
Note

Up or Out: Why “Sufficiently Reliable” Statistical Risk Assessment Is Appropriate at Sentencing and Inappropriate at Parole

*Pari McGarraugh**

Imagine you are sitting before a parole board. Now imagine that your parole hearing is one of over 8000 the board will hear this year—more than thirty cases each working day¹—and that you have no lawyer to advocate for your cause.² You know that state law requires the parole board to consider the results of a statistical risk assessment to determine your chances of recidivism and thus the appropriateness of your release.³ Given the rushed nature of the proceeding, it occurs to you that the board is likely to weigh these results heavily.⁴ You ask for a copy of the risk assessment instrument or even a copy of your risk report and the parole board refuses.⁵ You start to wonder: What

* J.D. and M.P.H. Candidate, University of Minnesota. Thanks to Professor Kevin Reitz for his insight and guidance. Thanks also to Michel Boudreaux for sharing his expertise in statistics. Endless thanks to my husband, Jordan Miller, for his unwavering support, shockingly convincing dinosaur impressions, and commitment to making fun a way of life—without him, I'd be a complete bore. Copyright © 2013 by Pari McGarraugh.

1. SUSAN C. KINNEVY & JOEL M. CAPLAN, CTR. FOR RESEARCH ON YOUTH & SOC. POLICY, FINDINGS FROM THE APAI INTERNATIONAL SURVEY OF RELEASING AUTHORITIES 9 (2008), available at http://paroleboard.arkansas.gov/Resources/Documents/Publications/2008APAI_Survey.pdf.

2. See *Holup v. Gates*, 544 F.2d 82, 85 (2d Cir. 1976) (holding that inmates do not have a constitutional right to have counsel or other advocates attend parole hearings).

3. *E.g.*, COLO. REV. STAT. § 17-22.5-404(4)(a)(II) (2012); CONN. GEN. STAT. § 54-124a(m) (2009).

4. *Cf.* Vincent O'Leary & Daniel Glaser, *The Assessment of Risk in Parole Decision Making*, in THE FUTURE OF PAROLE: COMMENTARIES ON SYSTEMS IN BRITAIN AND U.S.A. 135, 136 (D.J. West ed., 1972) (finding that 79.4% of parole board members consider risk of recidivism if released to be a consideration which “significantly influenced their decision”).

5. Some states treat documents examined by the parole board as confidential and refuse access even to the inmate whose release is under consideration. *E.g.*, GA. CODE ANN. § 42-9-53 (2012).

factors is that risk assessment considering? Exactly what do the results mean? And most importantly, can a statistical model really predict the future with enough reliability that it ought to seal anyone's fate?

A statistical risk assessment instrument provides a structured way for the criminal justice system to evaluate the probability that an individual will reoffend.⁶ Such an instrument uses the statistical relationships between individual risk factors and criminal behavior to develop "explicit rules" that assign weights to risk factors and combine the weights to produce "an objective estimate of violence risk."⁷ A wide range of instruments are now in use, which accordingly have variable power to predict future dangerousness.⁸ States have been using risk assessment with increasing frequency and are unlikely to halt their efforts on this front.⁹ Because the continued and accelerated use of risk assessment instruments in criminal justice is inevitable, the relevant question is not if statistical risk assessment ought to be used, but how it can be responsibly used.

Presently, both sentencing judges and parole boards in states using indeterminate sentencing models, those jurisdictions that allow the parole board discretionary release authority,¹⁰ routinely use statistical risk assessment instruments when evaluating offenders.¹¹ Inmates seeking parole release face the board and statistical risk assessment largely without legal pro-

6. John Monahan, *A Jurisprudence of Risk Assessment: Forecasting Harm Among Prisoners, Predators, and Patients*, 92 VA. L. REV. 391, 405 (2006).

7. *Id.* at 405–06.

8. Paul Gendreau et al., *A Meta-Analysis of the Predictors of Adult Offender Recidivism: What Works!*, 34 CRIMINOLOGY 575, 590 (1996).

9. This increasing emphasis on empirically-measured risk is apparent from policy players with a wide range of political affiliations. *See, e.g.*, AM. CIVIL LIBERTIES UNION, SMART REFORM IS POSSIBLE: STATES REDUCING INCARCERATION RATES AND COSTS WHILE PROTECTING COMMUNITIES 9 (2011), available at <http://www.aclu.org/files/assets/smartreformispossible.pdf> (arguing that states should "[r]equire [e]vidence-[b]ased [c]riminal [j]ustice [p]ractices and [r]isk [a]ssessment [i]nstruments"); *Priority Issues: Adult Probation*, RIGHT ON CRIME, <http://www.rightoncrime.com/priority-issues/adult-probation/> (last visited Nov. 27, 2012) ("Probation can be made particularly efficient through the use of risk assessments . . .").

10. *See* Steven L. Chanenson, *The Next Era of Sentencing Reform*, 54 EMORY L.J. 377, 382–83 (2005).

11. KINNEVY & CAPLAN, *supra* note 1, at 12 (describing risk assessment at parole); Roger K. Warren, *Evidence-Based Sentencing: Are We Up to the Task?*, 23 FED. SENT'G REP. 153, 157 (2010) (reporting that in at least ten states, sentencing judges use risk assessment instruments).

tection¹²—they are not typically represented by counsel,¹³ the rules of evidence do not apply,¹⁴ and only the barest procedure is required to satisfy due process.¹⁵ By contrast, more robust protections are available to defendants post-conviction at sentencing. These protections include, critically, the right to counsel¹⁶ and the requirement that evidence introduced must meet a minimum standard of reliability.¹⁷

This Note argues in support of a recent draft of section 6B.09 of the Model Penal Code (MPC) which suggests limiting the use of statistical risk assessment to the sentencing stage.¹⁸ Because there is meaningful protection for criminal defendants at sentencing that is not present at parole,¹⁹ using statistical risk assessment at sentencing polices a basic line of fairness to criminal defendants. This Note expands upon the reasoning of the MPC by examining, in depth, the statistical and legal weaknesses of risk assessment instruments. Part I sets forth the basic protections available to defendants and inmates in indeterminate sentencing jurisdictions at parole and sentencing and describes the panoply of statistical risk assessment instruments and their use. Part II explores the practical and legal limitations of these risk assessment instruments. Part III articulates the reasons why section 6B.09 is a reasonable compromise between fairness to defendants and protecting public safety and argues for an addition to the section establishing statutory criteria for determining if a risk assessment instrument is “sufficiently reliable.”

12. Not all states stick to the bare constitutional minimum in parole proceedings. Some states provide for greater protection through statute. *E.g.*, HAW. REV. STAT. § 706-670(3)(b) (2011) (allowing inmates to be represented in parole hearings).

13. N.M. CODE R. § 22.510.2.8(A)(3) (2001) (prohibiting legal counsel from attending parole hearings); *Holup v. Gates*, 544 F.2d 82, 85 (2d Cir. 1976).

14. *See Davis v. Brown*, 311 F. Supp. 2d 110, 115 (D.D.C. 2004) (noting that the parole board may rely on hearsay evidence).

15. *See Greenholtz v. Inmates of the Neb. Penal & Corr. Complex*, 442 U.S. 1, 16 (1979).

16. *Gardener v. Florida*, 430 U.S. 349, 358 (1977).

17. *E.g.*, *People v. Johnson*, 499 N.E.2d 1355, 1371 (Ill. 1986).

18. MODEL PENAL CODE: SENTENCING § 6B.09 (Tentative Draft No. 2, 2011).

19. *Compare, e.g.*, *Gardener v. Florida*, 430 U.S. 349, 358 (1977) (right to counsel at sentencing), *with Holup v. Gates*, 544 F.2d 82, 85 (2d Cir. 1976) (no counsel at parole hearing).

I. PROCEDURAL PROTECTION AND RISK ASSESSMENT INSTRUMENTS

In indeterminate sentencing jurisdictions,²⁰ the exact length of a defendant's sentence is not known at the time of sentencing because the judge only designates a minimum and maximum sentence.²¹ Once the inmate has served the minimum term, a parole board determines the actual length of a sentence because the board decides when parole release is appropriate for each offender.²² Together, the parole board and the sentencing judge share responsibility for determining the length of an offender's sentence.²³ This sharing of decision-making power is designed to ensure that "no man be imprisoned unless it is clear that his freedom is dangerous to others, and that when once imprisoned, no man be freed until the danger has ceased."²⁴ Because a defendant has a more substantial liberty interest in the character of the sentencing procedure than an inmate has in the procedures for making parole release decisions, the Constitution demands distinctly different protections at each stage.²⁵ Despite these differing levels of protection, both sentencing judges²⁶ and parole release boards²⁷ are likely to use statistical risk assessment to guide their decisions.

20. By Professor Kevin Reitz's count, over half of the states have sentencing and parole systems that incorporate indeterminate decision-making mechanisms. Kevin R. Reitz, *The "Traditional" Indeterminate Sentencing Model*, in THE OXFORD HANDBOOK OF SENTENCING AND CORRECTIONS 270, 289 n.1 (Joan Petersilia ed., 2012) [hereinafter Reitz, *Indeterminate Sentencing Model*].

21. LYNN S. BRANHAM, THE LAW AND POLICY OF SENTENCING AND CORRECTIONS IN A NUTSHELL 93 (8th ed. 2010).

22. DEAN J. CHAMPION, PROBATION, PAROLE AND COMMUNITY CORRECTIONS 241 (2d ed. 1996); Chanenson, *supra* note 10, at 382–83.

23. See Kevin R. Reitz, *Sentencing*, in CRIME AND PUBLIC POLICY 467 (James Q. Wilson & Joan Petersilia eds., 2011) (noting the multiple decision-makers involved in sentencing).

24. THE TWENTIETH CENTURY FUND TASK FORCE ON CRIMINAL SENTENCING, FAIR AND CERTAIN PUNISHMENT 94 (1976) (quoting Charlton T. Lewis).

25. Reitz, *Indeterminate Sentencing Model*, *supra* note 20, at 283.

26. NAT'L CTR. FOR STATE COURTS & VA. CRIMINAL SENTENCING COMM'N, OFFENDER RISK ASSESSMENT IN VIRGINIA: A THREE-STAGE EVALUATION 11–12 (2002).

27. JAMES AUSTIN ET AL., INST. ON CRIME, JUSTICE, AND CORR. AT GEORGE WASH. UNIV., RELIABILITY AND VALIDITY STUDY OF THE LSI-R RISK ASSESSMENT INSTRUMENT 5 (2003) ("Once an inmate becomes eligible for parole, the decision to release that inmate is guided by an assessment of the inmate's risk to public safety."); KINNEVY & CAPLAN, *supra* note 1, at 12 (finding 88% of releasing authorities nationally use risk assessment instruments).

A. PROCEDURAL PROTECTION AT SENTENCING AND AT PAROLE

Individuals are certainly interested in both sentencing decisions and parole release decisions because both affect the amount of time an individual will spend incarcerated. But the Supreme Court has made clear that inmates have much less legal interest in parole decisions than criminal defendants have in sentencing decisions.²⁸ At sentencing, the defendant has a constitutional liberty interest and corresponding due process right in the character of the procedure employed.²⁹ At parole, courts typically have found that inmates' liberty interest in parole release is constrained by the language of the state statute governing parole.³⁰ Not all state parole systems create liberty interests in parole release.³¹ To give rise to a liberty interest, the statute must do more than establish a system for granting parole release.³² It must create an entitlement to such release, which would lead to a grievous loss if parole release were denied.³³ A statute might create an entitlement and the corresponding grievous loss by, for example, providing that an inmate "shall" be released if certain criteria are met,³⁴ or by substantially limiting the parole board's discretion in making release decisions.³⁵ Based on the line of cases establishing this rule, state legislatures have deliberately phrased parole statutes so that they do not create liberty interests.³⁶ Accordingly,

28. Compare *Scott v. Illinois*, 440 U.S. 367, 373–74 (1979) (finding that the Constitution requires the state to provide counsel to an indigent defendant facing potential incarceration), with *Greenholtz v. Inmates of the Neb. Penal & Corr. Complex*, 442 U.S. 1, 9 (1979) ("There is a crucial distinction between being deprived of a liberty one has . . . and being denied a conditional liberty that one desires.").

29. *E.g.*, *Estelle v. Smith*, 451 U.S. 454, 471 (1981) (analyzing a defendant's Sixth Amendment right to counsel); *Williams v. Currie*, 103 F. Supp. 2d 858, 868 (M.D.N.C. 2000) (finding that imposing longer sentences on men than on women, without a justifying government interest, violates the Equal Protection Clause).

30. *E.g.*, *Bd. of Pardons v. Allen*, 482 U.S. 369, 372 (1987).

31. *E.g.*, *id.* (noting that Montana does not provide a liberty interest in parole).

32. See *Sandin v. Conner*, 515 U.S. 472, 480–81, 483–84 (1995).

33. *Id.*

34. *E.g.*, *Allen*, 482 U.S. at 377–78.

35. *E.g.*, *Ky. Dep't of Corr. v. Thompson*, 490 U.S. 454, 462 (1989), *partially overruled on other grounds by Sandin*, 515 U.S. at 484 n.5.

36. See, *e.g.*, *Worden v. Mont. Bd. of Pardons & Parole*, 962 P.2d 1157, 1165 (Mont. 1998) ("In response to *Allen*, the Montana legislature amended [the parole statute] to state that 'the Board *may* release' inmates on parole . . ." (citing MONT. CODE ANN. § 46-23-201 (1989))).

states typically afford inmates considerably less procedural protection at parole than at sentencing.³⁷

The courts have not examined the outside limits on the types of procedures a parole board may or may not use for two reasons. First, inmates do not challenge the denial of parole as often as defendants challenge sentences because “meaningful review . . . is lacking in virtually all American parole systems.”³⁸ Second, because state law frequently prohibits inmates from examining their own dossiers³⁹ and parole boards are not typically required to articulate their reasons for denying parole,⁴⁰ an inmate may not have enough information about the parole board’s decision to object to it.⁴¹ In a similar vein, the parole process itself is often opaque—to inmates, to judges, and to the general public.⁴² Even though states typically lay out parole processes in statute,⁴³ it may be difficult to verify if parole boards are following statutory procedures because state law may not require the parole board to produce a record of its proceedings.⁴⁴ The lack of both judicial guidance and transparency means that an inmate’s rights at parole are not as well-defined as a defendant’s rights are at sentencing.⁴⁵ Despite the uncertainty regarding the precise limits of required procedural protection, several things are clear: the right to counsel, evidentiary standards, and the qualifications of the decision-maker are distinctly different at sentencing and parole.

37. Reitz, *Indeterminate Sentencing Model*, *supra* note 20, at 283 (noting that if the procedural protections at sentencing are “second-string,” the “procedural accoutrements of parole release are of the *third-* or *fourth-*string variety”).

38. *Id.* at 285.

39. *E.g.*, KY. REV. STAT. ANN. § 439.510 (LexisNexis 2010).

40. *See* *Glover v. Mich. Parole Bd.*, 596 N.W.2d 598, 603–04 (Mich. 1999).

41. The American Law Institute argues for placing risk assessment at sentencing because “substantive concerns . . . will no doubt be brought forward in the courtroom setting—issues that were never raised in the low-visibility, low-process forums of parole release.” MODEL PENAL CODE: SENTENCING § 6B.09 cmt. a (Tentative Draft No. 2, 2011).

42. *See, e.g.*, Michael A. Wolff, *Evidence-Based Judicial Discretion: Promoting Public Safety Through State Sentencing Reform*, 83 N.Y.U. L. REV. 1389, 1404 (2008) (noting that sentencing judges are consistently ignorant of parole procedures and practices).

43. *E.g.*, S.D. CODIFIED LAWS §§ 24-15-1 to 30 (2012).

44. Reitz, *Indeterminate Sentencing Model*, *supra* note 20, at 285.

45. *See id.*

1. The Right to Counsel

The distinction between the right to counsel at the sentencing stage and at the parole stage is straightforward. At sentencing, the defendant facing potential imprisonment has a Sixth Amendment right to the assistance of an attorney.⁴⁶ Moreover, an indigent defendant is entitled to appointed counsel at sentencing.⁴⁷ By contrast, there is no constitutional right to counsel at parole hearings.⁴⁸ This difference appears to be premised on the fundamentally different liberty interests involved at each stage.⁴⁹

At sentencing, courts emphasize that even though the convicted defendant does not have a legal interest in what sentence the court imposes, the defendant does have an interest in the quality of the decision-making process.⁵⁰ Representation at this “critical stage of the proceedings”⁵¹ ensures that defendants take the appropriate steps to preserve their right to appeal sentencing decisions.⁵² Further, a vigorous defense at sentencing evens the playing field in a way that is “essential to the truth-seeking function . . . which may influence the sentencing decision.”⁵³ A defense lawyer is responsible for ensuring that the “sentence [is] not predicated on misinformation.”⁵⁴ In this way, the right to counsel goes hand in hand with the application of evidentiary standards, discussed below.⁵⁵

46. *Mempa v. Rhay*, 389 U.S. 128, 137 (1967).

47. *Scott v. Illinois*, 440 U.S. 367, 373–74 (1979).

48. *See Franciosi v. Mich. Parole Bd.*, 604 N.W.2d 675, 676 (Mich. 2000) (finding that, without a discussion constitutional rights, the state statute did not permit lawyers to “act as a legal representative of a prisoner during a parole interview”).

49. *Compare Rhay*, 389 U.S. at 137 (framing the right to counsel in constitutional terms), *with Conn. Bd. of Pardons v. Dumschat*, 452 U.S. 458, 466 (1981) (framing the liberty interest at parole in relation to the procedural protections offered by the state statute).

50. *E.g.*, *Perdue v. Commonwealth*, 82 S.W.3d 909, 913 (Ky. 2002).

51. *Lavallee v. Justices in the Hampden Superior Court*, 812 N.E.2d 895, 903 n.13 (Mass. 2004).

52. *Rhay*, 389 U.S. at 135 (finding that legal representation is important because “certain legal rights may be lost if not exercised at this stage”).

53. *Gardener v. Florida*, 430 U.S. 349, 360 (1977).

54. *Townsend v. Burke*, 334 U.S. 736, 741 (1948).

55. *See id.* (connecting the right to counsel at sentencing with protection against receiving a sentence based on inaccurate evidence).

2. Evidentiary Standards

Although the rules of evidence applicable at the guilt stage of a trial do not attach in either sentencing or parole hearings,⁵⁶ defendants and inmates are entitled to some protection because neither sentencing judges nor parole board members may consider certain kinds of evidence.⁵⁷ Insofar as there are evidentiary limits, they fall into two broad categories. First, at least at sentencing, evidence must meet a minimal standard of reliability.⁵⁸ Second, sentencing judges and parole boards may make only limited use of certain characteristics of the defendant or inmate such as race, nationality, and gender.⁵⁹

a. *Reliability of Evidence*

In general, sentencing judges have wide discretion to consider whatever evidence they find useful.⁶⁰ However, this discretion is not unlimited.⁶¹ Evidence considered at sentencing must be accurate enough to guarantee the defendant's due process rights.⁶² A sentencing judge may not base her decision on information that is "extensively and materially false"⁶³ and she may only consider evidence that is "relevant and reliable."⁶⁴ Defense attorneys can, and do, challenge expert evidence offered

56. *E.g.*, *Malenchik v. State*, 928 N.E.2d 564, 573 (Ind. 2010) (sentencing); *see, e.g.*, *Davis v. Brown*, 311 F. Supp. 2d 110, 115 (D.D.C. 2004) (parole).

57. *People v. Morgan*, 492 N.E.2d 1303, 1317 (Ill. 1986).

58. *People v. Owens*, 464 N.E.2d 261, 271 (Ill. 1984).

59. *E.g.*, *United States v. Kaba*, 480 F.3d 152, 156 (2d Cir. 2007) (the court may not consider the defendant's race when imposing a sentence); *Tex. Supporters of Workers World Party Pres. Candidates v. Strake*, 511 F. Supp. 149, 155 (D.C. Tex. 1981) (recognizing the right to be considered for parole "without invidious discrimination based on race or national origin"). *But see Wildermuth v. Furlong*, 147 F.3d 1234, 1239 & n.7 (10th Cir. 1998) (noting circuit disagreement on whether a substantive due process right can be violated in the absence of a liberty interest at parole which gives rise to a procedural due process right).

60. *State v. Conn*, 669 P.2d 581, 583 (Ariz. 1983).

61. *Id.*

62. *Roberts v. United States*, 445 U.S. 552, 556 (1980) ("We have . . . sustained due process objections to sentences imposed on the basis of misinformation of constitutional magnitude." (quoting *United States v. Tucker*, 404 U.S. 443, 447 (1972))). Note that evidentiary standards tend to use "reliable" to mean "accurate" or "valid," despite the fact that the two terms have distinct meanings in the scientific and statistical literature. *See Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 594–95 (1993).

63. *Townsend v. Burke*, 334 U.S. 736, 741 (1948).

64. *People v. Morgan*, 492 N.E.2d 1303, 1317 (Ill. 1986).

by the prosecution at sentencing as unreliable.⁶⁵ Even when the sentencing judge finds the evidence reliable and admits it, the effort in raising the objection is not wasted. The objection sensitizes the court to the limitations of the evidence and may inspire the court to afford the evidence less weight.⁶⁶ The situation at parole is markedly different. Although the contours of an inmate's due process right to have a release decision based on reliable information are not defined,⁶⁷ it is clear that parole boards routinely consider exceptionally unreliable evidence like "unsubstantiated rumors."⁶⁸

Only one⁶⁹ appellate court has considered whether statistical risk assessment to predict future dangerousness is reliable enough to be admitted in a sentencing hearing.⁷⁰ The Indiana Supreme Court reviewed the literature on statistical risk assessment and concluded that the instrument in question, the Level of Service Inventory-Revised (LSI-R), was "statistically valid, reliable, and effective in forecasting recidivism."⁷¹ Though the evidence was admitted, the defendant's objection was productive because the court explicitly opined that a sentencing judge may not base the sentence solely on the results of the instrument.⁷²

b. Prohibited Categories of Evidence

There is extensive case law suggesting that neither sentencing judges, nor parole boards may base decisions on a de-

65. *E.g.*, *Malenchik v. State*, 928 N.E.2d 564, 567 (Ind. 2010); see *Barefoot v. Estelle*, 463 U.S. 880, 898 (1983), *superseded by statute on other grounds*, Antiterrorism and Effective Death Penalty Act of 1996, Pub. L. No. 104-132, § 102, 110 Stat. 1214 (1996).

66. See *State v. Woomer*, 299 S.E.2d 317, 320 (S.C. 1982) ("[T]he appellant's objection was addressed to the weight of the evidence and not to its admissibility.").

67. Cf. MODEL PENAL CODE: SENTENCING § 6B.09 cmt. a (Tentative Draft No. 2, 2011) (noting that parole release is "low-visibility, low-process" and unlikely to encourage meaningful review).

68. Reitz, *Indeterminate Sentencing Model*, *supra* note 20, at 284 (internal quotation marks omitted); see also KINNEVY & CAPLAN, *supra* note 1, at 16 ("[I]t seems that [Parole Boards] are willing to listen to anyone with knowledge of or interest in the offender's case.").

69. Warren, *supra* note 11, at 156 (noting that only the Indiana Supreme Court has considered the issue).

70. In two opinions issued the same day, the Indiana Supreme Court allowed the admission of risk assessment results in sentencing hearings. *Malenchik*, 928 N.E.2d at 573; *J.S. v. State*, 928 N.E.2d 576, 577 (Ind. 2010).

71. *Malenchik*, 928 N.E.2d at 573.

72. *Id.* at 568.

defendant's or inmate's race,⁷³ gender,⁷⁴ or national origin without adequate justification.⁷⁵ In the sentencing context, the Constitution does not categorically prohibit consideration of race, gender or nationality; it simply prohibits certain kinds of consideration.⁷⁶ In particular, consideration of gender at sentencing "must pass heightened scrutiny by substantially furthering a legitimate government interest."⁷⁷ For example, in *United States v. Maples*, the reviewing court found that the sentencing judge unconstitutionally imposed a longer sentence on a male defendant than on a female co-defendant.⁷⁸ Specifically, the court found that because there was no link between female gender and the policy goals of rehabilitation or deterrence, it could not use gender as a factor in sentencing.⁷⁹

Classifications based on race or ethnicity in sentencing must typically pass an even more rigorous test.⁸⁰ In that instance, the court requires more than a simple connection between the protected category and a general propensity to deterrence or rehabilitation to legitimate the racial or ethnic classification. The Ninth Circuit, in *United States v. Borrero-Isaza*, found that the trial judge unconstitutionally considered the defendant's Colombian nationality when setting his sen-

73. *United States v. Kaba*, 480 F.3d 152, 156 (2d Cir. 2007) (sentencing); *Tex. Supporters of Workers World Party Presidential Candidates v. Strake*, 511 F. Supp. 149, 155 (S.D. Tex. 1981) (parole).

74. *Williams v. Currie*, 103 F. Supp. 2d 858, 868 (M.D. N.C. 2000) (sentencing); see *Walker v. Luther*, 644 F. Supp. 76, 81 (D. Conn. 1986) (finding parole standards that established different standards for men and women violated the Equal Protection Clause because the state failed to articulate a rational basis for the disparate treatment).

75. *United States v. Borrero-Isaza*, 887 F.2d 1349, 1355-56 (9th Cir. 1989) (sentencing); *Parisie v. Morris*, 873 F. Supp. 1560, 1563 (N.D. Ga. 1995) (parole).

76. *E.g.*, *United States v. Gomez*, 797 F.2d 417, 420 (7th Cir. 1986) (finding that a sentencing judge may consider a defendant's nationality when determining "the identity of the countries which are recognized as often the source" of the drugs involved in the underlying crime).

77. *Williams*, 103 F. Supp. 2d at 862 (citing *J.E.B. v. Alabama ex rel. T.B.*, 511 U.S. 127, 137 (1993)).

78. 501 F.2d 985, 987 (4th Cir. 1974).

79. *Id.*

80. Challenges to the use of race at sentencing are raised under both the Equal Protection Clause and the Due Process Clause. See *United States v. Smart*, 518 F.3d 800, 804 n.1 (10th Cir. 2008). While the tests for these two constitutional protections are different, the result is largely the same as race is not generally a permissible consideration at sentencing under either test. See *id.*

tence following a conviction for a drug crime.⁸¹ The court suggested there would have to be specific evidence demonstrating a direct connection between the defendant's nationality and his crime, for example, if the defendant had trafficked drugs from Colombia, for his Colombian nationality to be relevant to his sentence.⁸² As discussed below, this prohibition on considering race, gender, and nationality characteristics that are unconnected to the commission of the convicted crime has significant ramifications in terms of the positive predictive power of statistical risk assessment instruments because gender and race are strong predictors of recidivism.⁸³

3. Qualifications of the Decision-Maker

The most obvious difference between sentence-length decisions made at parole and those made at sentencing is the identity of the decision-maker.⁸⁴ At sentencing, a judge is typically calling the shots.⁸⁵ Although they may be appointed or elected,⁸⁶ all trial judges are experienced in assessing the weight of evidence,⁸⁷ facilitating the adversarial truth-finding process,⁸⁸ and remaining objective.⁸⁹ By contrast, parole board members, typically appointed by the governor,⁹⁰ make the decision to release or not release a particular inmate.⁹¹ Because "political connec-

81. 887 F.2d 1349, 1355 (9th Cir. 1989).

82. *Id.* at 1356.

83. Gendreau et al., *supra* note 8, at 583.

84. See Reitz, *Sentencing*, *supra* note 23, at 467.

85. See, e.g., *Townsend v. Burke*, 334 U.S. 736, 740–41 (1948).

86. See Note, *Merit Appointment Versus Popular Election: A Reformer's Guide to Judicial Selection Methods in Florida*, 43 FLA. L. REV. 529, 530 (1991) (describing methods of selecting judges in the United States).

87. Cf. David L. Faigman, *To Have and Have Not: Assessing the Value of Social Science to the Law as Science and Policy*, 38 EMORY L.J. 1005, 1088 (1989) (noting that the judge's admissibility decision is followed by the trier-of-fact's determination of what weight the evidence deserves).

88. See *Gardner v. Florida*, 430 U.S. 349, 360 (1977) (noting the Court's "belief that debate between adversaries is often essential to the truth-seeking function of trials").

89. Not surprisingly, there is a mass of literature questioning the objectivity of judges. See, e.g., Shira J. Goodman et al., *What's More Important: Electing Judges or Judicial Independence? It's Time for Pennsylvania to Choose Judicial Independence*, 48 DUQ. L. REV. 859, 860 (2010). This preoccupation with the failures of judicial objectivity illuminates the centrality of the goal of impartiality.

90. KINNEVY & CAPLAN, *supra* note 1, at 7 (finding that governors appoint 85% of releasing authorities with fixed terms of office).

91. *Id.* at 10.

tions are often the main prerequisite for appointment” to a parole board,⁹² and because the process of parole is largely invisible to the public,⁹³ there is a risk that the parole board might be susceptible to political pressure. As one Arkansas parole board member put it: “If the governor likes you, you might get to keep your job.”⁹⁴ Although it is popular to say that parole board members are “experts” in determining when release is appropriate,⁹⁵ this may not be the case as many states have exceptionally minimal requirements for board membership.⁹⁶

Together, the availability of counsel, the requirement of minimally reliable evidence, and qualified judicial oversight provide a relatively robust system of procedural protection for defendants at sentencing. By contrast, the protection for inmates seeking parole release is flimsy. These differences in procedural protection offer a defendant seeking a fair sentencing process more points of intervention than an inmate hoping for parole release is likely to have.

B. STATISTICAL RISK ASSESSMENT

The details of state criminal sentencing systems vary widely, but both sentencing judges and parole boards may use statistical risk assessment to inform their decisions.⁹⁷ Before considering the appropriate procedural placement for such risk assessment instruments, it is worthwhile to examine the instruments themselves, the theory underlying their develop-

92. Reitz, *Indeterminate Sentencing Model*, *supra* note 20, at 285.

93. MODEL PENAL CODE: SENTENCING § 6B.09 cmt. a (Tentative Draft No. 2, 2011).

94. Reitz, *Indeterminate Sentencing Model*, *supra* note 20, at 285 (quoting Arkansas Times Staff & Max Brantley, *Web Special: Dumond Case Revisited*, ARK. TIMES (Sept. 1, 2005), <http://www.arktimes.com/arkansas/web-special-dumond-case-revisited/Content?oid=862759>).

95. See THE TWENTIETH CENTURY FUND TASK FORCE ON CRIMINAL SENTENCING, *supra* note 24, at 117 (noting that a parole board is to “administer [its] decision in accordance with its expertise”).

96. *E.g.*, MISS. CODE ANN. § 47-7-5(2) (2009) (requiring only that parole board members have “at least a bachelor’s degree or a high school diploma and four (4) years’ work experience”).

97. NAT’L CTR. FOR STATE COURTS & VA. CRIMINAL SENTENCING COMM’N, *supra* note 26, at 11–12 (sentencing); KINNEVY & CAPLAN, *supra* note 1, at 12 (parole); see also CONN. GEN. STAT. § 54-125a(a) (“A person . . . may be allowed to go at large on parole in the discretion of the panel of the Board of Parole . . . if (1) it appears . . . that there is a reasonable probability that such inmate will live and remain at liberty without violating the law”); AUSTIN, *supra* note 27, at 5 (“Once an inmate becomes eligible for parole, the decision to release that inmate is guided by an assessment of the inmate’s risk to public safety.”).

ment, and their purposes in the punishment process. The focus of this Note is on the use of statistical risk assessment to aid in sentencing and parole release decisions. These are not the only possible uses of statistical instruments. State departments of corrections and parole also use these instruments to assess the rehabilitative needs of inmates, probationers, and those on parole release.⁹⁸ Although statistical models are sometimes used to assess both needs and risk⁹⁹ and the two are occasionally treated interchangeably,¹⁰⁰ needs assessment is distinct from risk assessment.¹⁰¹ Needs assessment has a broader focus and attempts to match an offender with optimal corrections programming.¹⁰² By contrast, risk assessment focuses on measuring an individual's chances of endangering public safety by reoffending.¹⁰³

States use a wide variety of instruments for risk assessment. Some use commercially-developed instruments, like the LSI-R¹⁰⁴ or its progeny.¹⁰⁵ Others use instruments developed in-house by state agencies.¹⁰⁶ A survey conducted by the Center for Research on Youth and Social Policy reported that eighteen states developed their own instrument in-house, twelve states use the LSI-R instrument, and nine states use one or more of an array of other instruments.¹⁰⁷ In general, these instruments

98. See, e.g., STATE OF CONN., DEP'T OF CORR., CRIMINAL JUSTICE RISK ASSESSMENT STRATEGY 5 (2009), available at <http://www.ct.gov/doc/lib/doc/PDF/PDFReport/RiskAssessmentStrategy.pdf> ("The following assessment instruments are utilized for the purposes of risk, program needs and interventions, classification, development of the Offender Accountability Plan, facility assignment and discharge planning.").

99. See *id.*

100. See D.A. Andrews et al., *The Recent Past and Near Future of Risk and/or Need Assessment*, 52 CRIME & DELINQUENCY 7 *passim* (2006).

101. This fact is frequently noted. E.g., MODEL PENAL CODE: SENTENCING § 6B.09 cmt. a (Tentative Draft No. 2, 2011) ("Needs and risk assessments are distinct tasks . . ."); see also NANCY M. CAMPBELL, NAT'L INST. OF CORR., COMPREHENSIVE FRAMEWORK FOR PAROLING AUTHORITIES IN AN ERA OF EVIDENCE-BASED PRACTICES 37 (2008) (describing the difference between the "needs principle" and the "risk principle").

102. NAT'L CTR. FOR STATE COURTS & VA. CRIMINAL SENTENCING COMM'N, *supra* note 26, at 44.

103. *Id.*

104. See, e.g., D.A. ANDREWS & JAMES L. BONTA, MULTI-HEALTH SYS., INC., LSI-R: LEVEL OF SERVICE INVENTORY-REVISED (2008), available at http://www.assessments.com/assessments_documentation/LSI-R%20Technical%20Brochure.pdf.

105. KINNEVY & CAPLAN, *supra* note 1, at 13.

106. *Id.*

107. *Id.*

require input of a number of data points about the defendant or inmate including factors related to criminal history, family life, education, and employment history.¹⁰⁸ These data points may be either static factors, meaning that they will not change during the individual's incarceration, or dynamic factors, meaning that the input value may change if the individual's behavior or attitudes change.¹⁰⁹ For example, the STATIC-99 risk assessment instrument, used by seventeen states to evaluate sex offenders' risk of recidivism,¹¹⁰ uses static factors like whether the individual has prior convictions.¹¹¹ Dynamic factors take account of the inmate's post-incarceration behavior. Factors might include, for example, whether an inmate has completed correctional programming while incarcerated.¹¹² Although the vast majority of instruments do not explicitly consider race,¹¹³ many consider other factors that have considerable overlap with race, most notably neighborhood of residence.¹¹⁴ Some states employ instruments that use gender as a predictor,¹¹⁵ or have developed separate instruments for men and women.¹¹⁶

108. See, e.g., ANDREWS & BONTA, *supra* note 104.

109. Gendreau et al., *supra* note 8, at 575–76.

110. KINNEVY & CAPLAN, *supra* note 1, at 13.

111. *STATIC-99 Tally Sheet*, STATIC-99 CLEARINGHOUSE, http://www.static99.org/pdfdocs/static-99-coding-rules_e71.pdf (last visited Nov. 27, 2012).

112. Maryland's instrument includes this measure. SHAMIR RATANSI & STEPHEN M. COX, STATE OF CONN., ASSESSMENT AND VALIDATION OF CONNECTICUT'S SALIENT FACTOR SCORE unpublished Appendix C (2007) (on file with author) [hereinafter Appendix C].

113. Reitz, *Indeterminate Sentencing Model*, *supra* note 20, at 280. One judge in Oregon has developed a "sentencing support" software program that considers race/ethnicity in developing a risk profile for a defendant. MICHAEL MARCUS, SENTENCING SUPPORT TOOLS: USER MANUAL FOR JUDGES 10 (2009). This strategy appears to be unique.

114. See ANDREWS & BONTA, *supra* note 104 (including neighborhood of residence in the "accommodation" factor of the LSI-R); Alexander M. Holsinger et al., *Ethnicity, Gender, and the Level of Service Inventory-Revised*, 31 J. CRIM. JUST. 309, 315, 318 (2003) (noting racial trends in seven of ten areas measured by the LSI-R).

115. E.g., THE NAT'L CTR. FOR STATE COURTS & VA. CRIM. SENTENCING COMM'N, *supra* note 26, at 27. Interestingly, members of parole boards report that they only very rarely consider the offender's gender in making their parole release decisions. KINNEVY & CAPLAN, *supra* note 1, at 19. This may indicate that the parole board is unaware of whether the risk assessment rolls gender into the score an inmate receives.

116. THOMAS BLOMBERG ET AL., VALIDATION OF THE COMPAS RISK ASSESSMENT CLASSIFICATION INSTRUMENT 13 (2010), available at <http://www.criminologycenter.fsu.edu/pdf/pretrial/Broward%20Co.%20COMPAS%20Validation%202010.pdf>.

Risk assessment scores may be obtained based on interviews,¹¹⁷ pen and paper checklists,¹¹⁸ or by automatically extracting data from the inmate or defendant's record.¹¹⁹ The personnel responsible for administering risk assessments varies state by state;¹²⁰ but, in any case, significant staff training is necessary to administer the instruments competently and with consistent proficiency.¹²¹

Methods of writing statistical risk assessment instruments fall into two categories.¹²² First, an instrument may be based on empirical data.¹²³ This type of instrument is developed by drawing a sample of offenders, observing recidivism rates as well as other offender characteristics and then using the statistical relationships between these characteristics and recidivism to build the instrument.¹²⁴ Once built, the risk assessment instrument produces individual numeric risk scores depending on the characteristics of the individual being assessed.¹²⁵ For example, Virginia used a common statistical method, multivariate logistic regression, to construct a risk assessment instrument that estimates risk of recidivism based on eleven factors including gender, age, employment status, and factors related to prior criminal behavior.¹²⁶ Risk assessment instruments may

117. Holsinger et al., *supra* note 114, at 310 (describing administration of the LSI-R).

118. Tammy Meredith et al., *Developing and Implementing Automated Risk Assessments in Parole*, 9 JUST. RESEARCH & POL'Y 1, 2 (2007).

119. *Id.*

120. For example, in Connecticut, prison personnel are responsible for administering risk assessment instruments, see STATE OF CONN., DEP'T OF CORR., *supra* note 98, at 5, while in Los Angeles County, probation officers are responsible for ongoing risk and needs assessment with their probationers, see SUSAN TURNER & TERRY FAIN, VALIDATION OF THE LOS ANGELES COUNTY PROBATION DEPARTMENT'S RISK AND NEEDS ASSESSMENT INSTRUMENTS, at xi (2003), available at <https://www.ncjrs.gov/pdffiles1/nij/grants/201303.pdf>.

121. Anthony W. Flores et al., *Predicting Outcome with the Level of Service Inventory-Revised: The Importance of Implementation Integrity*, 34 J. CRIM. JUST. 523, 528 (2006); see also AUSTIN ET AL., *supra* note 27, at 15.

122. DORIS LAYTON MACKENZIE, WHAT WORKS IN CORRECTIONS: REDUCING THE CRIMINAL ACTIVITIES OF OFFENDERS AND DELINQUENTS 63 (2006).

123. *Id.*

124. *Id.*

125. See Kirk Heilbrun et al., *Violence Risk Assessment Tools: Overview and Critical Analysis*, in HANDBOOK OF VIOLENCE RISK ASSESSMENT 1, 5 (Randy K. Otto & Kevin S. Douglas eds., 2010) (noting that empirically derived risk assessment instruments produce reproducible scores based on the "predictor variables").

126. NAT'L CTR. FOR STATE COURTS & VA. CRIM. SENTENCING COMM'N, *supra* note 26, at 26–27.

also be developed theoretically.¹²⁷ Theoretical risk assessment instruments measure variables that are, according to psychology or criminology experts, “theoretically related” to recidivism¹²⁸ and then measure the statistical relationship between these variables and recidivism in populations.¹²⁹ Theoretical risk assessment instruments, including the popular LSI-R,¹³⁰ typically use the correlation between values of variables and recidivism risk to assign weights to inmates’ responses.¹³¹ These weights are added to produce numeric risk scores.¹³² Risk assessment instruments developed through either method typically produce risk scores on a finite scale, for example, from zero to twelve.¹³³ Risk instruments themselves do not necessarily resolve downstream policy choices of determining meaningful cutoffs in risk scores.¹³⁴

In general it is difficult for a lay person to access information about a state’s policy with respect to risk assessment instruments.¹³⁵ Often states do not readily provide information about which risk assessment they use, how the instruments were developed, or how they are used in practice. For example, a request for the “risk assessment instrument . . . along with

127. MACKENZIE, *supra* note 122, at 63.

128. *Id.*; see also PETER RAYNOR ET AL., RISK AND NEED ASSESSMENT IN PROBATION SERVICES: AN EVALUATION 9 (2000), available at <http://library.npia.police.uk/docs/hors/hors211.pdf> (describing the selection of items in the LSI-R as “intended to have a theoretical and professional rationale justifying the selection of these particular items as relevant to offending”).

129. MACKENZIE, *supra* note 124, at 63.

130. See, e.g., Holsinger et al., *supra* note 114, at 310.

131. See, e.g., *Risk to Re-offend Score Chart*, GA. STATE BOARD PARDONS & PAROLES, http://www.pap.state.ga.us/opencms/export/sites/default/resources/RISK_TO_RE.pdf (last visited Nov. 27, 2012) (laying out Georgia’s summation scoring method).

132. *Id.*

133. *STATIC-99 Tally Sheet*, *supra* note 111.

134. Although the responsibility for determining the acceptable level of risk rests with the states, the developers of risk assessment instruments may provide guidance on how state officials, including parole boards and judges ought to interpret the results. See, e.g., D.A. Andrews, James Bonta & Stephen Wormith, *The Level of Service (LS) Assessment of Adults and Older Adolescents*, in HANDBOOK OF VIOLENCE RISK ASSESSMENT, *supra* note 125, at 199, 205.

135. See Susan J. Sachsenmaier & Stephen J. Lally, *Toward a Scientific Foundation of Sex Offender Risk Assessment*, in SEX AND VIOLENCE: THE PSYCHOLOGY OF CRIME AND RISK ASSESSMENT 26–27 (David P. Farrington et al. eds., 2001).

any accompanying documentation or manuals for its use”¹³⁶ made under New York’s Freedom of Information Law yields only a copy of a memorandum from the Chairwoman of New York’s Board of Parole explaining appropriate use of the state’s newest instrument,¹³⁷ and a printout of the risk assessment scoring sheet.¹³⁸ No information about the design or validation of the instrument is provided. Given the range of possible methods by which a statistical risk assessment instrument may be prepared and the nuanced ways in which this method affects the validity and reliability of the instrument, as discussed below, the unavailability of this type of information is troubling.

II. STATISTICAL RISK ASSESSMENT’S LEGAL AND EMPIRICAL LIMITATIONS

True risk assessment instruments are creatures of statistics and thus are subject to a number of inherent limitations. No matter how thorough the analysis, no statistical model can ever be completely accurate.¹³⁹ Uncertainty in the result produced is inevitable. For example, a hypothetical risk assessment instrument might predict that a particular offender has a 25% probability of reoffending. That 25% is a point estimate. Such a point estimate does not express how confident a statistician is in that the estimate is accurate. To understand the degree of confidence, statisticians look to a confidence interval indicating the range of possible values around the point estimate that the true value is likely to be.¹⁴⁰ Perhaps the instrument is very robust and produces a very narrow confidence interval—say, that evaluator is 95% sure that the true value of the offender’s risk of reoffending is between 24% and 26%. Alternatively, the instrument might be very weak and produce a 95% confidence interval of between 10% and 40%.

Because perfection is not possible, states ought to face head-on the issue of uncertainty and develop strategies for

136. Letter from Pari McGarraugh, to Patrick Lawlor, N.Y. Dep’t of Corr. & Comm. Supervision, Counsel’s Office (Nov. 2, 2011) (on file with author).

137. Memorandum from Andrea W. Evans, Chairwoman, N.Y. Dep’t of Corr. & Comm. Supervision, to the Members of the Board of Parole (Oct. 5, 2011) (on file with author).

138. COMPAS REENTRY ASSESSMENT (2007) (on file with author).

139. Emma J. Palmer, *Risk Assessment: Review of Psychometric Measures*, in *SEX AND VIOLENCE: THE PSYCHOLOGY OF CRIME AND RISK ASSESSMENT*, *supra* note 135, at 7.

140. RAND R. WILCOX, *BASIC STATISTICS: UNDERSTANDING CONVENTIONAL METHODS AND MODERN INSIGHTS* 103 (2009).

managing it.¹⁴¹ Despite how important it is to resolve questions of predictive power, not all states directly address the problem. In fact, only about 85% of states using risk assessment instruments report having validated the instruments in the state's own population in any way.¹⁴² The following sections outline these concerns by examining weaknesses of statistical risk assessment instruments both theoretically and in practice.

A. THE ASSUMPTIONS UNDERLYING RISK ASSESSMENT INSTRUMENTS ARE OFTEN UNJUSTIFIED

As discussed above, risk assessments are typically based on the results of regression models or are constructed as weighted sums.¹⁴³ Instruments based on regression models make several assumptions related to their external validity. First, regression models assume random selection, meaning that each individual in the population of interest has an equal chance of being included in the sample.¹⁴⁴ Second, regression models assume independence of observations, meaning that the behavior of each individual in the population is entirely unrelated to the behavior of any of the others.¹⁴⁵ In the context of risk assessment, neither of these assumptions is likely to be justified.¹⁴⁶ High quality data on recidivism is scarce and certain types of outcomes, like whether an incarcerated person would have reoffended if released, are counterfactual and inherently impossible to measure directly.¹⁴⁷ For this reason, data sets underlying risk assessment instruments tend to be "convenience samples" rather than true random samples drawn from the population of interest.¹⁴⁸ Furthermore, individuals in these samples are rarely independent because they are likely to

141. Cf. NAT'L CTR. FOR STATE COURTS & VA. CRIM. SENTENCING COMM'N, *supra* note 26, at 29–30 (describing Virginia's method of determining an appropriate risk threshold for diverting an offender from incarceration).

142. KINNEVY & CAPLAN, *supra* note 1, at 13.

143. MACKENZIE, *supra* note 122, at 63.

144. Richard A. Berk & David A. Freedman, *Statistical Assumptions as Empirical Commitments*, in PUNISHMENT AND SOCIAL CONTROL: ESSAYS IN HONOR OF SHELDON L. MESSINGER 245 n.1 (Thomas G. Blomberg & Stanley Cohen eds., 1995).

145. *Id.* at 245 n.1, n.4.

146. *Id.* at 246.

147. Cf. Christopher T. Lowenkamp et al., *Risk/Need Assessment, Offender Classification, and the Role of Childhood Abuse*, 28 CRIM. JUST. & BEHAVIOR 543, 548–49 (2001) (including only those inmates released on parole in the validation study).

148. See Berk & Freedman, *supra* note 144, at 246.

have shared influences like supervision by the same parole officer.¹⁴⁹ Indeed, such samples are likely to be inherently clustered around particular programs or geographic locations.¹⁵⁰ That such clustered observations are likely not to be independent because individuals systematically share characteristics does not necessarily bias the result of an analysis, but it does result in an overestimate of certainty.¹⁵¹

Even if the sample is properly drawn and the analysis perfectly conducted, statistical models only predict probable behavior for a typical member of the population from which the original sample was drawn.¹⁵² This poses obvious problems when the sample population is very different than the one in which the risk assessment instrument is to be used.¹⁵³ For example, the LSI-R was developed in populations of Canadian inmates, and, although the instrument was validated in inmate populations in the United States, the validation may provide false comfort as its sample included predominantly Caucasian inmates.¹⁵⁴ This methodological quirk is not benign; other research shows that the LSI-R predicts recidivism much less effectively for African-Americans and Hispanics than for Caucasians.¹⁵⁵ Assuming a close match between the sample population and the population in which the instrument is to be applied, a good risk assessment instrument still only predicts average behaviors.¹⁵⁶ The instrument cannot predict the behavior of any one individ-

149. *See id.* at 250.

150. *See id.* at 246 (noting “the data in hand are simply the data most readily available”).

151. *See* Jeromy Anglim, *Clustered Samples and Assuming Independence of Observations*, JEROMY ANGLIM’S BLOG: PSYCHOLOGY AND STATISTICS (Feb. 26, 2010), <http://jeromyanglim.blogspot.com/2010/02/clustered-samples-and-assuming.html>.

152. *See* Berk & Freedman, *supra* note 144, at 245.

153. *See* Lowenkamp et al., *supra* note 147, at 560 (conceding that the results of the validation study are limited if the sample was not representative).

154. Melinda D. Schlager & David J. Simourd, *Validity of the Level of Service Inventory-Revised (LSI-R) Among African American and Hispanic Male Offenders*, 34 CRIM. JUST. & BEHAVIOR 545, 546 (2007).

155. *Id.* at 553.

156. Kathleen Auerhahn, *Selective Incapacitation and the Problem of Prediction*, 37 CRIMINOLOGY 703, 708 (1999).

ual.¹⁵⁷ By definition, these risk assessment instruments fail to account for atypical behavior.¹⁵⁸

In addition, risk assessment instruments based on either summation models, such as the LSI-R, or regressions are subject to an additional key limitation. Unless specifically designed to do so, the instrument cannot account for interaction between input measures. They assume that the uptick in risk of recidivism associated with each input factor remains constant regardless of inputs on other risk factors.¹⁵⁹ To illustrate, Georgia's risk assessment instrument, used to guide parole release decisions, uses a summation model and assigns two risk "points" to an offender who has a history of drug or alcohol abuse no matter the inmate's age when he entered prison.¹⁶⁰ Georgia's model assumes that at every age, a history of drug or alcohol abuse has the same effect on risk of recidivism.¹⁶¹ The failure to account for interaction between substance abuse and age is significant because there is evidence indicating that the risk of recidivism associated with prior substance abuse does indeed vary with age.¹⁶² This shortcoming is relevant for risk factors other than history of substance abuse because interactions between multiple risk factors are likely to have a significant impact on the predictive power of the instrument,¹⁶³ especially in offender subpopulations.¹⁶⁴

Risk assessment instruments frequently assume random selection, independence of individuals and a lack of interaction between variables.¹⁶⁵ Because, as discussed above, these assumptions are unlikely to be justified in practice, published

157. *Id.*; see also O'Leary & Glaser, *supra* note 4, at 140 (noting that although older inmates are less likely to reoffend than younger inmates, "[b]y no means should it be inferred that all old prisoners are good risks or all youngsters poor risks").

158. O'Leary & Glaser, *supra* note 4, at 140.

159. See RAYNOR ET AL., *supra* note 128, at 65.

160. See *Risk to Re-offend Score Chart*, *supra* note 131.

161. See *id.*

162. Cf. Darrell J. Steffensmeier et al., *Age and the Distribution of Crime*, 94 AM. J. SOC. 803, 821 (finding that the risk of committing a substance-related crime is dramatically skewed toward younger offenders).

163. E.g., MICHAEL SHADER, U.S. DEP'T OF JUSTICE, RISK FACTORS FOR DELINQUENCY: AN OVERVIEW 7 (2004), available at <https://www.ncjrs.gov/pdffiles1/ojdp/frd030127.pdf>; cf. David A. Wolfe & Robin McGee, *Dimensions of Child Maltreatment and Their Relationship to Adolescent Adjustment*, 6 DEV. & PSYCHOPATHOLOGY 165, 178 (1994) (noting that interactions between risk factors are "powerful predictors" of child maltreatment).

164. See RAYNOR ET AL., *supra* note 128, at 65.

165. See *id.*; Berk & Freedman, *supra* note 144, at 245 n.1.

studies evaluating risk assessment instruments tend to overstate their predictive power.¹⁶⁶

B. GARBAGE IN, GARBAGE OUT

It is well-recognized that the output of a statistical model is only as good as the input data.¹⁶⁷ In the risk assessment context, the “garbage in, garbage out” problem comes up when the instrument requires subjective input factors that are difficult to measure accurately,¹⁶⁸ even in a controlled research setting.¹⁶⁹ Many risk assessment models incorporate subjective variables¹⁷⁰ which are notoriously difficult to measure consistently.¹⁷¹ The garbage in, garbage out problem is especially likely to apply to dynamic factors not apparent until the individual has been observed in the prison environment because such dynamic factors tend to focus on subjective evaluations rather than on easily categorized characteristics.¹⁷² Given the unavoidable difficulty in consistently and accurately measuring such factors, there is likely to be non-negligible error in these data points.¹⁷³ For example, Alabama’s risk assessment instrument, used to guide parole decision-making, asks the rater to evaluate where on a four-point scale an inmate’s marital and family relationships fall—from “[g]ood support and influence” to “[s]erious

166. See Patrick Keenan, *The New Deterrence: Crime and Policy in the Age of Globalization*, 91 IOWA L. REV. 505, 524 (2006) (noting “most models incorporate too many simplifications and occasionally unrealistic assumptions to be accurate”).

167. See e.g., Clive R. Hollin, *Treatment Programs for Offenders: Meta-Analysis, “What Works,” and Beyond*, 22 INT’L J.L. & PSYCHIATRY 361, 367 (1999) (noting the “garbage in, garbage out” problem in using meta-analysis to evaluate the efficacy of correctional programs).

168. Reitz, *Indeterminate Sentencing Model*, *supra* note 20, at 280.

169. 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, 1875 (2003) (pointing out that the developer of one risk assessment instrument “recognizes its subjective nature and recommends that at least two independent ratings be obtained and averaged” (quotation marks omitted)).

170. For example, the LSI-R takes into account multiple subjective risk factors including the offender’s “peer interactions,” “authority interactions,” and whether the offender is “supportive of crime.” Holsinger, *supra* note 114, at 312–13.

171. See Lisa M. Dennis, *Constitutionality, Accuracy, Admissibility: Assessing Expert Predictions of Future Violence in Capital Sentencing Proceedings*, 10 VA. J. SOC. POL’Y & L. 292, 307 (2002).

172. Gendreau et al., *supra* note 8, at 575–76.

173. See *id.*

domestic discord or domestic violence.”¹⁷⁴ It may be that family support, when perfectly measured, is an excellent predictor of recidivism;¹⁷⁵ it is less clear that the same question item administered in practice produces high-quality data points. In keeping with the garbage in, garbage out principle, the inclusion of such error-ridden inputs produces risk scores that are unreliable in unpredictable ways.¹⁷⁶ This concern is amplified when instruments are administered by untrained individuals, in real-world settings, with uncooperative subjects.¹⁷⁷

C. THE UNKNOWN PREDICTIVE POWER OF CONSTITUTIONALLY COMPLIANT MODELING

Apart from the statistical limitations of risk assessment, its use may be subject to legal limitations. Several factors typically used in statistical risk assessment, if tested through the appeals process, may not pass constitutional muster.¹⁷⁸ Recent drafts of the MPC note, without explanation, that considering race and ethnicity “raises serious constitutional concerns,”¹⁷⁹ but find, again without discussion, that consideration of gender in risk assessment is permissible.¹⁸⁰ Upon closer examination, the MPC drafters are spot-on in suggesting that including race or ethnicity into the list of factors built into a risk assessment instrument used at either parole or sentencing is unconstitutional.¹⁸¹ Their position on gender, however, is not convincing under the case law.¹⁸²

174. Appendix C, *supra* note 112.

175. Indeed, empirical evidence shows that family structure and family criminality are predictors of recidivism. Gendreau et al., *supra* note 8, at 583.

176. See Hollin, *supra* note 167, at 367.

177. See Flores et al., *supra* note 121, at 528.

178. See MODEL PENAL CODE: SENTENCING § 6B.09 reporter’s note e (Tentative Draft No. 2, 2011).

179. *Id.* at reporter’s note i; see MODEL PENAL CODE: SENTENCING § 6B.06(2)(a) (Tentative Draft No. 1, 2007). *But see* MARCUS, *supra* note 113, at 10 (incorporating race/ethnicity into his risk assessment tool).

180. MODEL PENAL CODE: SENTENCING § 6B.06(4)(b) (Tentative Draft No. 1, 2007). The drafters do not justify this statement with an explicit constitutional discussion nor do they explicate exactly what “grave concerns” trouble them.

181. *E.g.*, United States v. Kaba, 480 F.3d 152, 156 (2d Cir. 2007) (holding that national origin cannot be the basis for determining a sentence).

182. See, *e.g.*, Williams v. Currie, 103 F. Supp. 2d 858, 868 (M.D.N.C. 2000) (holding that considering gender in sentencing violated the defendant’s equal protection rights).

Because the vast majority of risk assessment instruments do not consider race directly,¹⁸³ there remains a question of whether analysis of factors that have nearly complete overlap with race is constitutional.¹⁸⁴ For example, the frequently-used LSI-R takes into consideration whether the offender lives in a high-crime neighborhood.¹⁸⁵ Because, at least in some cities, neighborhood of residence correlates nearly perfectly with race,¹⁸⁶ this factor, at least in those geographic locations, operates as a proxy for race.

While constitutional case law does not strictly prohibit all consideration of race or nationality at sentencing,¹⁸⁷ it minimally requires the government to have a good reason for such consideration. A satisfactory reason may be that the defendant's race or ethnicity is relevant to rehabilitation or deterrence.¹⁸⁸ A stricter court may even require the defendant's race or ethnicity to have been specifically connected to the commission of the defendant's crime.¹⁸⁹ Risk factors that overlap closely with race are unlikely to meet either of these standards. The empirical literature simply cannot demonstrate a causal relationship between race or ethnicity and propensity toward recidivism or rehabilitation; at best, validated risk factors are correlated with outcomes.¹⁹⁰ Further, even if such a causal relationship could be conclusively proven at the population level, it is not reasonable to conflate population-level causation with individual-level cau-

183. *But see, e.g.*, MARCUS, *supra* note 113.

184. This concern is raised, and explained away by the MPC's drafters. MODEL PENAL CODE § 6B.06 cmt. c (Tentative Draft No. 1, 1997).

185. D.A. ANDREWS & JAMES L. BONTA, LSI-R PROFILE REPORT FOR REX DARLINGTON 4 (2001), *available at* http://www.psychassessments.com.au/products/59/prod59_report1.pdf.

186. *See, e.g.*, EDWARD L. GLAESER ET AL., THE BROOKINGS INST. CTR. ON URBAN & METRO. POLICY, RACIAL SEGREGATION IN THE 2000 CENSUS: PROMISING NEWS 5–7 (2001), *available at* <http://www.brookings.edu/es/urban/census/glaeser.pdf> (noting that the Midwest and Northeast remain more segregated than the South and West and that the largest cities are “significantly more segregated” than the national average).

187. *See, e.g.*, United States v. Borrero-Isaza, 887 F.2d 1349, 1356 (9th Cir. 1989) (noting that geography may be considered if related to the crime).

188. *E.g.*, United States v. Maples, 501 F.2d 985, 987 (4th Cir. 1974).

189. *E.g.*, Borrero-Isaza, 887 F.2d at 1356.

190. *See* John Q. La Fond, *Clinical, Legal and Ethical Issues for Mental Health Professionals in Implementing a Sexual Predator Law in the United States*, in SEX AND VIOLENCE: THE PSYCHOLOGY OF CRIME AND RISK ASSESSMENT 116 (David P. Farrington et al. eds., 2001) (“Even validated risk factors only establish correlations; they do not establish causation.”).

sation.¹⁹¹ Because a population-level causal relationship between race or ethnicity or proxy factors is only dubiously supported by the literature,¹⁹² and a risk assessment instrument's reliance on population-level correlation could never show individual-level causation,¹⁹³ a court would be unlikely to find that inclusion of risk factors that stand proxy for race demonstrates a general or specific propensity for rehabilitation or recidivism.¹⁹⁴ Even aside from the constitutional concern, there is a distinct danger of bad public policy when correlation is confused with causation in this context. This is evident in one study's reasoning that the correlation between minority race and criminal behavior indicates that minority populations "should have additional correctional resources made available to them"¹⁹⁵ but does not consider what underlying societal conditions might have produced the correlation. Similarly, gender is not generally a permissible consideration at either sentencing or parole because the statistical relationship between gender and recidivism does not indicate a propensity for recidivism or rehabilitation and it is not specifically connected with the commission of a particular crime.¹⁹⁶

In order to create a risk assessment instrument that does not offend the Constitution, race and ethnicity, factors closely overlapping with race and ethnicity, and gender must be purged from the list of inputs. But because race¹⁹⁷ and gender¹⁹⁸ are fairly reliable predictors of criminal behavior,¹⁹⁹ removing them will reduce the predictive capability of risk assessments.²⁰⁰ Nailing down the exact magnitude of this reduction in

191. Such an unfounded conclusion is known as an "ecological fallacy." See Sharon Schwartz, *The Fallacy of the Ecological Fallacy: The Potential Misuse of a Concept and the Consequences*, 84 AM. J. PUB. HEALTH 819, 820–22 (1994).

192. *Id.* at 821.

193. *Id.*

194. See *Borrero-Isaza*, 887 F.2d at 1356 (requiring a specific connection to the crime); *United States v. Maples*, 501 F.2d 985, 987 (4th Cir. 1974) (requiring only a general connection).

195. See Holsinger, *supra* note 114, at 318. It is not entirely clear what such "correctional resources" might include, but certainly one reasonable interpretation is that correctional resources is code for more incarceration.

196. See *Williams v. Currie*, 103 F. Supp. 2d 858, 868 (M.D.N.C. 2000).

197. Reitz, *Sentencing*, *supra* note 23, at 486.

198. Patrick A. Langan & David J. Levin, *Recidivism of Prisoners Released in 1994*, 15 FED. SENT'G REP. 58, 61 (2002).

199. Gendreau et al., *supra* note 8, at 583.

200. Cf. NAT'L CTR. FOR STATE COURTS & THE VA. CRIMINAL SENTENCING COMM'N, *supra* note 26, at 27–28 (describing the Virginia Sentencing Commis-

predictive power is difficult for two reasons. First, extracting constitutionally dubious input factors from instruments that consider interaction between factors may be impossible without rerunning the initial analyses.²⁰¹ Second, even if the initial development studies are susceptible to manipulation that would reveal the predictive power of only a subset of the included risk factors, these development studies may not be published or may not be published with enough methodological detail to allow this type of calibration.²⁰²

The exclusion of gender presents an added wrinkle not present when factors overlapping with race or ethnicity are excluded.²⁰³ Specifically, because women recidivate at a lower rate than men do,²⁰⁴ removing gender from the model artificially inflates women's risk assessment scores.²⁰⁵ To some extent, this effect is moderated by the fact that there are systematic differences between male and female offenders which mean female offenders tend to receive lower risk scores even when the instrument does not explicitly consider gender.²⁰⁶ However, the problem remains, leaving parole boards and sentencing judges with two bad choices: (1) despite the constitutional concerns, use an instrument that considers gender because it more accurately predicts women's risk or (2) comply with the constitutional prohibition on unwarranted consideration of gender and tolerate that women will receive higher risk scores.

D. RISK ASSESSMENT INSTRUMENTS IN PRACTICE

The foregoing critiques are primarily based on the empirical literature examining statistical risk assessment. In practice, these limitations continue to apply and are amplified by likely errors caused by improper implementation and administration of the instruments as well by the manner in which results of risk assessments are communicated to decision-

sion's efforts to reassess the validity of its instrument after choosing not to include race, despite its predictive power).

201. *See id.*

202. *See* Sachsenmaier & Lally, *supra* note 135, at 26–27.

203. Because non-white race is positively associated with recidivism, see Gendreau et al., *supra* note 8, at 583, not considering race tends to pull risk scores for non-white offenders down.

204. *See* Langan & Levin, *supra* note 198, at 61.

205. MODEL PENAL CODE: SENTENCING § 6B.06(2)(a) cmt. d (Tentative Draft No. 1, 2007).

206. *See* Holsinger, *supra* note 114, at 313.

makers.²⁰⁷ Problems with implementation might include failure to conduct validation studies in the population in which the instrument is to be used²⁰⁸ or alterations to the original model on which the instrument is based. For example, Texas uses a risk assessment instrument originally developed by the state of Wisconsin.²⁰⁹ Texas alters the instrument by weighting one risk factor, whether the inmate has committed an assault in the previous five years, more heavily than the factor is weighted in the original instrument.²¹⁰ Alterations, including adding risk factors or reweighting risk factors, are particularly troubling because such alterations interrupt the integrity of the underlying analysis.²¹¹

The predictive power of risk assessment instruments may also be limited by suboptimal administration. The empirical literature repeatedly emphasizes the importance of proper staff training to reach the full predictive potential of risk assessment instruments.²¹² In practice, however, the staff members administering risk assessment instruments²¹³ and the individuals interpreting the results may not have the required training.²¹⁴ Moreover, uneven staff skill levels caused by inadequate training diminish inter-rater reliability by introducing systematic differences in risk scores between different administrators.²¹⁵ It may be especially difficult to detect problems caused by poor administration when instruments are used at stages of the criminal process that have little knowledgeable oversight.²¹⁶

207. See Flores et al., *supra* note 121, at 528.

208. See KINNEVY & CAPLAN, *supra* note 1, at 13.

209. MARC LEVIN, POLICY PERSPECTIVE: THE ROLE OF RISK ASSESSMENT IN ENHANCING PUBLIC SAFETY AND EFFICIENCY IN TEXAS CORRECTIONS 2 (2010), available at www.texaspolicy.com/sites/default/files/documents/2010-07-PP15-RiskAssessment_ml.pdf.

210. *Id.*

211. Cf. Berk & Freedman, *supra* note 144, at 245–46 (discussing the assumptions underlying statistical models).

212. E.g., *id.*; CAMPBELL, *supra* note 101, at 38; Holsinger, *supra* note 114, at 310.

213. See Flores et al., *supra* note 121, at 526 (comparing predictive power of the LSI-R when used by formally trained and untrained staff members).

214. For example, parole board members interpreting risk assessment scores may not have expertise in criminal psychology. See MISS. CODE ANN. § 47-7-5(2) (2009) (amended 2012).

215. AUSTIN ET AL., *supra* note 27, at 15.

216. Cf. D.A. Andrews et al., *Classification for Effective Rehabilitation: Rediscovering Psychology*, 17 CRIM. JUST. & BEHAV. 19, 26 (1990) (expressing dismay at the frequency with which criminal justice professionals are unfamiliar with simple concepts of risk and recidivism).

Finally, the manner in which the results of risk assessment instruments are typically communicated to decision-makers impacts their utility. Decision-makers generally receive the results of a risk assessment expressed as a point estimate of a probability of recidivism.²¹⁷ By relying on a point estimate of probability of recidivism, the decision-maker is blind to the level of error inherent in that probability calculation.²¹⁸ So while a parole board member or sentencing judge may get a report that indicates an offender has a “high” risk of recidivism, she is not likely to be informed of the confidence interval surrounding that point estimate.

In sum, statistical risk assessments suffer from several key deficiencies that limit their predictive power. Even if these instruments are administered and implemented perfectly—a practical impossibility—they can only predict the expected population behavior and this prediction is inevitably uncertain. In no instance can a risk assessment instrument ever predict with certainty what any one person will do in the future.²¹⁹

III. BALANCING PREDICTIVE POWER AND PROCEDURAL PROTECTION

Because statistical risk assessment is both imperfect and useful to the criminal justice system,²²⁰ the use of statistical risk assessment to make choices about incarceration should be carefully controlled. A recent draft of the MPC proposes language to this effect: “The commission shall develop instruments or processes, supported by current and ongoing recidivism research, that will estimate the relative risks that individual offenders pose to public safety through their future criminal conduct.”²²¹ The Code further clarifies that instruments may be

217. See NAT’L CTR. FOR STATE COURTS & THE VA. CRIMINAL SENTENCING COMM’N, *supra* note 26, at 29 (describing Virginia’s policy decision to recommend diversion from incarceration for those offenders whose scores indicate a less than 12% probability of reconviction within three years).

218. See 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 s60, s61 (2007).

219. See La Fond, *supra* note 190, at 115–16 (“It should be noted that actuarial prediction tools only identify offenders who are members of a group that has these probabilities of committing another sex offense. They do not necessarily identify which individual members of the group will reoffend.”).

220. See, e.g., Wolff, *supra* note 42, at 1406 (noting that statistical risk assessment predicts risk of recidivism more accurately than intuition alone).

221. MODEL PENAL CODE: SENTENCING § 6B.09(2) (Tentative Draft No. 2, 2011).

incorporated into sentencing guidelines of the state when they “prove sufficiently reliable.”²²² The Code’s drafters argue that placing risk assessment at sentencing “domesticates’ the use of risk assessments by repositioning them in the open forum of the courtroom, where the tools devised . . . are available for inspection, and where the constitution guarantees the offender legal representation to contest any adverse findings.”²²³

This Part expands on the reasoning of the drafters of the MPC and argues that statistical risk assessment potentially provides sentencing judges with useful information relevant to determining an appropriate sentence and that the procedures available to defendants at sentencing protect defendants from overreliance on these imperfect instruments better than the procedures available at parole. Indeed, one major advantage of this procedural adjustment is that a skillful defense attorney is able to draw the sentencing judge’s attention to statistical weaknesses discussed above. This Part further suggests that the MPC be modified to include a definition of “sufficiently reliable” and that sentencing judges be specifically instructed to consider the uncertainty of risk estimates produced by any risk assessment instruments in order to provide defendants with even more robust protection.

A. WHY USE RISK ASSESSMENT AT ALL?

Despite their limitations, risk assessment instruments have considerable value. Recidivist criminals are a serious concern,²²⁴ and scientists have empirically demonstrated that statistical risk assessment much more accurately predicts recidivism than do individuals relying on intuition and experience.²²⁵ For this reason, judges should be allowed to use the results of statistical risk assessment, as long as the procedural protections available to defendants are adequate and the instrument sufficiently reliable to guard against injustice.²²⁶

222. *Id.*

223. MODEL PENAL CODE: SENTENCING § 6B.09 cmt. (Tentative Draft No. 2, 2011).

224. See, e.g., Alfred Blumstein & Soumyo Moitra, *The Identification of “Career Criminals” from “Chronic Offenders” in a Cohort*, 2 LAW & POL’Y Q. 321, 322 (1980) (citing a study finding a small proportion of offenders are responsible for a large portion of criminal offenses).

225. Wolff, *supra* note 42, at 1406 n.73.

226. MODEL PENAL CODE: SENTENCING § 6B.09 cmt. a (Tentative Draft No. 2, 2011).

Setting aside for a moment the issue of procedural protection, risk assessment may actually better prevent recidivism when it is used at sentencing than when it issued at parole. Because spending any time in prison at all is a strong predictor of recidivism,²²⁷ especially for low risk offenders,²²⁸ using risk assessment to identify low-risk offenders appropriate for work release or other less restrictive sentences may reduce crime by preventing this risk factor from ever developing.²²⁹ The second draft of the MPC resonates with this reasoning, advising judges and sentencing commissions to use low risk scores as a mitigating factor and grounds for a downward departure in sentencing.²³⁰ Thus, moving the use of statistical risk assessment forward in the criminal process may reduce state expenditures on unnecessary incarceration, prevent recidivism, and promote the liberty of low-risk offenders. This strategy has proven effective in Virginia. The state reports that using such a strategy diverted 555 offenders from incarceration and saved state and local governments over \$8.5 million dollars between 1997 and 1999.²³¹

B. SENTENCING, NOT PAROLE, IS THE APPROPRIATE FORUM TO CONSIDER THE RESULTS OF STATISTICAL RISK ASSESSMENT

Using risk assessment at sentencing rather than at parole both better serves the public interest in preventing recidivism and is fairer to defendants. The major argument for waiting until parole to use risk assessment is that dynamic factors related to inmates' in-prison behavior are predictive of recidivism.²³² This argument is limited because there is considerable doubt regarding whether subjective dynamic factors can be or

227. See Gendreau et al., *supra* note 8, at 583 (reporting criminal history as a predictor of recidivism).

228. CHRISTOPHER T. LOWENKAMP & EDWARD J. LATESSA, UNDERSTANDING THE RISK PRINCIPLE: HOW AND WHY CORRECTIONAL INTERVENTIONS CAN HARM LOW-RISK OFFENDERS 7 (2004), available at <http://www.yourhonor.com/dwi/sentencing/RiskPrinciple.pdf>.

229. For example, Virginia now uses risk assessment at sentencing with the goal of diverting offenders from the prison system and reducing overall inmate populations. NAT'L CTR. FOR STATE COURTS & THE VA. CRIMINAL SENTENCING COMM'N, *supra* note 26, at 9–10.

230. MODEL PENAL CODE: SENTENCING § 6B.09 cmt. c (Tentative Draft No. 2, 2011).

231. NAT'L CTR. FOR STATE COURTS & THE VA. CRIMINAL SENTENCING COMM'N, *supra* note 26, at 100.

232. See Gendreau et al., *supra* note 8, at 591 (arguing dynamic factors “must be included” in risk assessment).

are measured reliably.²³³ The argument for including dynamic risk factors is further undercut by the relatively undeveloped and untested theoretical basis for measuring such dynamic factors.²³⁴ Even if we assume that such factors can be accurately and reliably measured, there is reasonable reluctance to connect an individual's behavior while incarcerated to his or her behavior while integrated into society.²³⁵ Incarceration is a controlled atmosphere and very different from the outside world, so some experts question whether in-custody behavior gives a reliable read on how an offender will behave once exposed to old temptations, habits, and associates.²³⁶ In the real world, the exclusion of dynamic risk factors may make little difference. According to a 2007 survey conducted by the state of Connecticut, the majority of state parole boards use instruments that include only one or no dynamic factors.²³⁷ For the foregoing reasons, very little predictive power is lost if states use only information that is available at sentencing.²³⁸

Moreover, because the procedural protections at these stages are different, it is fairer to defendants to limit the use of risk assessment to sentencing. An individual's opportunity to challenge the results or admission of a score yielded from the use of a risk assessment instrument is dramatically different depending on whether that person is a defendant facing sentencing or an inmate hoping for parole release.²³⁹ In crafting a policy for appropriate use of risk assessment, policymakers must balance the social value of preventing crime and the potential unfairness to defendants resulting from the very serious limitations of risk assessment.²⁴⁰ Allowing defendants to avail themselves of the procedural protections available at the sen-

233. See, e.g., Reitz, *Indeterminate Sentencing Model*, *supra* note 20, at 280–81.

234. See Gendreau et al., *supra* note 8, at 576 (expressing concern at the lack of empirical focus on dynamic factors).

235. O'Leary & Glaser, *supra* note 4, at 157–58.

236. See *id.*

237. RATANSI & COX, *supra* note 112, at 10–11 (2007).

238. See CHAMPION, *supra* note 22, at 243 (arguing that better success predictors come outside of prison).

239. See, e.g., *Franklin v. Shields*, 569 F.2d 784, 801 (4th Cir. 1977) (*per curiam*) (finding inmates are not entitled to personal parole hearings, access to their own files, or the ability to call witnesses to appear in their behalf).

240. See NAT'L CTR. FOR STATE COURTS & THE VA. CRIMINAL SENTENCING COMM'N, *supra* note 26, at 29–30 (describing the policy rationale behind selecting a risk threshold for diversion).

tencing stage appropriately achieves this balance. This is so even without the statutory innovations described below.

The single most important procedural check on the use of risk assessment is the availability of a defense attorney at sentencing.²⁴¹ At sentencing, the defense lawyer can advocate to obtain the statistical risk assessment instrument, the instrument's underlying theory and methodology, and the defendant's risk assessment report.²⁴² Without this kind of access, the defendant could not even begin to challenge the accuracy or appropriateness of the risk assessment.²⁴³ Once a defense attorney has access to information related to the risk assessment instrument, he can carefully examine whether the assumptions underlying the statistical model are justified and check whether the instrument has been validated in the local population. Similarly, an attorney can gather information about whether the instrument relies on inputs that are particularly difficult to measure and whether the method in which the instrument was administered might damage its reliability. If his investigation raises concerns, an attorney can draw the court's attention to them and advocate for the assessment to be considered in light of its weaknesses.

Even if a state chose not to adopt a statute establishing a minimum level of reliability for risk assessment instruments, a skillful defense lawyer could still effectively advocate for his or her client by articulating a constitutional challenge to particular factors included in the assessment.²⁴⁴ From there, a defense lawyer might argue that because constitutionally compliant instruments have limited predictive power, the instrument is not reliable enough to be admitted at all. Although the only court to examine the issue has found that statistical risk assessment is appropriate at sentencing,²⁴⁵ the defendant in that case did not specifically challenge the instrument's reliability or its consti-

241. See *Townsend v. Burke*, 334 U.S. 736, 741 (1948) (finding that a lawyer can "take steps" to protect the defendant's due process rights).

242. See *Wolff*, *supra* note 42, at 1408–09 ("Risk assessment methodology—whatever its components—ought to be shared among . . . prosecutors, defense attorneys, probation officers, judges, prison officials, parole boards, and parole officers.").

243. See *Sachsenmaier & Lally*, *supra* note 135, at 27 (detailing the process of verifying risk assessment accuracy).

244. As discussed in Part II.C, the inclusion of some factors in risk assessments, including gender and close proxies for race and gender, is likely unconstitutional. See MODEL PENAL CODE: SENTENCING § 6B.09 reporter's note (Tentative Draft No. 2, 2001).

245. *Malenchik v. State*, 928 N.E.2d 564, 573 (Ind. 2010).

tutionality.²⁴⁶ Instead, that defendant argued that the judge failed to provide an individualized sentence by relying too heavily on the risk score.²⁴⁷ It is certainly possible, in the absence of a statutory reliability standard, that a court would find risk assessment admissible because it is objectively more reliable than the psychiatric prediction of future dangerousness²⁴⁸ held admissible at sentencing by the United States Supreme Court.²⁴⁹ Even so, admissibility decisions involving scientific evidence are notoriously unpredictable²⁵⁰ and so it is conceivable that a court, prompted by a defense attorney, might find that a statistical risk assessment instrument is so unreliable that it is inadmissible.²⁵¹ Even if the court ultimately admitted the risk score, a skillful attack on its reliability or constitutionality might draw the court's attention to the limitations of the instrument and lead it to afford the risk score less weight.²⁵²

At sentencing, the defendant also has the advantage of the opportunity to appeal the sentence²⁵³ and a qualified finder-of-fact.²⁵⁴ There is ample evidence indicating that people tend to over-rely on scientific evidence such as statistical risk assessment because it appears objective and conclusive.²⁵⁵ Judges are

246. See Brief of Appellee at 10, *Malenchik v. State*, 928 N.E.2d 564 (Ind. 2010) (No. 79S02-0908-CR-365).

247. *Id.*

248. See Palmer, *supra* note 139, at 8 (stating that "it is generally accepted" that statistical prediction is more reliable than clinical prediction).

249. *Barefoot v. Estelle*, 463 U.S. 880, 898 (1983).

250. See Cassandra Welch, *Flexible Standards, Deferential Review: Daubert's Legacy of Confusion*, 29 HARV. J.L. & PUB. POL'Y 1085, 1086 (2006).

251. See Eric S. Janus & Robert A. Prentky, *Forensic Use of Actuarial Risk Assessment: How a Developing Science Can Enhance Accuracy and Accountability*, 16 FED. SENT'G REP. 176, 177 (2004) (arguing admissibility decisions are unpredictable "because 'reliability' is a continuous variable (no method is perfectly reliable), and the evidentiary tests do not specify how much reliability is required. It is the underlying substantive law—the legal context—that determines how much reliability is necessary").

252. See *State v. Woomer*, 299 S.E.2d 317, 320 (S.C. 1982) (opining that the defendant's objection to the reliability of the evidence was relevant to the weight the evidence ought to receive).

253. *E.g.*, *United States v. Kaba*, 480 F.3d 152 (2d Cir. 2007) (challenging the constitutionality of a sentence).

254. See *Gardner v. Florida*, 430 U.S. 349, 360 (1977).

255. This phenomenon has been particularly explored in the context of the admissibility of scientific evidence at trial under *Daubert v. Merrell Dow Pharm.*, 509 U.S. 579 (1993). Although the *Daubert* test is inapplicable at sentencing hearings, the academic critique of that test also applies at sentencing. See Paul S. Milich, *Controversial Science in the Courtroom: Daubert and the Law's Hubris*, 43 EMORY L.J. 913, 920 (1994) (arguing that no matter what the

qualified to understand evidence of all sorts, and so are well-suited to handle the challenge of reasonably weighing the value and limitations of statistical risk assessments. In particular, judges have extensive experience with expert scientific evidence.²⁵⁶ In most legal contexts, judges are trusted to be the gatekeepers for all scientific evidence,²⁵⁷ so interpreting the risk scores produced by risk assessment instruments fits neatly in the judicial wheelhouse.²⁵⁸

By contrast, at parole, an inmate is unlikely to be able to meaningfully object to the use of a statistical risk assessment instrument. The inmate's likely pro se status and the lack of an admissibility test at parole mean that an inmate will rarely have the professional support necessary or the legal hook required to challenge the reliability of a risk assessment instrument.²⁵⁹ Along these same lines, the lack of formal process and legal representation mean that an inmate is unlikely to be able to argue a risk assessment instrument violates the Constitution, even though the prohibition on consideration of certain characteristics applies just as much at parole as it does at sentencing.²⁶⁰ Even if an inmate manages to overcome these obstacles and effectively express an objection to the use of risk assessment, the inmate is at the mercy of the members of the parole board who are much less qualified than judges to assess the quality of complex evidence.²⁶¹ Furthermore, inmates at parole do not generally have a meaningful avenue for appeal.²⁶²

Recognizing that statistical risk assessment has real value to protect the public, reduce expenditures, and divert low-risk

standard of admissibility, judges will defer to "what certain experts think").

256. Sophia I. Gatowski et al., *Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-Daubert World*, 25 LAW & HUM. BEHAV. 433, 434 (2001) (noting that "judges are central and active figures in admissibility decision-making").

257. *See id.*

258. *See* MODEL PENAL CODE: SENTENCING § 6B.09 cmt. c (Tentative Draft No. 2, 2011) (arguing that the appropriate interpretation of risk assessment instruments "will reside in the discretion of the trial judge").

259. *See id.* § 6B.09 cmt. a (Tentative Draft No. 2, 2011).

260. *See* *United States v. Kaba*, 480 F.3d 152, 156 (2d Cir. 2007) (finding consideration of race impermissible at sentencing); *Tex. Supporters of Workers World Party Presidential Candidates v. Strake*, 511 F.Supp. 149, 155 (S.D. Tex. 1981) (finding a parole board may not consider race in making its release decision).

261. *See* *Andrews et al.*, *supra* note 216, at 26.

262. *See* *Reitz, Indeterminate Sentencing Model*, *supra* note 20, at 277–79 (describing the wide discretion granted to parole boards).

offenders from incarceration, it is good public policy to make use of the technology. The array of procedural protections available to defendants at sentencing, and unavailable to inmates seeking parole release, all point to using statistical risk assessment exclusively at sentencing. The requirement of counsel, evidentiary standards, and oversight by qualified trial judges combine to provide a forum that is well-equipped to manage the advantages and limitations of statistical risk assessment.²⁶³

C. PROPOSED ADDITIONS TO THE MPC

Because statistical risk assessment is unreliable in some circumstances²⁶⁴ and the stakes are high in this context—no less than the difference between freedom and incarceration—the relatively minimal evidentiary reliability standard required by the Constitution²⁶⁵ may be inadequate. Although the MPC's suggestion that “sufficiently reliable” risk assessment instruments ought to be used at sentencing is sound given the potential upsides of using risk assessment, policymakers and defendants may benefit from a statutory definition of “sufficiently reliable.”²⁶⁶

In particular, the MPC should specify that a risk assessment instrument must be based on analyses of the statistical relationships between objective criteria and recidivism in a sample drawn from the population of the state in which the instrument is to be used or in a substantially similar population. Requiring objective criteria instead of subjective criteria reduces the risk of error associated with inexpert administration²⁶⁷ and low inter-rater reliability.²⁶⁸ Similarly, by using an instrument based on the state's population or a population similar to it, states can avoid inaccuracies in risk scores resulting from a poor match between the population used in the underlying analysis and the population in which the state actually uses the instrument.²⁶⁹ In either case, the instrument should be validat-

263. See MODEL PENAL CODE § 6B.09 cmt. a (Tentative Draft No. 2, 2011).

264. See, e.g., Palmer, *supra* note 139, at 7.

265. See, e.g., People v. Owens, 464 N.E.2d 262, 271 (Ill. 1984).

266. Cf. Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557, 621 (1992) (arguing that when an issue arises frequently, it is reasonable to anticipate at the outset how to resolve the question by adopting a specific rule).

267. See Flores et al., *supra* note 121, at 526.

268. AUSTIN ET AL., *supra* note 27, at 15.

269. See Lowenkamp et al., *supra* note 147, at 560.

ed periodically in the state's population to assure that it remains effective even when the state's population changes.²⁷⁰ In addition, the MPC should specify that sentencing judges considering the results of a risk assessment instrument must be provided with the confidence intervals associated with the results and a copy of any assessment that has been conducted.

These two additions to the MPC's draft section 6B.09 would enhance the section's purpose by providing structure to the delicate balance between protecting public safety and responsibly stewarding public funds and protecting defendants from unfair use of risk assessment. Although the MPC is not binding in any jurisdiction, it is highly influential: over two-thirds of states have adopted at least some part of the Code with some states adopting it nearly in full.²⁷¹ The proposed revision to the MPC, if adopted, is similarly likely to influence state legislatures developing their own policies on the responsible use of statistical risk assessment.

CONCLUSION

Statistical risk assessment has social value because it prevents crime by identifying high risk offenders and reduces societal costs by diverting low-risk offenders from incarceration. Although risk assessment instruments are presently used by both judges and parole boards, states should, as a recent draft of the MPC suggests in section 6B.09, limit the use of sufficiently reliable instruments to the sentencing stage. Doing so would allow society to reap the benefits of risk assessment and would control against inappropriate use of such instruments by allowing defendants to avail themselves of the procedural protections available at sentencing. Because statistical risk assessment instruments have limited predictive power and are susceptible to improper use, the MPC should modify section 6B.09 to include a definition of "sufficiently reliable," which states should, in turn, adopt. Setting a statutory standard for minimally reliable risk assessment instruments would allow states to ensure a fair balance between the social value of preventing recidivism by predicting risk and protecting defendants by refraining from unnecessarily labeling them as high risk based on uncertain statistics.

270. See Schlager & Simourd, *supra* note 154, at 546 (revealing through a validation study the limitations of the LSI-R in certain populations).

271. Paul H. Robinson, *A Functional Analysis of Criminal Law*, 88 NW. U. L. REV. 857, 858 n.3 (1994).