the]

287. *Burton*, 884 So. 2d at 1119–20 (Altenbernd, C.J., concurring).

288. *In re Murrell*, 215 S.W.3d 96, 115–17 (Mo. 2007) (Wolf, C.J., dissenting).

289. *See, e.g., In re Anderson*, 730 N.W.2d 570, 583–90 (N.D. 2007) (Kapsner, J., dissenting) (providing citations to peer-reviewed studies and lamenting uncritical use of actuarial assessments which have become substitutes for judicial judgments); *In re Murrell*, 215 S.W.3d at 115–16 (Wolf, C.J., dissenting) (analyzing methodology and mechanics of STATIC-99 and referencing publication regarding actuarial assessments); *see also Burton*, 884 So. 2d at 1119–20 (Altenbernd, C.J., concurring) (discussing generally the issue of potential false positives for RRASOR). *But see State v. Rosado*, 889 N.Y.S.2d 369, 397 (N.Y. App. Div. 2009) (reviewing empirical faults of STATIC-99 and ruling it inadmissible per *Frye* to inform factual finding of mental abnormality).

290. 512 F.3d 421 (7th Cir. 2008).

291. *McIlrath*, 512 F.3d at 424.

292. *Id.*

293. *Id.* at 425.

294. *See id.* (citing Janus & Prentky, *supra* note 40, at 1455–58).

295. *In re Commitment of Simons*, 821 N.E.2d 1184, 1194 (Ill. 2004).

legislature with respect to the trustworthiness of psychiatric opinion evidence in cases involving sexually dangerous persons.”²⁹⁶ Other courts, after ruling that neither *Daubert* nor *Frye* applied, have repeated that, “where the trier of fact is required by statute to determine whether a person is dangerous or likely to be dangerous, expert prediction may be the only evidence available.”²⁹⁷

An unusual case involving a conflicted expert helps to further illustrate the role of pragmatism. In *State v. Nichols*,²⁹⁸ a sex-offender registration case, the defendant was classified based, in part, on being assigned a score of 5 on STATIC-99.²⁹⁹ In appealing the risk designation to the court, the defendant called as an expert a mental health professional who was at that time a member of the Sex Offender Review Board (SORB), the state agency responsible for making initial determinations of a sex offender’s classification level.³⁰⁰ When challenged about the STATIC-99 score, the expert responded “I’ve never been a Static-99 fan.”³⁰¹ The judges were clearly appalled:

This response the Court finds to be somewhat shocking in that the SORB consistently uses this risk assessment tool in deciding the risk to re-offend. One would think that as the professional, a “sex offender treatment specialist” appointed to the board because of his expertise would have challenged the use of this tool in assessing risk.³⁰²

As a result, the court found that the defense expert’s testimony was not credible and it upheld the state’s classification level.³⁰³

B. *The Standard of Likelihood to Sexually Reoffend*

The intersection between legal language and scientific knowledge used in the mental health fields has proven challenging in SVP cases. The terminology used in the SVP laws varies somewhat by state and it is unclear whether there is any meaningful difference between them. For example, in civil commitment laws, the future dangerousness concept for sexual violence is variously described in statutes as “likely,”³⁰⁴ “more likely than not,”³⁰⁵ “substantial probability,”³⁰⁶ and “irresponsible

296. *Commonwealth v. Parks*, 2005 Mass. Super. LEXIS 225, at *12 (Mar. 9, 2005) (citing *Commonwealth v. McGruder*, 205 N.E.2d 726 (Mass. 1965)).

297. *People v. Ward*, 71 Cal. App. 4th 368, 374 (1999), quoted in *Black v. Voss*, 557 F. Supp. 2d 1100, 1108 (C.D. Cal. 2008); *Carmony v. Hunter*, No. CIV S-03-0927 DFL JFM P, 2006 U.S. Dist. LEXIS 92099, at *51 (E.D. Cal. 2006); *People v. Meyers*, No. C042511, 2005 Cal. App. LEXIS 4872, at *10–11 (June 2, 2005).

298. No. PM 04-3211, 2006 R.I. Super. LEXIS 187 (Dec. 18, 2006).

299. *Nichols*, 2006 R.I. Super. LEXIS 187, at *3.

300. *Id.* at *5–6.

301. *Id.* at *8.

302. *Id.*

303. *Id.* at *9.

304. ARIZ. REV. STAT. § 36-3701(7)(b) (2010); CAL. WELF. & INST. CODE § 6600(a)(1) (West 2010); MASS. GEN. LAWS ANN. ch. 6, § 178C (West 2004 & Supp. 2010); 42 PA. CONS. STAT. ANN. § 6403(a)(3) (2010); S.C. CODE ANN. § 44-48-30(1)(b) (2009); TEX. HEALTH & SAFETY CODE ANN. § 841.003(a)(2) (2007); VA. CODE ANN. 37.2-900 (2005 & Supp. 2010).

305. MO. ANN. STAT. § 632.480(5) (West 2011).

for personal conduct with respect to sexual matters.”³⁰⁷ Two statutes with the “likely” term define it further as “more likely than not”³⁰⁸ while other laws define it to be a “propensity . . . of such a degree as to pose a menace to the health and safety of others.”³⁰⁹ These definitions are hardly clear. Does the “more likely than not” mean fifty-one percent chance? Some mental health experts working with sexual predator assessments seem to think so.³¹⁰ The California Supreme Court rejected the idea that “likely” was synonymous with “more likely than not,” ruling instead that it referred to a “*serious and well-founded risk*.”³¹¹ This court implied this standard was less burdensome than “more likely than not” but provided no substantive guidance. Another court defined the statutory “likely” language to mean “probable rather than merely possible.”³¹²

With the allure of the seeming certainty from the definitive numbers derived from the actuarial tools, one could reasonably desire—considering the dramatic consequences of SVP laws—that legislatures provide more specific guidelines on the threshold levels of risk that equate to likelihood of reoffending. However, “while some states” have quantified this requisite threshold, “most have not.”³¹³ The normative question of what the threshold should be is one “that science simply cannot answer.”³¹⁴ Would a one-percent risk of recidivism be sufficient to justify civil commitment, or would the risk need to be well above fifty percent?³¹⁵ The lack of any meaningful articulation of the legal standards, coupled with the empirically doubtful nature of future-dangerousness actuarial tools, creates a risk of arbitrary and capricious decisions as courts apply vague laws with inadequate science.

The criteria for sex offender registry tiers are often no more helpful. Several state registration acts provide three levels of classification that relate to the specific registration requirements, differentiating them with simple categorizations of risk of

306. 725 ILL. COMP. STAT. ANN. § 207/15(b)(5) (West 2006). Wisconsin’s legislature changed its standard from “substantially probable” to “likely.” Wis. Stat. § 980.01 (2010), *amended by* S.B. 441, 2003 Leg. (Wisc. 2004).

307. MINN. STAT. § 253B.02 Subd. 18b. (West 2011).

308. IOWA CODE ANN. § 229A.2(4) (West 2011); WIS. STAT. § 980.01(1m) (2010).

309. FLA. STAT. ANN. § 394.912(4) (West 2008); KANSAS STAT. ANN. § 59-29a02(c) (2008); *see also* NEB. REV. STAT. ANN. § 83-174.01(2) (2008) (containing slight variation from cited definition); N.H. REV. STAT. ANN. § 135-E:2(VI) (2010) (same); N.J. STAT. ANN. § 30:4-27.26 (West 2008) (same).

310. *See generally* Abbott, *supra* note 176, at 18 (defining “more likely than not” as greater than fifty percent chance of reoffending).

311. *Cooley v. Superior Court*, 57 P.3d 654, 670–71 (Cal. 2002).

312. *Martin v. Reinstein*, 987 P.2d 779, 800 (Ariz. Ct. App. 1999).

313. Erickson, *supra* note 191.

314. *Id.*

315. Fennel, *supra* note 58, at 39–40. Fennel cites RON SUSKIND, *THE ONE PERCENT DOCTRINE: DEEP INSIDE AMERICA’S PURSUIT OF ITS ENEMIES SINCE 9/11* 62 (2006) (referring to then Vice-President Dick Cheney’s stance that preventive measures are justifiable if there is even a one percent chance of a security threat) and George G. Woodworth & Joseph B. Kadane, *Expert Testimony Supporting Post-Sentence Civil Incarceration of Violent Sexual Offenders*, 3 L. PROBABILITY & RISK 221, 227 (2004) (arguing that language in state SVP statutes requires risk levels “substantially above 50%”).

sexual recidivism (“low,” “moderate,” and “high”), without further statutory definition.³¹⁶

The statutes also leave open the question about the relevant time period for which future dangerousness is assessed: one day; one year; five years; twenty-five years; life? It is further unclear whether the relevant time period differs depending on whether the case concerns civil commitment or registration. To date, few courts have addressed these issues. This is true despite RRASOR and STATIC-99 experience tables being typically based on five-, ten-, and fifteen-year follow-up periods. Of the few opinions addressing the time issue, one court, noting that the legislature had not specified any particular time period of risk for its SVP civil commitment statute, expressly declined to specify one.³¹⁷ Instead, the court, referring to the legislature’s recognition that sexual predators need long-term care, ruled that it would not adopt the one-year risk period that the defendant suggested.³¹⁸ On the other hand, other courts that have reviewed civil commitment statutes have acknowledged a lack of temporal specificity in the laws, but have noted that the present-tense language should be construed to mean dangerous at the time of the proposed commitment.³¹⁹ Still, these courts have summarily accepted the long-term actuarial evidence as relevant to determining the defendants’ immediate risk.³²⁰

*In re Civil Commitment of K.S.*³²¹ is a case that illustrates the temporal risk conundrum. The court considered the “high risk” assessment from STATIC-99 as relevant to the defendant’s current risk of reoffending, despite acknowledging that STATIC-99 measured long-term risk potential.³²² The risk assessment, however, was provided at a 2007 hearing for a continued civil commitment, while the high-risk score was based on offenses occurring in 1975, 1982, and 1989, and the defendant had been continuously incarcerated since his 1989 conviction.³²³ The court’s almost complete disregard of the temporal issue—by making a determination of immediate risk based on actuarial assessments for five to fifteen year risk periods—suggests a judicial concern for pragmatism, since no known empirical tools adequately assesses imminent risk. Yet, assessing immediate risk based upon long-term projections also suggests a certain arbitrary and capricious quality to SVP determinations.

With respect to the likelihood standard, there has been a strong tendency to highlight actuarial predictions that were around or exceeded the 50% mark, even if they

316. MASS GEN. LAWS ANN. ch.6 § 178K (West 2010); MONT. CODE ANN. § 46-23-509 (2009); N.Y. CORRECT. LAW § 168-l(6) (McKinney 2011); OKLA. STAT. ANN. tit. 57, § 582.5(c) (West 2011).

317. *In re Det. of Ewoldt*, 634 N.W.2d 622, 624 (Iowa 2001).

318. *Id.*; see also *In re Det. of Pierce*, 748 N.W.2d 509, 513 (Iowa 2008) (determining that state need not show immediate risk, so measures of lifetime risk were acceptable).

319. *In re Murrell*, 215 S.W.3d 96, 104–05 (Mo. 2007) (en banc); cf. *Beasley v. Molett*, 95 S.W.3d 590, 600 (Tex. Ct. App. 2002) (construing state statute language “menace” as “a threat or imminent danger”).

320. See, e.g., *In re Det. of Wilson*, No. 7-832/ 06-1625, 2007 Iowa App. LEXIS 1333, at *5 (Dec. 28, 2007) (contending that while civil commitment requires assessment of likelihood of reoffending at time of commitment, actuarial estimates of future risk remain relevant to that question).

321. No. A-1451-07T2, 2008 N.J. Super. Unpub. LEXIS 627 (March 11, 2008).

322. *K.S.*, 2008 N.J. Super. Unpub. LEXIS 627, at *6–7, 10.

323. *Id.* at *1–3, 9–10.

used fifteen-year experience scores to reach it.³²⁴ Opinions have clearly stated that the over 50% actuarial scores legally met the “more likely than not” standard.³²⁵ A STATIC-99 score of six can be enough to conclude the defendant “comes out over 50 percent,” thereby meeting the “likely” to sexually reoffend threshold.³²⁶ In a unique case, one court upheld the expert’s statement that an actuarial score of 52% meant the defendant was “still above the threshold that is represented by the term, ‘likely,’” despite the defendant’s voluntary castration.³²⁷ Other courts have been clear that high actuarial scores strongly influence their determination of SVP status.³²⁸

Notwithstanding the accepted relevance of 50% actuarial scores in supporting sufficiency of the evidence queries, many opinions have upheld sex-offender classifications in the face of actuarial scores of far less than 50%. Several opinions, for example, have referred to the state expert as starting their analysis by calculating the actuarial-based risk estimate and then adding percentage points, reportedly based on additional risk factors beyond the scope of the testing instruments.³²⁹ Some experts have described the actuarial numbers as providing a baseline for which the experts can adjust based on other factors.³³⁰ In contrast, another expert testified that “there was no empirical method of adding variables” to individually tailor the Static-99 results.³³¹ There is, indeed, no empirical evidence that modifying actuarial scores improves the accuracy of predictions.³³² Instead, it appears that the modified scores end up being

324. *E.g.*, *In re Rush*, 766 N.W.2d 720, 723 (N.D. 2009); *People v. Deberry*, No. D047707, D049686, 2008 Cal. App. Unpub. LEXIS 11, at *8 n.3 (Jan. 2, 2008); *Wilson*, 2007 Iowa App. LEXIS 1333, at *2; *Murrell*, 215 S.W.3d at 109; *In re Det. of Harless*, No. 6-1017/06-0033, 2007 Iowa App. LEXIS 66, at *6 (Jan. 31, 2007); *In re G.R.H.*, 758 N.W.2d 719, 723 (N.D. 2008); *People v. Therrian*, 6 Cal. Rptr. 3d 415, 417 (Ct. App. 2003); *In re Commitment of Lalor*, 661 N.W.2d 898, 903-04 (Wis. Ct. App. 2003). *But see* *Commonwealth v. Squire*, 685 S.E.2d 631, 633 (Va. 2009) (upholding trial court’s determination that defendant was not likely to reoffend despite STATIC-99’s 52% risk percentage where there was no evidence of any reoffending during prior ten years).

325. *E.g.*, *Wilson*, 2007 Iowa App. LEXIS 1333, at *6; *Harless*, 2007 Iowa App. LEXIS 66, at *11; *Murrell*, 215 S.W.3d at 103, 113; *In re Det. of Taylor*, 134 P.3d 254, 257 (Wash. Ct. App. 2006).

326. *Cooley v. Superior Court*, 57 P.3d 654, 659 (Cal. 2002) (quoting testimony of licensed psychologist regarding STATIC-99 data).

327. *People v. Flores*, 50 Cal. Rptr. 3d 567, 570-71, 573 (Ct. App. 2006) (quoting prosecutor’s expert) (internal quotation marks omitted).

328. *See State v. Dyer*, No. L-06-1357, 2007 Ohio App. LEXIS 4351, at *11 (Sept. 21, 2007) (stating lower court was heavily influenced by state experts’ agreement on high STATIC-99 scores in SVP status case); *State v. Twiggs*, No. 88142, 2007 Ohio App. LEXIS 1192, at *4 (Mar. 22, 2007) (acknowledging lower court’s belief that STATIC-99 score was “most clear and convincing evidence” to establish sexual predator status); *State v. Butler*, No. 86554, Ohio App. LEXIS 4412, at *9 (Aug. 31, 2006) (indicating that STATIC-99 score correlating to 52% risk “weighed heavily” in lower court’s determination of sexual predator status).

329. *See, e.g.*, *Garcetti v. Superior Court*, 102 Cal. Rptr. 2d 214, 228 (Ct. App. 2000) (noting state expert derived fifty-two percent score from STATIC-99 and then added five percentage points based on other risk factors), *rev’d on other grounds*, *Cooley*, 57 P.3d at 673-74; *Commonwealth v. Cowen*, 897 N.E.2d 586, 590 n.3 (Mass. 2008) (describing testimony by experts that STATIC-99 scores may be “adjusted upwards or downwards depending on risk factors”).

330. *E.g.*, *United States v. Shields*, 597 F. Supp. 2d 224, 236 (D. Mass. 2009); *People v. Stewart*, No. C037873, 2003 Cal. App. Unpub. LEXIS 4207, at *8 (Apr. 29, 2003); *People v. Pacini*, No. H022565, 2003 Cal. App. Unpub. LEXIS 325, at *33-34 (Jan. 10, 2003).

331. *People v. Langhorne*, No. H027495, 2005 Cal. App. Unpub. LEXIS 11150, at *13 (Dec. 1, 2005).

332. *Hart et al.*, *supra* note 183, at s64.

“little more than empirical window dressing for clinical judgment,” which the actuarial tools were designed to improve upon.³³³

Many other courts have affirmed sex-offender status despite actuarial-based risk estimates below fifty percent.³³⁴ In *Pedroza v. State*,³³⁵ for example, the sexual predator label was upheld in the face of low actuarial scores. In the case, the prosecutor had argued that even low scores from actuarial tools are sufficient to constitute the legal standard of “likely” to reoffend:

Even taking [the expert’s] tests, the RRASOR, about 11% failure rate after ten years. The SORAG was, I think—did he say 17% after seven years. Another one, . . . about 12% he’s likely to reoffend. The base rate for all offenders was 22%. All that sounds like ‘likely to reoffend,’ you know. The only—it’s only about 12% chance here that you have cancer or are going to die. Whoa, whoa, whoa! That’s pretty scary when we’re talking about human lives and behavior. That’s “likely.”³³⁶

The jury agreed and Pedroza was civilly committed. The appellate court affirmed, with a strong dissent arguing that the risk percentages were too low to legally constitute “likely” in face of the defendant’s liberty interest.³³⁷

C. *Battle of the Experts*

Case law provides considerable support for the criticism that actuarial assessments remain subject to the influences of adversarial allegiance and subjective bias. One area of disagreement has been whether or not the specific actuarial tools are generally accepted. Where there has been contention, state experts have tended to argue that the tools are generally accepted, while defense experts have asserted the opposite.³³⁸ When asked about professional acceptance of actuarial instruments (including RRASOR), one defense expert warned:

I think it’s a real concern here that these instruments promise something they don’t deliver. And they have an incredible aura of scientific certainty and preciseness that’s just not there if you peel away the second layer of the onion.

Therefore, I think psychologists do a disservice to the profession and psychiatrists, too, for that matter, when they use them and act as if there’s this precision and with a scientific basis that’s not really there.³³⁹

333. Prentky et al., *supra* note 6, at 380.

334. See *infra* notes 334–37, 350–52, and accompanying text for a discussion of cases where courts have upheld sex offender status despite actuarial-based risk estimates below fifty percent.

335. 773 So. 2d 639 (Fla. Dist. Ct. App. 2000).

336. *Pedroza*, 773 So. 2d at 640 n.3.

337. *Id.* at 643. (W. Sharp, J., dissenting) (internal quotation marks omitted).

338. *E.g.*, *In re Det. of Cain*, 931 N.E.2d 337, 339–41 (Ill. App. Ct. 2010); *People v. Avila*, No. B193902, 2008 Cal. App. Unpub. LEXIS 6125, at *23 (July 28, 2008); *People v. Wallace*, No. H027890, 2006 Cal. App. Unpub. LEXIS 1039, at *4–5 (Feb. 2, 2006); *People v. Jackson*, No. A100091, 2004 Cal. App. Unpub. LEXIS 2390, at *7–8 (Mar. 16, 2004); *In re Commitment of Field*, 813 N.E.2d 319, 322, 325 (Ill. App. Ct. 2004); *In re Commitment of R.S.*, 773 A.2d 72, 81–82 (N.J. Super. Ct. App. Div. 2001).

339. *R.S.*, 773 A.2d at 82. In accord with this statement, the defense argued that these actuarial tools may be unethical. *Id.* at 91–94.

Another defense expert more dramatically charged that the state's experts were conducting themselves unethically in using the actuarial tools.³⁴⁰

There has also been some disagreement between opposing experts in computing scores using the same actuarial instrument, in directions consistent with party affiliation. Almost always, the state experts give higher risk predictions than defense experts.³⁴¹ One state expert averred that a California study showed prisoners like the defendant with a STATIC-99 score of eight "had a 100 percent rate of sexually reoffending."³⁴² Another state expert opined that the high STATIC-99 score meant simply that the question was not if the defendant would sexually reoffend, but when.³⁴³

Many courts have permitted the experts themselves to opine as to whether the particular defendant is likely to reoffend. This is not surprising since many jurisdictions' evidence rules permit experts to testify on the ultimate issue,³⁴⁴ here the individual defendant's risk of recidivism. State experts have opined that the defendant is "highly likely,"³⁴⁵ has a "high likelihood,"³⁴⁶ and is "more likely than not"³⁴⁷ to sexually reoffend; that the defendant poses an "unacceptable public risk",³⁴⁸ and that the risks of reoffense are "substantially probable."³⁴⁹

As previously noted, many cases have found sufficient evidence of the defendant's likelihood of reoffending to assign SVP status when the actuarial results are around or above fifty percent. Even, however, when the actuarial results are below fifty percent, courts have generally upheld the states' cases for sex-offender status. In one case affirming the SVP classification for registration purposes (with a "likely" to reoffend criterion), the court was unconcerned with a STATIC-99 score that associated with a nineteen percent chance of reoffending in fifteen years: "[W]e believe that his one in five chance of committing a sexually oriented offense in the next 15 years could be viewed as unacceptably high."³⁵⁰ Judicial affirmance of SVP classifications has occurred, interestingly enough, even when the state's own expert(s) have given low actuarial scores.³⁵¹ Notably, in these cases, the state experts have supported the

340. *People v. Stewart*, No. C037873, 2003 Cal. App. Unpub. LEXIS 4207, at *9 (Apr. 29, 2003).

341. *E.g.*, *Sigman v. Rogers*, No. 07-1383(DMC), 2008 U.S. Dist. LEXIS 71127, at *22-23 (D.N.J. Sept. 3, 2008); *United States v. Abregana*, 574 F. Supp.2d 1145, 1155 (D. Haw. 2008); *People v. McKee*, 73 Cal. Rptr. 3d 661, 688-89 n.23-25 (Ct. App. 2008); *Kilgore v. Hurst*, No. L03-2611, 64 Va. Cir. 376, 2004 Va. Cir. LEXIS 178, at *7-8 (Apr. 20, 2004); *In re Mental Commitment of Good*, No. 2007AP279, 2008 WI App 17, ¶ 7, 2007 WL 4390443, at *2 (Wis. Ct. App. Dec. 18, 2007).

342. *People v. Rose*, No. H030534, 2007 Cal. App. Unpub. LEXIS 10017, at *7-8 (Dec. 12, 2007) (internal quotation marks omitted).

343. *People v. Torres*, 119 Cal. Rptr. 2d 597, 600 (Ct. App. 2002).

344. *E.g.*, FED. R. EVID. 704(a).

345. *In re Civil Commitment of Stringer*, No. A05-2240, 2006 Minn. App. Unpub. LEXIS 831, at *10 (Aug. 1, 2006).

346. *State v. Bieck*, No. 2003-CA-66, 2004 Ohio App. LEXIS 3223, at *25 (July 2, 2004).

347. *State v. Mruk*, No. L-04-1213, 2006 Ohio App. LEXIS 532, at *4-5 (Feb. 10, 2006); *In re Det. of Savala*, No. 20658-1-III, 2004 Wash. App. Lexis 1443, at *10 (Wash. Ct. App. July 13, 2004).

348. *In re Civil Commitment of Rustman*, No. A06-266, 2006 Minn. App. Unpub. LEXIS 683, at *6 (June 27, 2006).

349. *In re Commitment of Simons*, 821 N.E.2d 1184, 1188 (Ill. 2004).

350. *State v. Carter*, No. 88380, 2007 Ohio App. LEXIS 2476, at *8 (May 31, 2007).

351. *E.g.*, *In re A.B.*, 334 S.W.3d 746, 753-54 (Mo. Ct. App. 2011).

continued use and practical abilities of the actuarial instruments.³⁵² Presumably this widespread advocacy by state experts is due to the states' interest in maintaining the role of actuarial tools in sex offender proceedings. In order to accomplish this seemingly contradictory approach, the experts have employed, and the courts have adopted, two general tactics, often in combination: downplaying the absolute need for high actual-based risk estimates in every case, while endorsing clinical judgments to support raising the actuarial-based risk estimations.

Thus, while opining that the defendant is likely to reoffend, state experts have not claimed the actuarial tools are wholly flawed. Instead, their argument has been one directional: the experience tables underestimate risk, either as a whole or as applied to the individual. In other words, the state experts in these cases tend to convey that the instruments themselves are acceptable but that the experience tables underestimate (never overestimate) risk.³⁵³ Some experts have explained that consistent underestimates are possible because the actuarial tools are based on studies that used convictions, rather than charges, as the endpoint for measuring recidivism.³⁵⁴ There has been virtually no counter challenge that, in fact, the developers of RRASOR and STATIC-99 have made it clear the recidivism experience tables are based on samples with recidivism definitions that variously include charges, convictions, and readmissions.³⁵⁵

While most experts have not tried to estimate the size of the purported underestimation, there are some exceptions. One researcher, for example, contends that actual reoffense rates are three times higher than the rates reported in the actuarial tables,³⁵⁶ while another expert claims that the actual rates are five times those reported by the STATIC-99 instrument.³⁵⁷ In *People v. McGee*,³⁵⁸ the expert justified his hugely inflated assessment of the defendant's recidivism risk by arguing that, although the defendant only scored a three on STATIC-99 (for which the experience table provided a nineteen percent risk of sexual reoffense in fifteen years), he "was one of the thirty

352. *People v. Williams*, No. A100942, 2003 Cal. App. Unpub. LEXIS 11678, at *15 (Dec. 15, 2003); *A.B.*, 334 S.W.3d at 753–54.

353. *E.g.*, *People v. Castillo*, 89 Cal. Rptr. 3d 71, 77 (Ct. App. 2009); *In re Det. of Pierce*, 748 N.W.2d 509, 513–14 (Iowa 2008); *People v. Jackson*, No. A100091, 2004 Cal. App. Unpub. LEXIS 2390, at *6–7 (Mar. 16, 2004); *People v. Pacini*, No. H022565, 2003 Cal. App. Unpub. LEXIS 325, at *29–30 (Jan. 10, 2003); *see also* *Garcetti v. Superior Court*, 102 Cal. Rptr. 2d 214, 224, 231 (Ct. App. 2000) (noting state expert's testimony that STATIC-99 reflected "minimum potential for reoffense").

354. *See Huftile v. Hunter*, No. CIV S-05-0174, 2009 U.S. Dist. LEXIS 5311, at *15–16 (E.D. Cal. Jan. 15, 2009) (discussing state expert's testimony that STATIC-99 score does not consider possible offenses—only convictions—and therefore may not consider all of defendant's predatory acts); *People v. Orozco*, No. B199439, 2008 Cal. App. Unpub. LEXIS 7697, at *10 (Sept. 30, 2008) (discussing expert testimony that STATIC-99 underestimates because it considers only convictions, rather than offenses which may go unreported).

355. *See Hanson & Thornton, supra* note 168, at 122–23 tbl.2 (including and differentiating convictions from charges/readmissions in recidivism criteria).

356. *In re Det. of Shearer*, No. 5-858/05-0048, 2006 Iowa App. LEXIS 34, at *11 (Iowa Ct. App. Jan. 19, 2006).

357. *In re Civil Commitment of Luhmann*, No. A07-912, 2007 Minn. App. Unpub. LEXIS 890, at *10 (Aug. 28, 2007). The testifying expert also noted that defendant's high scores for likelihood to reoffend were supported by his socioeconomic status among other factors. *Id.*

358. 73 Cal. Rptr. 3d 661 (Ct. App. 2008).

percent of persons for whom the STATIC-99 test did not diagnose properly.”³⁵⁹ In another case, the expert argued that the STATIC-99 rate was inaccurately low as to the defendant since STATIC-99 is “slanted towards people who are not very good at hiding their offenses and a clever offender tends to get a lower score.”³⁶⁰

In many cases involving lower actuarial scores, state experts have cited the importance of other risk factors, observed in their clinical assessment, which are not considered by the instruments.³⁶¹ In accepting the clinically-derived risk factors as evidence to support a higher risk of future dangerousness, courts have concluded that the actuarial evidence, while relevant, is not dispositive as a matter of law.³⁶² This is evident where courts have upheld sex offender status decisions despite actuarial scoring of zero points.³⁶³ One court, which upheld a sexual-predator classification for a defendant with an actuarial score of zero (based on expert testimony about the presence of other risk factors), noted that even with an actuarial score of zero, the tools indicated at least a five percent risk level of reoffending within five years.³⁶⁴

There have also been differences in how both sides of the adversarial process have articulated the benefits and limitations of the actuarial tools. Defense experts have been much more critical about the accuracy of the actuarial instruments as a whole and much more likely to be cited as specifying weaknesses in the tests.³⁶⁵ Also, while state

359. *McKee*, 73 Cal. Rptr. 3d at 688 n.23; see also *In re Hehn*, 745 N.W.2d 631, 632–33 (N.D. 2008) (noting that one of two state experts declined to apply RRASOR and STATIC-99 because she believed they were “spuriously low” in defendant’s case since they did not account for allegations against defendant).

360. *People v. Langhorne*, No. H027495, 2005 Cal. App. Unpub. LEXIS 11150, at *17 (Dec. 1, 2005).

361. See, e.g., *People v. Clotfelter*, No. C038351, 2002 Cal. App. Unpub. LEXIS 8921, at *24–26 (Sept. 25, 2002) (noting state expert’s testimony that STATIC-99 has “shortcomings,” including propensity to underestimate risk of defendant reoffending, largely because it does not account for factors relevant to recidivism); cf. *In re Care and Treatment of Burgess*, 147 S.W.3d 822, 829 (Mo. Ct. App. 2004) (noting state expert’s argument that “more likely than not” standard was met despite actuarial results far less than 50% based on defendant’s failure to participate in treatment or reassessment, among other factors).

362. See, e.g., *Hehn*, 745 N.W.2d at 636 (“We have previously made clear that we will not engage in a ‘contest over percentage points’ when it comes to determining whether an individual meets the requirements for civil commitment.” (quoting *In re Interest of P.F.*, 712 N.W.2d 610, 617 (N.D. 2006) (Kapsner, J., concurring)) (internal quotation marks omitted)); *In re M.D.*, 757 N.W.2d 559, 562 (N.D. 2008) (declining to find that low actuarial scores were dispositive); *State v. Garner*, No. 89840, 2008 Ohio App. Lexis 1663, at *43–46 (Apr. 24, 2008) (holding that actuarial assessment can not be only factor considered in SVP status as Ohio law requires an individualized determination of risk); *State v. Vanek*, No. 89125, 2007 Ohio App. LEXIS 5433, at *12 (Nov. 21, 2007) (finding that “utility” of STATIC-99 to individual risk assessment is “open to question”); *State v. McKinniss*, 795 N.E.2d 160, 165 (Ohio Ct. App. 2003) (noting that experts argued “all psychological evaluations should generally be reviewed in light of the entirety of evidence”); *People v. Murphy*, No. H021758, 2001 Cal. App. Unpub. LEXIS 2770, at *19–20 (Nov. 21, 2001) (holding trial court was justified in relying on experts’ opinion that actuarial tools do not take into account all relevant factors for predicting recidivism).

363. E.g., *In re Interest of J.R.*, 762 N.W.2d 305, 326 (Neb. 2009) (affirming sexual-predator classification, despite actuarial score of zero by state expert, based on presence of other risk factors).

364. *State v. Johns*, No. 03 MA 61, 2004 Ohio App. LEXIS 4673, at *2–3 (Sept. 20, 2004). *Contra Commonwealth v. Monroe*, No. BRCV08-00895, 2008 Mass. Super. LEXIS 480, at *14–15, 22–24 (Dec. 26, 2008) (ruling in favor of defendant where state’s expert conceded score of zero on STATIC-99 and finding expert’s clinical assessment that defendant was likely to recidivate was unsupported by objective evidence).

365. See, e.g., *Cooley v. Superior Court*, 57 P.3d 654, 660 (Cal. 2002) (noting state expert’s concession that STATIC-99 is “only moderately reliable as an overall instrument,” but is “best we can come up with,” and

experts have often claimed that the experience tables underestimate the risk of reoffense, defense experts have argued the opposite.³⁶⁶ Indeed, various defense experts argue that the experience tables overestimate risk because the test overincludes the types of sex offenses that qualify under SVP laws (such as those involving violence),³⁶⁷ or because they are based largely on non-diverse samples of high-risk offenders from outside the United States³⁶⁸ where base rates of offending exceed the United States.³⁶⁹ Defense experts have also been more likely to decline calculating an actuarial score for defendants, based on their belief that the group-based models should not be used for individualized assessments,³⁷⁰ are generally inaccurate,³⁷¹ and/or are not validated on juveniles³⁷² or other subpopulations of which the defendant is a member.³⁷³ As noted

defense expert's testimony that STATIC-99 is "weak and inexact," with no known base rate of recidivism, overlapping predictors that might overestimate risk, and lack of standardization for Hispanic population (internal quotation marks omitted)); *In re Det. of Traynoff*, 831 N.E.2d 709, 714–16 (Ill. App. Ct. 2005) (citing state expert testimony that STATIC-99 is "work in progress but strongly relied upon" and defense expert's testimony that actuarial tools are controversial within mental health field and best understood, not as tests with manuals, but as research instruments (internal quotation marks omitted)); *People v. Taylor*, 830 N.E.2d 855, 857 (Ill. App. Ct. 2005) (citing state expert's comment that RRASOR and STATIC-99 are "reliable tool[s]," and defense expert's comment that these tools are "young pioneering efforts of novel science" that have "not been accepted by the psychological or psychiatric community" (internal quotation marks omitted)). *But see* *People v. Therrian*, 6 Cal. Rptr. 3d 415, 418 (Ct. App. 2003) (illustrating agreement between state and defense experts, with state expert testifying that STATIC-99's developer "never said it was perfect," and defense expert calling test "work-in-progress" with unknown reliability).

366. See *supra* notes 353–54 and accompanying text for a discussion of the frequent testimony from state experts that actuarial tools underestimate risk. See also *United States v. Shields*, 597 F. Supp.2d 224, 238–39 (D. Mass. 2009) (indicating disagreement on STATIC-99's risk estimation whereby state expert concluded age was not protective factor to warrant belief that tool overestimated defendant's risk while defense expert reached opposite conclusion).

367. *People v. Thomas*, No. C035786, 2002 Cal. App. Unpub. LEXIS 11548, at *8 (Dec. 13, 2002).

368. See *People v. Avila*, No. B193902, 2008 Cal. App. Unpub. LEXIS 6125, at *23 (July 28, 2008) (noting defense expert's testimony that STATIC-99 "is based on old cases from outside of the United States"); *People v. Rose*, No. H030534, 2007 Cal. App. Unpub. LEXIS 10017, at *14–15 (Dec. 12, 2007) (noting defense expert's testimony that STATIC-99 is mostly based on young adult samples of child molesters in England and Canada); *Thomas*, 2002 Cal. App. Unpub. LEXIS, at *8 (stating defense expert testified that actuarial tools provide group data not individual assessments, and that STATIC-99 is based on highly disturbed group, so should be used cautiously).

369. See *People v. Hernan*, No. B199439, 2008 Cal. App. Unpub. LEXIS 7697, at *7–8 (Sept. 30, 2008) (citing defense expert's testimony that STATIC-99 is based on data from Canada and England, where sexual recidivism is three to four times higher than in United States).

370. See, e.g., *State v. Fowler*, 694 N.W.2d 446, 451 (Wisc. Ct. App. 2005) (stating that defense expert had disputed actuarial scores could be used to predict recidivism).

371. E.g., *People v. Marentez*, No. B206971, 2010 Cal. App. Unpub. LEXIS 8001, at *17 (Oct. 6, 2010); *People v. Edmonton*, 126 Cal. Rptr. 2d 836, 839 (Ct. App. 2002); *People v. Vercolio*, 843 N.E.2d 417, 420 (Ill. App. Ct. 2006).

372. E.g., *In re Commitment of Sandry*, 857 N.E.2d 295, 302 (Ill. App. Ct. 2006); *In re Commitment of Stevens*, 803 N.E.2d 1036, 1043 (Ill. App. Ct. 2004).

373. For examples of subpopulations to which defendants belonged, see *Cooley v. Superior Court*, 57 P.3d 654, 660 (Cal. 2002) (Hispanics); *People v. Flores*, 50 Cal. Rptr. 3d 567, 570 (Ct. App. 2006) (castrated men); *People v. Jackson*, No. A10091, 2004 Cal. App. Unpub. LEXIS 2390, at *11–12 (Mar. 16, 2004) (African Americans); *People v. Wolfenbaerger*, No. C043071, 2003 Cal. App. Unpub. LEXIS 9850, at *12 (Oct. 21, 2003) (California state population).

by one defense expert, STATIC-99 is “confusing, distracting, and intellectually dishonest.”³⁷⁴

In sum, the cases to date indicate that scoring the actuarial factors may not be as objective and simple as desired since differences of opinion in the coding rules have emerged, despite the few factors that are involved. In addition, adversarial allegiance is evident in the divergent paths that state and defense experts have taken in their approach to, and assessment of, actuarial tools. Courts, however, have consistently sided with the state experts’ approach. The overall trend suggests a certain arbitrariness with respect to the influence of actuarial evidence. When the actuarial tool shows a high risk (greater than fifty percent), little additional support has been needed to support an SVP determination. But even if the actuarial score has been less than fifty percent, as low as zero, state experts and courts have still supported sexual-predator classifications, based on non-actuarial evidence. This trend has been consistent despite many defense experts providing specific information undermining the reliability and validity of actuarial assessments.

D. (Mis)Interpretation of Actuarial Prediction

By critically analyzing the appellate opinions to date, it is clear that many experts and courts have erroneously interpreted the actuarial tools and their purposes. This section outlines four types of interpretive problems. The first, which is quite common among the cases, is the improper interpretation that group-based scores provide risk-assessment estimates that are individualized to specific defendants.³⁷⁵ In describing the general usefulness of actuarial results to individual predictions, for example, one opinion stated that the STATIC-99 actuarial instrument “calculated *defendant’s* risk of reoffense,”³⁷⁶ while another court referred to an expert stating that actuarial instruments are “commonly used to assess an individual’s risk of recidivism.”³⁷⁷ Further illustrating this interpretative error, one opinion described actuarial tools as “instruments that

374. *Pritchett v. Hunter*, No. C 05-0764, 2008 U.S. Dist. LEXIS 106899, at *12 (N.D. Cal. Sept. 9, 2008).

375. Defendants have not successfully challenged the actuarial instruments as irrelevant due to their basis in group data rather than individual risk. *See, e.g., In re* care and treatment of Murrell, 215 S.W.3d 96, 112 (Mo. 2007) (en banc) (rejecting irrelevance argument because data is reasonably relied upon by experts in field). However, courts do acknowledge that actuarial scores cannot predict a given individual’s future conduct. *See, e.g., State v. Garner*, No. 89840, 2008 Ohio App. LEXIS 1663, at *20 n.16 (Ct. App. Apr. 24, 2008) (citing precedent that STATIC-99 cannot provide individualized prediction of future offenses).

376. *People v. Hubbart*, 106 Cal. Rptr. 2d 490, 498 (Ct. App. 2001) (emphasis added).

377. *In re* Det. of Taylor, 134 P.3d 254, 257 (Wash. Ct. App. 2006); *see also Murrell*, 215 S.W.3d at 11 (referring to expert claiming actuarial tool was “relevant to the risk of recidivism for the individual as well as the test group”); *People v. Suarez*, No. E042936, 2008 Cal. App. Unpub. LEXIS 9144, at *5 (Oct. 28, 2008) (describing STATIC-99 as measuring “likelihood a person will engage in sexually violent behavior”); *People v. Troglin*, No. H027162, 2005 Cal. App. Unpub. LEXIS 9649, at *3 (Oct. 24, 2005) (describing STATIC-99 as “actuarial tool that estimates an individual’s risk for sexual reoffense”); *People v. Therrian*, 6 Cal. Rptr. 3d 415, 417 (Ct. App. 2004) (“The total score of the [STATIC-99] is a percentage chance of the defendant’s likelihood of being convicted for a future sexual offense.”).

experts use[] to determine whether a defendant qualifies as a sexually violent predator.”³⁷⁸

In addition, in many cases, the expert has directly imputed the actuarial risk statistic from the experience table to the specific defendant.³⁷⁹ Examples include an expert testifying that the defendant’s score of seven on STATIC-99 “means that the likelihood of [the defendant] being convicted of a new sex offense is 39% within 5 years”³⁸⁰ and another expert stating that the results of RRASOR and STATIC-99 “indicated that [the defendant] was likely to reoffend.”³⁸¹ An expert in another recent case contended that STATIC-99 could “diagnose properly” an individual’s risk, a judgment which the court adopted.³⁸² This individualization of actuarial scores is in direct conflict with the instructions of STATIC-99’s developers that the tool cannot be used in this manner.³⁸³

Conveying the experience-table result as a group-based statistic (rather than an individually-tailored one) has also generally favored the state as well. While courts have tended to accept the individualized statistic when the actuarial-based risk estimate is high, they have tended to highlight the group-based nature of actuarial scores when the results are low.³⁸⁴

378. *In re* Commitment of Burton, 884 So. 2d 1112, 1113 (Fla. Dist. Ct. App. 2004) (describing prior trial court proceedings at *Frye* hearing to determine admissibility of testimony on actuarial tools’ results).

379. *See* Helm v. Ahlin, No. 1:08-CV-01648, 2009 U.S. Dist. LEXIS 41572, at *7 (E.D. Cal. May 15, 2009) (summarizing state expert’s testimony that STATIC-99 “score indicates [defendant] has ‘high risk’ of reoffending” (alteration in original)); Cooley v. Superior Court, 57 P.3d 654, 659 (Cal. 2002) (“[S]core of 6 on the [STATIC-99] test put[s] defendant in the ‘high risk’ category for reoffense, with a 52 percent minimum risk of reoffense within 15 years.”); People v. Zinlu, No. A114563, 2007 Cal. App. Unpub. LEXIS 8730, at *5 (Oct. 29, 2007) (relaying expert’s testimony that STATIC-99 score indicated defendant “had a 39 percent risk of reoffense in five years”); People v. Tribble, No. F043573, 2004 Cal. App. Unpub. LEXIS 10133, at *5 (Nov. 5, 2004) (describing expert testimony that STATIC-99 score indicated defendant had “38 percent probability of reconviction over the next 10 years”); Garcetti v. Superior Court, 102 Cal. Rptr. 2d 214, 226 (Ct. App. 2000) (quoting expert as testifying that, based on actuarial tools, defendant “‘comes out over [a] 50 percent’ risk of reoffending” (alteration in original)); *see also* Therrian, 6 Cal. Rptr. 3d at 417 (describing STATIC-99 total score as being “percentage chance of the defendant’s likelihood of being convicted for a future sexual offense”).

380. *State v. P.H.*, 874 N.Y.S.2d 733, 740 (N.Y. Sup. Ct. 2008).

381. *People v. Taylor*, 830 N.E.2d 855, 857–58 (Ill. App. Ct. 2005).

382. *See* *People v. McKee*, 73 Cal. Rptr. 3d 661, 688 n.23, 689 (Ct. App. 2008) (concluding that defendant was not diagnosed properly but relying on expert’s belief that defendant would reoffend).

383. HARRIS ET AL., *supra* note 170, at 71 (“The recidivism estimates provided by the STATIC-99 are group estimates based upon reconvictions and were derived from groups of individuals with these characteristics. As such, these estimates do not directly correspond to the recidivism risk of an individual offender.”).

384. *See In re* Civil Commitment of K.S., No. A-1451-07T2, 2008 N.J. Super. LEXIS 627, at *6–9 (Mar. 11, 2008) (noting “high risk” score on STATIC-99 as part of determination that commitment was proper); *State v. Vanek*, No. 89125, 2007 Ohio App. LEXIS 5433, at *6–7 (Nov. 21, 2007) (finding for state despite low score while noting that “the utility of the Static-99 evaluation as a diagnostic tool for individual risk assessment is open to question” (internal quotation marks omitted)); *State v. Ellison*, No. 78256, 2002 Ohio App. LEXIS 4216, at *5–6, 24 (Aug. 8, 2002) (ruling that court need not give much weight to STATIC-99 evidence showing defendant to be “low-to-medium risk,” since actuarial-based risk estimates could “be at odds with Ohio’s statutory scheme” requiring individualized determinations).

The second interpretive problem is evident in opinions that present erroneous representations of the actuarial tools. Many appellate opinions refer to expert testimony about the actuarial test results that appears inconsistent with, even contradictory to, the actual tests the experts purported to use. One expert, for example, used the meta-analysis that helped form the resulting RRASOR and STATIC-99 tests as if the meta-analysis were itself an actuarial tool, contrary to the developer's specific intention.³⁸⁵ Another state expert used STATIC-99 to support his assessment that the young defendant was likely to reoffend, despite the other state expert's critique that the tool was not appropriate for adolescents.³⁸⁶ Regardless, the court upheld the evidentiary use of the actuarial score to civilly commit the defendant, even though the defendant was fourteen-years-old at the time of his sexual offense and had been incarcerated or in detention ever since, without further charges.³⁸⁷

In another case, the opinion erroneously indicated that STATIC-99 was a tool to diagnose psychiatric disorders, explaining that STATIC-99 "measures both paraphilia and antisocial personality disorder."³⁸⁸ Additionally, some experts have created new hierarchies, asserting that STATIC-99 placed the defendant in a category of "very high risk"³⁸⁹ or "extremely high" risk,³⁹⁰ despite the top STATIC-99 ordinal grouping of "high risk."³⁹¹ Another expert has stated that, according to STATIC-99, the defendant was "in a class almost by himself."³⁹² Other misrepresentations have involved the specifics of how the actuarial tests were developed,³⁹³ scoring methodology,³⁹⁴ the

385. See *In re* Commitment of Field, 813 N.E.2d 319, 322–24 (Ill. App. Ct. 2004). The expert stated that he used the Hanson & Bussiere meta-analysis, which he described as itself an actuarial instrument, and concluded that the defendant's future dangerousness was indicated by having seven of the twenty-two identified correlates in the meta-analysis. *Id.* However, the meta-analysis he cited was not intended as an instrument and, indeed, the authors themselves describe the work as "a quantitative review of the sexual offender recidivism literature." R. Karl Hanson & Monique T. Bussiere, *Predicting Relapse: A Meta-Analysis of Sexual Offender Recidivism Studies*, 66 J. CONSULTING & CLINICAL PSYCHOL. 348, 349 (1998). While Hanson and Bussiere discuss the various (far more than twenty-two) correlates with sexual recidivism, they "do not recommend simply summing the items." *Id.* at 358.

386. *In re* Commitment of Sandry, 857 N.E.2d 295, 300, 302 (Ill. App. Ct. 2006); see also *B.W. v. State*, 909 N.E.2d 471, 474–75 n.6 (Ind. Ct. App. 2009) (upholding registration requirement with expert basing risk assessment of juvenile in part on STATIC-99).

387. *Sandry*, 857 N.E.2d at 297–98, 316.

388. *In re* Anderson, 730 N.W.2d 570, 582 (N.D. 2007).

389. *In re* Commitment of Stevens, 803 N.E.2d 1036, 1042 (Ill. App. Ct. 2004); *In re* Commitment of Simons, 821 N.E.2d 1184, 1187 (Ill. 2004) (internal quotation marks omitted) (attributing such categorization to STATIC-99's creators).

390. *In re* Det. of Lieberman, 884 N.E.2d 160, 167 (Ill. App. Ct. 2007) (internal quotation marks omitted).

391. HARRIS ET AL., *supra* note 170, at 3. In addition to using language that is more severe than that provided by STATIC-99 Coding Rules, courts have also miscategorized defendants altogether. In one case, the expert stated that the STATIC-99 score of four placed the defendant in the "high risk" category, despite STATIC-99's classification of four as "medium-high." *In re* Commitment of Fisher, 164 S.W.3d 637, 643 (Tex. 2005).

392. *People v. Therrian*, 6 Cal. Rptr. 3d 415, 417 (Ct. App. 2003) (internal quotation marks omitted).

393. Experts have represented the development process for different actuarial tests in widely divergent manners. See *Murrell v. State*, 215 S.W.3d 96, 108–09 (Mo. 2007) (en banc) (recounting expert's testimony that STATIC-99 was "developed by looking at the characteristics of approximately 4,000 sex offenders to see which characteristics they possessed were associated with the likelihood or reoffense within 15 years, as

tools' definition of recidivism,³⁹⁵ and what type of recidivism a specific actuarial tool was designed to predict.³⁹⁶ While none of these, on their own, may be hugely important to the resulting SVP-status decisions, they cause one to pause in considering the quality of the training, experience, and care experts have taken with respect to their actuarial assessments. These examples are also helpful in that the opinions lack any indication that such errors had been revealed by *Daubert's* and *Barefoot's* adversarial check of cross-examination or the appellate review process.

The third interpretive issue involves scoring discrepancies. Despite the purported ease of RRASOR and STATIC-99 scoring, several cases document state experts changing their scores of the defendant.³⁹⁷ In addition, there have been discrepancies in the scoring of defendants, both between experts on competing sides³⁹⁸ and between the two or more experts on the same side.³⁹⁹ The courts in these cases have generally

defined by reconviction"); *People v. Flores*, 50 Cal. Rptr. 3d 567, 570 (Ct. App. 2006) (stating that STATIC-99 is "based on data from over 31,000 sex offenders"); *People v. Torres*, 119 Cal. Rptr. 2d 597, 600 n.4 (Ct. App. 2002) (indicating that expert distinguished RRASOR and STATIC-99 by declaring that RRASOR was not developed using actual recidivism statistics, but STATIC-99 was); *Simons*, 821 N.E.2d at 1187 (discussing expert testimony and stating STATIC-99 is "based upon a study of thousands of sex offenders from England, Canada, and the United States"); *see also* *United States v. Shields*, 597 F. Supp. 2d 224, 237 (D. Mass. 2009) (relying on expert's testimony that RRASOR's five-year rate of 49.8% was observed rate, while ten-year rate was extrapolated).

394. *See* *Cooley v. Superior Court*, 57 P.3d 654, 659 (Cal. 2002) (stating expert testified that a factor of STATIC-99 was "whether the most recent sex offense involved the use of violence"); *People v. Elam*, No. C039689, 2003 Cal. App. Unpub. LEXIS 5621, at *8 (June 10, 2003) (stating expert testified that defendant's score was "6 out of 6" on STATIC-99); *see also In re Commitment of R.S.*, 773 A.2d 72, 79 (N.J. Super. Ct. App. Div. 2001) (relating expert testimony that STATIC-99 includes substance abuse as factor).

395. *See Therrian*, 6 Cal. Rptr. 3d at 417 (indicating that STATIC-99's experience table is based on convictions for new sexual offenses); *State v. Hornack*, No. 81021, 2003 Ohio App. LEXIS 415, at *8 (Jan. 30, 2003) (referring to state expert's statement that STATIC-99 experience tables do not include charges).

396. *See Black v. Voss*, 557 F. Supp. 2d 1100, 1106 (C.D. Cal. 2008) (describing defense expert's testimony that she used RRASOR but not STATIC-99 because latter measures risk for criminal offending and not sexual deviance).

397. *See People v. Allen*, 80 Cal. Rptr. 3d 183, 193 (Cal. 2008) (describing state expert who originally assigned score of six on STATIC-99, but later raised score to eight after receiving additional information); *Shelton v. Commonwealth*, 645 S.E.2d 914, 915 (Va. 2007) (describing state expert who, while initially giving defendant a RRASOR score of five, testified at probable cause hearing that it should be two, and later adjusted score to three at trial); *Miles v. Commonwealth*, 634 S.E.2d 330, 334 (Va. 2006) (finding state conceded defendant incorrectly received RRASOR score of four, an additional point that was not warranted by number of prior convictions, resulting in defendant's release from civil commitment as statute specifically required minimum RRASOR score before initiating proceedings); *People v. Clotfelter*, No. C038351, 2002 Cal. App. Unpub. LEXIS 8921, at *24-25 (Sept. 25, 2002) (noting that, although state expert had previously scored defendant with a three on STATIC-99, she later realized while preparing for trial that she could have added two more points due to evidence of prior sex offenses, thus increasing recidivism risk from thirteen to forty percent); *In re Commitment of J.P.*, 772 A.2d 54, 59 (N.J. Super. Ct. App. Div. 2001) (referring to state expert who admitted scoring STATIC-99 incorrectly by double counting offense).

398. *See United States v. Abregana*, 574 F. Supp. 2d 1145, 1155 (D. Haw. 2008) (discussing two point gap in state and defense experts' STATIC-99 scores and one point gap in RRASOR scores).

399. *See People v. Stewart*, No. C037873, 2003 Cal. App. Unpub. LEXIS 4207, at *6-7 (Apr. 29, 2003) (noting that different state experts gave RRASOR scores of three and five and STATIC-99 scores of six and seven to same defendant); *Therrian*, 6 Cal. Rptr. 3d at 417 (explaining that two state experts gave defendant different scores using STATIC-99); *Clotfelter*, 2002 Cal. App. Unpub. LEXIS 8921, at *24-25, 28-29 (noting states experts' disagreement on which offense to classify as "index offense"); *In re Commitment of Combs*,

treated these discrepancies as issues of fact rather than as problems inherent in the instruments and their scoring rules.

The fourth interpretive problem involves how the experts convey the significance of using multiple actuarial tools. Here, a common empirical flaw has been to characterize the use of multiple actuarial tools as somehow strengthening the reliability of scores from the individual tools, at least when they are relatively consistent in the direction of their risk predictions. For example, in *People v. Calderon*,⁴⁰⁰ the court stated that the expert “confirm[ed] the accuracy” of STATIC-99 results by comparing them with those obtained from RRASOR and noted that each tool “relies on a different basis of prediction.”⁴⁰¹ Other experts have also based their predictions of future dangerousness on the reinforcing RRASOR and STATIC-99 scores, with no mention in the appellate opinions of the tests’ overlapping nature.⁴⁰² Issues with intercorrelation between variables makes this problematic as RRASOR and STATIC-99 share several factors⁴⁰³ derived from the same offender-recidivism literature by the same developers.⁴⁰⁴ An empirical study found that the practice of using multiple tools does not result in a statistically significant better prediction than would be provided by using the single best actuarial risk scale.⁴⁰⁵ On the other hand, a study found that less than five percent of its sample received consistent rankings of high risk or low risk using each of five actuarial tools, including RRASOR and STATIC-99.⁴⁰⁶

Despite the multiple criticisms of actuarial risk assessment as conveyed in this Part, defendants have been unable to prevail on ineffective assistance of counsel claims when their lawyers fail to properly challenge the admissibility, validity, and interpretation of the actuarial tools.⁴⁰⁷ Criticisms of actuarial risk assessments, however, including the vagueness and variability of the likelihood standard, adversarial allegiance, and the frequent misinterpretations of the actuarial tests merit further

720 N.W.2d 684, 687, 689 (Wis. Ct. App. 2006) (identifying situation where four state experts gave same defendant RRASOR scores of two, three, and four; and STATIC-99 scores from four to eight, with one expert giving a “six slash seven slash eight” (internal quotation marks omitted)); *see also In re Commitment of Simons*, 821 N.E.2d 1184, 1188 (Ill. 2004) (observing that one state expert scored defendant “a little bit higher” than other state expert (internal quotation marks omitted)).

400. No. B206734, 2009 Cal. App. Unpub. LEXIS 1427 (Feb. 23, 2009).

401. *Calderon*, 2009 Cal. App. Unpub. LEXIS 1427, at *14.

402. *E.g.*, *People v. Paredes*, No. B200935, 2009 Cal. App. LEXIS 3274, at *3 (Apr. 28, 2009); *People v. Lopez*, 53 Cal. Rptr. 3d 549, 553 (Ct. App. 2006); *People v. Edmonton*, 126 Cal. Rptr. 2d 836, 838–39 (Ct. App. 2002).

403. *See Barbaree et al.*, *supra* note 220, at 500 (noting that the ten items measured by STATIC-99 include the four used in RRASOR); HARRIS ET AL., *supra* note 170, at 3 (noting that STATIC-99 was created by merging RRASOR and another tool).

404. Michael C. Seto, *Is More Better? Combining Actuarial Risk Scales to Predict Recidivism Among Adult Sex Offenders*, 17 PSYCHOL. ASSESSMENT 156, 156 (2005); *see also HANSON*, *supra* note 158, at 2–4 (describing statistical studies that RRASOR relied upon in developing guidelines for actuarial factors).

405. Seto, *supra* note 404, at 165.

406. Howard E. Barbaree et al., *Different Actuarial Risk Measures Produce Different Risk Rankings for Sexual Offenders*, 18 SEXUAL ABUSE 423, 431 (2006).

407. *See, e.g.*, *Sigman v. Rogers*, No. 07-1383(DMC), 2008 U.S. Dist. LEXIS 71127, at *43–45 (D.N.J. Sept. 3, 2008) (finding plaintiff’s argument “nothing more than Monday-morning quarter-backing” because based on difference of opinion over actuarial tools’ reliability); *People v. Langhorne*, No. H027495, 2005 Cal. App. Unpub. LEXIS 11150, at *48 (Dec. 1, 2005).

attention. The disconnect between legal language and the scientific judgments by mental health professionals does a disservice to the interests of justice and underscores the inconsistency of the legal standards among cases.

VI. CONCLUSION

In order to balance the goals of protecting the public while adhering to the constitutional principles of liberty and privacy, legal actors in the criminal justice process have an ethical duty to critically evaluate scientific testimony that is offered as expert evidence. Some practitioners and academics with joint law and doctoral credentials have argued for a ban on actuarial risk assessment as unreliable science in SVP proceedings.⁴⁰⁸ This author supports such a ban, not only because of the significant problems with using group-based actuarial tools to predict individual behavior, but also because of the absence of legal criteria to guide the use of such tools under current SVP laws.

There are a number of reasons why legal personnel may be loath to challenge mental health experts' evidence of future dangerousness. First, it could be the natural "desire for authoritative methods for generating knowledge."⁴⁰⁹ The aura of objectivity and science serves this positivist aim.⁴¹⁰ Second, judges may be relying upon the adversarial process, pursuant to *Daubert* and *Barefoot*, to weed out bad science. By doing so, however, judges are eschewing their gatekeeping obligations onto attorneys, who may also be ill-equipped in terms of scientific literacy.⁴¹¹ The discussion of existing case law in Part V emphasizes the failure of both the adversarial process and judicial gatekeeping in challenging this suspect science. Third, SVP laws are an area of law which invokes passionate responses and moral judgments.⁴¹² And, finally, some argue that the legislatures created sexual predator laws with the future dangerousness concept and, thus, we (mental health experts, judges, and lawyers) need to use the best available evidence to make those decision—even if the legal standards remain vague, and even though current models of actuarial risk assessment suffer large gaps in validity and reliability.⁴¹³ This reasoning, however, exalts the political over the legal.

408. See Fennel, *supra* note 58, at 61 (calling STATIC-99 unreliable and saying that laws relying on actuarial tools ask science to "perform a task it cannot reliably perform").

409. Jennifer L. Mnookin, *Idealizing Science and Demonizing Experts: An Intellectual History of Expert Evidence*, 52 VILL. L. REV. 763, 793 (2007) (capitalization omitted).

410. See D. Michael Risinger, *The Irrelevance, and Central Relevance, of the Boundary Between Science and Non-Science in the Evaluation of Expert Witness Reliability*, 52 VILL. L. REV. 679, 683 (2007) (describing "heroic positivist" view that "when properly applied, [science] allows a practitioner to discover new details of the contours of . . . external reality").

411. See George, *supra* note 119, at 252 (arguing that this places opposing attorney in position of either having to challenge admissibility of expert's opinion based on reliability grounds or submit her own expert and then challenge weight of evidence offered by original expert).

412. See *infra* Part II for a discussion of the "moral panic" that has given rise to current SVP laws.

413. See Janus & Prentky, *supra* note 40, at 1445–46 (arguing for widespread use of actuarial risk assessment despite uncertainty and other concerns surrounding their accuracy and impact). *But see In re Commitment of Burton*, 884 So. 2d 1112, 1120 (Fla. Dist. Ct. App. 2004) (Altenbernd, C.J., concurring) (referring to high rate of "false positives" from actuarial tests and arguing that with regard to sexual predators, "[w]e have embarked on the first steps into a new world, arguably a science fiction world, in which judges and juries are asked to prevent crimes years before they occur" (internal quotation marks omitted)).

For purposes of my point, assume there is empirical evidence that an astrology-based model of risk assessment is shown to be better than chance. Then assume that the astrology model is empirically shown to be a better predictor than a psychic model of risk assessment. Would these facts justify the acceptance of the astrology-based risk tool as scientifically valid? And, would this be enough to justify its use as evidence in judicial proceedings in which an individual's liberty and privacy is at stake?⁴¹⁴

Advocating for a new sex offender policy is beyond the scope of this Article. Still, considering the myths about the recidivism risk of sex offenders that led to the rise of SVP laws, there is a strong basis for simply getting back to the basics of transparent sentencing policy, if the intent is to rely directly on punitive incarceration. There is also potential to engage the rehabilitative possibilities offered by much work done in recent years in sexual-offender treatment.⁴¹⁵

In the end, officials will likely continue to believe that political and public interests are best served by the current SVP regime's reliance on actuarial risk assessment tools.⁴¹⁶ The naturalistic fallacy (i.e., confusing what is with what one thinks ought to be), in terms of having an objective ability to predict risk, is not an uncommon tendency. Historical juridical authority, though, should not be entirely ceded in protecting what constitutes expert evidence in law. The justice system must use the *Daubert* and *Frye* standards to critically analyze predictions of future dangerousness in the SVP context. Actuarial assessments of risk are couched in terms of science and objectivity and thus should be evaluated regularly for their reliability and validity.

Education is key for those involved in SVP-status determinations, not only as to the ability of actuarial science to accurately predict future dangerousness, but also how best to challenge actuarial assessments in court to ensure that only those tools that are sufficiently reliable and valid are admissible.⁴¹⁷ The ability of legal professionals to better understand and critically analyze science is a matter of additional learning; guidance is available.⁴¹⁸ If this Article is not convincing on the suspect nature of the

414. Cf. Boer, *supra* note 246, at 1 (arguing that there is more concern in literature about which method of risk assessment is preferred than ethics of, and whether public safety is actually served by, mental health attempts at risk assessment); JANUS, *supra* note 3, at 6–7 (maintaining that actuarial assessment debate improperly focuses on small number of sex offenders who are at risk of being predators instead of focusing on larger patterns of sexual violence and more prevalent non-predator sex crimes, like acquaintance rape).

415. See NATHAN JAMES ET AL., CIVIL COMMITMENT OF SEXUALLY DANGEROUS PERSONS 32–35 (2008) (describing several states' experiments with outpatient civil commitment and other innovative monitoring programs); Molly T. Geissenhainer, *The \$62 Million Question: Is Virginia's New Center to House Sexually Violent Predators Money Well Spent?*, 42 U. RICH. L. REV. 1301, 1328–35 (2008) (discussing Virginia's civil commitment program and possible therapeutic alternatives); Lester, *supra* note 11, at 372–73 (arguing that residential and employment restrictions placed on sex offenders after punishment for their crimes are unjust and should be replaced by rehabilitation programs).

416. See Janus & Prentky, *supra* note 40, at 1459 (contending that legislatures will continue to mandate risk assessments in sexual predator laws and judges will “undoubtedly continue to oblige”).

417. See ALAN D. GOLD, EXPERT EVIDENCE IN CRIMINAL LAW: THE SCIENTIFIC APPROACH 17 (2003) (contending that frequency of junk science being offered as credible in criminal cases means that judges and lawyers must gain scientific literacy to distinguish good science from bad).

418. Post-*Daubert*, others have written general guides to educate judges and lawyers on matters of science and statistics. FED. JUDICIAL CTR., REFERENCE MANUAL ON SCIENTIFIC EVIDENCE (2d ed. 2000)

science of future-dangerousness evidence, at the very least it provides some assistance for digging deeper into the empirical nature of actuarial tests. Constitutional principles and evidentiary rules require adherence to rationality despite the mythical specter of a growing population of dangerous sexual predators.

(including chapters on statistics and multiple regression); *see also* BEECHER-MONAS, *supra* note 241, at 155–67 (explaining science and error rate behind predictions of sexual offenders’ dangerousness). Another paper offers specific questions for ascertaining the level of expertise the offered expert has in the subject at issue. David M. Godden & Douglas Walton, *Argument from Expert Opinion as Legal Evidence: Critical Questions and Admissibility Criteria of Expert Testimony in the American Legal System*, 19 *RATIO JURIS* 261, 277–79 (2006).