PREDICTING, UP AND DOWN:
A FRAMEWORK FOR LEGAL PREDICTION

*Joseph Avery**

ABSTRACT

Legal prediction has long been a feature of lawyers' and judges' decision making, not least in diversity cases and stay cases, and it has long formed a foundation for thinking about fundamental legal philosophical questions. Recall H.L.A. Hart, Roscoe Pound, and others debating Justice Oliver Wendell Holmes Jr.'s pronouncement that "[p]rophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law." With advances in computer science, new methods for making theories of prediction into reality are being implemented; thus, in addition to retaining its centrality in legal philosophical debates, legal prediction is waxing in practical significance. This Article provides the results of an empirical survey (nearly 200 currently sitting U.S. judges) of judicial thought on legal prediction, and it uses these results, along with an analysis of legal prediction cases in U.S. law, to argue that current legal scholarship has missed an important distinction in legal prediction problems that should be pulled apart: predicting up (What will the agent of a subsequent legal decision do? That is, what will the prosecutor decide? How will the court rule?) and predicting down (What will the object of the legal decision do? That is, will the defendant show up for court? Will he recidivate?). An example of a court predicting up is Barnette v. West Virginia State Board of Education, *where a lower court anticipated a shift in higher court thinking. An example of a court predicting down is* Miller v. Alabama, *where the Supreme Court limited life without parole sentences to only those juveniles who are considered "irreparably corrupt." There are unique issues that inhere to these distinct classes of prediction problems, such as technological-legal lock-in, the need for lay connection to legal processes, and the risk of racial bias.*

TABLE OF CONTENTS

INTRODUCTION

There is a deep debate that courses through legal scholarship, reflected in its various forms in arguments surrounding transcendentalism and functionalism,[1] positivism and realism,[2] the role of judges,[3] and whether law has some constraining force[4] or is lost to nihilism.[5] Surprisingly, amidst this debate from which we may have emerged as post-Realists,[6] there exists a parallel debate regarding the importance of legal prediction, a topic that has been controversial since it was brought to prominence by Justice Oliver Wendell Holmes Jr.[7] As entrée into the debate, let us consider an example.

---

1   Felix S. Cohen, *Transcendental Nonsense and the Functional Approach*, 35 COLUM. L. REV. 809 (1935).

2   Brian Leiter, *Legal Realism and Legal Positivism Reconsidered*, 111 ETHICS 278 (2001).

3   Michael C. Dorf, *Prediction and the Rule of Law*, 42 UCLA L. REV. 651, 655 (1995).

4   H. L. A. HART, THE CONCEPT OF LAW 141–42 (1961) (judges "are parts of a system the rules of which are determinate enough at the centre to supply standards of correct judicial decision").

5   Dorf, *supra* note 3, at 659 n.27.

6   *Id*. at 652.

7   Oliver W. Holmes, *The Path of the Law*, in COLLECTED LEGAL PAPERS 167, 173 (1920). For controversy and arguments surrounding the topic, see Dorf, *supra* note 3, at 652.

When a startup owner is faced with a legal question, does she, in seeking advice, merely want to predict the outcome of subsequent legal review of the behavior? Or is she concerned with a less practical matter: whether she has, in absolute or theoretical terms, violated applicable law? In addition, she might wonder if there is any space between potential answers to these two questions.

This debate impinges upon legal pragmatism, embodied perhaps best in the personage of Roscoe Pound, but also embodied by Justice Holmes, whether or not he personally embraced the term "pragmatism": "[P]rophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law."[8] This definition of the law has become known as the "prediction model."[9]

Staying with the above example regarding the startup owner, the idea is that, rather than wonder whether she has violated the statute in question (the answer to such a query might require review of the text of the statute, its legislative history, judicial decisions surrounding it, and so on), she might merely be wondering whether she is subject to censure. In other words, there is no metaphysical question about the law and infractions of it; there is only her utterly practical question: "Will my company be found guilty of violating the law?" This question, of course, involves prediction. What does one predict the legal decision maker (most likely a judge) will decide?

H. L. A. Hart vigorously pushed against Justice Holmes's proposed conception of the law, arguing that final authority does not create law by virtue of its authority; the law exists independently of the authority figure.[10] Because this point is so important, it is worth rehashing the analogy that Professor Hart made. He described an athletic competition in which the players initially enforced the rules themselves.[11] We might imagine a game of pickup basketball. Then, he had us imagine that a referee was introduced. Now, with the referee installed, would it be correct to reduce the rules to whatever the referee says they are? Surely not, argued Professor Hart. Just as the players once applied the rules, the referee now applies them. At one extreme, a nihilist might claim that the rules do not constrain the referee in any way; at the other extreme, one might claim that the rules, when

---

8   Holmes, *supra* note 7, at 173.

9   Dorf, *supra* note 3, at 653.

10   Hart, *supra* note 4, at 138–44.

11   Hart, *supra* note 4, at 138–42.

sufficiently clear, completely constrain the referee. The truth is certainly somewhere in the middle, and that has led to much consternation in legal scholarship on the topic.[12]

Largely on account of Professor Hart's argument, the prediction theory often is deemed applicable only to attorney decision making.[13] Jerome Frank took this position in his well-known 1931 article in which he emphasized that legal rules play a relatively minor role in case resolution.[14] As Murphy and Coleman wrote, "[I]t is really a cheap shot to take a programmatic remark such as 'law is a prediction of what courts will decide' . . . . Legal realism is, in large measure, the *lawyer's* perspective."[15] But still the point made by those antagonistic to prediction theory is rather astute; yes, it may be well for attorneys to predict how judges will rule, and perhaps we might classify this as "law" of some sort, for surely it is, but it makes no sense to extend this much further. After all, one would not argue that judges predict how they themselves will decide cases. As Professor Dorf wrote, "A court of last resort cannot sensibly 'predict' how it will rule."[16]

This is where the present article begins. With advances in artificial legal intelligence, prediction now can penetrate the work of all legal actors, including judges—and not just in diversity or stay cases. Thus, legal prediction, while retaining its centrality to debates about what the law is and what it is that lawyers and judges are doing when they practice the law, also grows in significance, as old limitations are cast aside and new computational methods for making theory reality are implemented. In this article, I provide an accounting for and explanation of legal prediction, and I do so as follows. In Part I of this article, I seek "reflective equilibrium" across ten cases that are pivotal in debates about the role of prediction in the law.[17] After stating

---

12 JEFFRIE G. MURPHY & JULES L. COLEMAN, PHILOSOPHY OF LAW: AN INTRODUCTION TO JURISPRUDENCE 35 (1990) (arguing that the insights of legal realism are mainly negative and destructive, rather than constructive).

13 LAURENCE H. TRIBE & MICHAEL C. DORF, ON READING THE CONSTITUTION 95 (1991) ("[A]lthough practicing lawyers may be wise to look not to what courts say, but what they do, as a means of *predicting* the outcome of a case, this is hardly an acceptable method for a judge to use in *deciding* a case.").

14 Jerome Frank, *Are Judges Human? Part One: The Effect on Legal Thinking of the Assumption that Judges Behave Like Human Beings*, 80 U. PA. L. REV. 17 (1931).

15 Murphy and Coleman, *supra* note 12, at 35.

16 Dorf, *supra* note 3, at 659.

17 For an overview of the methodology of reflective equilibrium, see JOHN RAWLS, A THEORY OF JUSTICE 48–51 (1971).

the cases and their resolutions, I consider possible explanatory principles. The desiderata that I set for a sound analysis are not only that the case results are explained, but that the role of prediction in U.S. law is also explained through the cases.

From this foundation, I present original empirical research exploring judicial conceptions of legal prediction, including the extent to which judges predict in their own acts of adjudication (Part II). This empirical research, which includes nearly 200 currently sitting U.S. judges as participants, provides important results regarding, *inter alia*, the workability of the *Miller v. Alabama* holding[18] and the usefulness of risk assessment tools.

In Part III, I develop a new framework for thinking about prediction in the law. First, I identify how legal prediction problems should be classified as predicting up (What will the prosecutor decide to do? What will a court decide?) or predicting down (Will the defendant show up for court? Will he recidivate while out on bail?). An example of a court predicting up is *Barnette v. West Virginia State Board of Education*, where a lower court anticipated a shift in higher court decision making.[19] An example of a court predicting down is *Miller v. Alabama*.[20] By limiting life without parole sentences to only those juveniles who are considered irreparably corrupt, the Court is asking sentencers to predict whether a juvenile will be a danger decades down the road and after a long prison sentence.[21]

I construct the framework so that it is useful for thinking about legal prediction tools in these two contexts and how responses to legal prediction ought to differ based on the class of prediction being undertaken. I outline problems that inhere to these distinct classes of legal prediction, including technological-legal lock-in[22] and the non-trivial need for lay connection to legal processes and outcomes.[23] Overall, I show that this new framework is necessary for addressing legal prediction problems. For instance, in the specific area of bias in the law, the failure to properly classify legal prediction problems has led to a lack of nuance in concluding when legal prediction

---

18	567 U.S. 460 (2012).

19	47 F. Supp. 251 (S.D.W. Va. 1942), *aff'd*, 319 U.S. 624 (1943).

20	567 U.S. 460 (2012).

21	Mary Marshall, Miller v. Alabama *and the Problem of Prediction*, 119 COLUM. L. REV. 1633 (2019).

22	Rebecca Crootof, *Cyborg Justice and the Risk of Technological-Legal Lock-in*, 119 COLUM. L. REV. ONLINE 233 (2019).

23	Roscoe Pound, *Mechanical Jurisprudence*, 8 COLUM. L. REV. 605 (1908).

might be harmful and lead to entrenched biases and when it might result in the opposite effect, that is, the lessening of bias in legal outcomes.

In Part IV, I show how the above-mentioned legal philosophical concerns manifest in this updated account. Drawing on Rudolph Ritter von Ihering's work, where he argued for a shift from *Begriffsjurisprudenz* (law of concepts) to *Wirklichkeitsjurisprudenz* (a law of outcomes/results),[24] I discuss how this bifurcation mirrors the divide between symbolic computational reasoning and more data-driven forms, such as those that make use of deep learning. The question is, do we predict as if we are drawing from "a rigid scheme of deductions from *a priori* conceptions,"[25] or do we rely instead on a Poundian notion of a law of principles.[26] If the latter, how can we advance neuro-symbolic AI so that it can accommodate Pound or Holmes's pragmatism? I explore the outer reaches of this approach.

Throughout this article, we might think of there being four overarching questions: in the context of the prediction model today, what are legal decision makers doing, what should they be doing, what are they likely to do moving forward, and what legal implications and considerations does their behavior generate? I take up all of these questions, including the second, although my discussion of the second is largely focused not on whether legal prediction should be pursued, but whether, when it is undertaken, what it ought to or might look like. In all, this article provides an unpacking of legal prediction, forming a cornerstone to support both theory-driven and applied thinking on the subject.

I. A LEGAL PREDICTION CANON

In Part I of this article, I seek "reflective equilibrium" across ten cases that I deem central to discussions on the role of prediction in the law.[27] After stating the cases and their legal resolutions, I consider possible explanatory principles. The desiderata that I set for a sound analysis are not only that I explain the results but that I explain the role of prediction in U.S. law. Moreover, in culling ten cases on which to focus, readers might be surprised that some of these cases are well-known ones, not the fringe or hyper-recent

---

24   RUDOLPH RITTER VON JHERING, LAW AS A MEANS TO AN END (1877; trans. 1999).

25   Pound, *supra* note 23, at 608.

26   *Id.* at 613.

27   *See* Rawls, *supra* note 17.

matters that one might expect to find in an article that considers artificial legal intelligence, predictive analytics, and other rather cutting-edge topics. However, I will show that these ten cases are what one needs to understand the philosophical and practical considerations that attach to the issue of the legal prediction, both for considering its history and for anticipating its near future.

So where and when does legal prediction emerge?[28] I will answer this question more fully in the discussion that follows the presentation of the ten cases. But, for now, to the knowledgeable reader, a few clear examples of judicial prediction should come to mind. To start, judges sometimes undertake prediction in diversity cases and in other similar matters.[29] In these cases, the federal court undertakes a specific prediction task: determining how the state court would rule. Stay cases are another good example. In such matters, for the purpose of quick decision making and workload reduction,[30] a single Justice is tasked with deciding whether to grant a stay of a lower court's judgment. The decision requires the Justice to "forecast whether four Justices would vote to grant certiorari when the petition is presented, predict the probable outcome of the case if certiorari were granted, and balance the traditional stay equities. All of this requires that a Justice cultivate some skill in the reading of tea leaves . . . ."[31] Thus acting as a surrogate for the Court, this solitary Justice likely must, as Professor Dorf has written, "hazard a prediction as to what her colleagues would do."[32]

---

[28]   Evan Caminker, *Precedent and Prediction: The Forward-Looking Aspects of Inferior Court Decisionmaking*, 73 TEX. L. REV. 1 (1994); Earl M. Maltz, *The Concept of the Doctrine of the Court in Constitutional Law*, 16 GA. L. REV. 357, 399 (1982) (discussing lower courts' obligation to "replicate the result that would be reached if the Supreme Court were faced with the same set of facts and allegations"). For an argument that prediction theory is inappropriate in diversity cases and that support for it rests on a flawed reading of Erie Railroad v. Tompkins, 304 U.S. 64 (1938), see Dorf, *supra* note 3, at 655 (arguing that Erie is about federalism, not legal realism).

[29]   *See* 28 U.S.C. § 1367 (Supp. V 1993) (authorizing supplemental jurisdiction); 28 U.S.C. § 1332 (1988) (authorizing jurisdiction in diversity cases); 28 U.S.C. § 1652 (1988) ("The laws of the several states, except where the Constitution or treaties of the United States or Acts of Congress otherwise require or provide, shall be regarded as rules of decision in civil actions in the courts of the United States, in cases where they apply."); Bishop v. Wood, 426 U.S. 341, 343 (1976) (involving a federal question case in which the federal court is tasked with deciding whether state law recognizes a property interest).

[30]   Lois J. Scali, *Comment, Prediction Making in the Supreme Court: The Granting of Stays by Individual Justices*, 32 UCLA L. REV. 1020, 1026–27 (1985).

[31]   Bd. of Educ. v. Superior Ct., 448 U.S. 1343, 1347 (1980) (Rehnquist, Circuit Justice).

[32]   Dorf, *supra* note 3, at 691.

There are other instances, as well, which I will discuss below, especially instances of prediction by non-judicial legal decision makers. But, without further ado, let us consider the ten cases that I deem essential for understanding legal prediction.

A. *The Cases*

Miller v. Alabama[33]

In summer 2003, Evan Miller and an accomplice murdered Cole Cannon.[34] Miller was 14 years old at the time; a year later, he was transferred from juvenile court to a county circuit court so that he might be tried as an adult.[35] In 2006, he was indicted and, later, found guilty and sentenced to life imprisonment. Miller filed a post-trial motion in which he argued that life imprisonment is cruel and unusual punishment, in violation of the Eighth Amendment, when the person subject to the punishment is merely 14 years old.[36]

In reviewing this matter, along with a companion case that had a similar fact pattern and legal posture, the Supreme Court of the United States addressed the question of whether the imposition of a life-without-parole sentence on a 14-year-old violates the Eighth and Fourteenth Amendments' prohibitions against cruel and unusual punishment?[37] The Court ruled that mandatory life without parole sentences for juveniles were unconstitutional unless the juvenile's "crime reflect[ed] irreparable corruption."[38] According to the decision, if a court concludes that a juvenile is incapable of change during and following a long prison sentence, then a life without parole sentence might be appropriate.[39] While this determination was just one that courts were charged with making and/or weighing—"an offender's age and the wealth of characteristics and circumstances attendant to it" should also be considered[40]—it is the only of the characteristics that requires significant

---

33   567 U.S. 460 (2012).

34   *Id.* at 467–68.

35   *Id.* at 467–69.

36   *Id.* at 469.

37   *Id.* at 470.

38   *Id.* at 479–80 (internal quotation marks omitted) (quoting Roper v. Simmons, 543 U.S. 551, 573 (2005)).

39   Marshall, *supra* note 21, at 1654.

40   Miller, 567 U.S. at 476.

prognostication. Moreover, as Mary Marshall has argued, and as I will vet in Part II of this article, the prediction that *Miller* requires is not just unique; it is very difficult, if not impossible, to make.[41]

State v. Loomis[42]

In early 2013, Eric Loomis was charged with five criminal counts related to a drive-by shooting in Wisconsin.[43] Loomis pleaded guilty to two lesser charges—ones relating to the unauthorized use of a vehicle and attempting to evade a traffic officer—and the remaining charges were dismissed, although they were provided to the court for consideration at sentencing.[44] In sentencing the defendant, the court relied on the COMPAS risk assessment tool.[45] This reliance is why *Loomis* has become a notable case.

COMPAS stands for "Correctional Offender Management Profiling for Alternative Sanctions."[46] It is a "risk–need assessment system . . . that incorporates a range of theoretically relevant criminogenic factors and key factors emerging from meta-analytic studies of recidivism."[47] It provides a prediction as to the likelihood that an offender will recidivate. The inputs for this prediction include both an interview of the offender and criminal history information.[48] COMPAS is furnished by a private company, and its methodology is a trade secret.[49]

Filing a motion for post-conviction relief in the trial court, Loomis argued that the court's reliance on COMPAS violated his due process rights.[50] He argued that the court's use of the tool violated his right to individualization to the extent that it relied on group-level statistics.[51] Loomis also argued that

---

41   Marshall, *supra* note 21, at 1654–64.

42   881 N.W.2d 749 (Wis. 2016).

43   "Brief of Defendant-Appellant," at 1-3, State v. Loomis, Crim. A. No. 2015AP157, 2015 WL 1724741, at 1–2 (Wis. Ct. App. Apr. 2, 2015).

44   Loomis, 881 N.W.2d at 754.

45   "Brief of Defendant-Appellant," at 1-3, State v. Loomis, Crim. A. No. 2015AP157, 2015 WL 1724741, at 9 (Wis. Ct. App. Apr. 2, 2015); *see also* Loomis, 881 N.W.2d at 755.

46   Tim Brennan et al., *Northpointe Inst. for Pub. Mgmt. Inc., Evaluating the Predictive Validity of the COMPAS Risk and Needs Assessment System*, 36 CRIM. JUST. & BEHAV. 21 (2009).

47   *Id.*

48   Loomis, 881 N.W.2d at 754.

49   *Id.* at 761.

50   *Id.* at 756.

51   *Id.* at 757.

the information contained in the report was not accurate.[52] The post-conviction motion was denied by the trial court.[53] After the Wisconsin Court of Appeals certified the appeal to the Wisconsin Supreme Court, the Wisconsin Supreme Court affirmed.[54] I will elide the court's discussion of the accuracy issue, as it is not germane to this article. As for the individualization claim, Justice Shirley S. Abrahamson agreed in her concurring opinion that the predictive tool provided only aggregate data on recidivism risk for groups similar to the offender and did not individualize; however, she wrote that, because the predictive tool was not the sole basis for the decision, the sentencing procedure was sufficiently individualized.[55]

Lessard v. Schmidt[56]

In the fall of 1971, Alberta Lessard was picked up in front of her Wisconsin residence by two police officers and taken to a mental health facility and held on an emergency basis.[57] As the result of rulings from various proceedings, Lessard was confined for 10 days, and this was soon extended an additional 30 days.[58] Lessard brought a class action in federal district court in which she made a number of arguments regarding the constitutionality of Wisconsin's involuntary commitment statute.[59] Most germane to our purposes, however, was the *Lessard*'s court's major substantive finding that the statute could only be constitutional if it required proof of "an extreme likelihood that if the [allegedly mentally ill] person is not confined he will do immediate harm to himself or others."[60] Specifically, the court ruled that civil confinement can be justified in some cases, but a finding of dangerousness must be based upon a "recent overt act, attempt, or threat to do substantial harm to oneself or another."[61]

---

52   *Id.*

53   *Id.*

54   *Id.* at 772.

55   *Id.* at 774–775 (Abrahamson, J., concurring).

56   349 F. Supp. 1078 (E.D. Wis. 1972).

57   Arnold H. Landis, *Civil Commitment of the Mentally Ill: Lessard v. Schmidt*, 23 DEPAUL L. REV. 1276, 1276 (1974).

58   *Id.*

59   Michael J. Remington, *Lessard v. Schmidt and Its Implications for Involuntary Civil Commitment in Wisconsin*, 57 MARQ. L. REV. 65, 68–69 (1973).

60   Lessard, 349 F. Supp. at 1093.

61   *See id.*

Lessard is recognized as the "first comprehensive federal court ruling on the substantive and procedural Constitutional limitations for civil commitment . . . ."[62] But it is more important here for how it deals with the issue of prediction. Part of the aim of involuntary commitment is to serve the best interests of the state and the individual, such that the individual is prevented from committing future acts that stand contrary to that aim. Of course, this is a predictive exercise: will the party subject to the involuntary commitment hearing commit detrimental or harmful acts if left unattended? The *Lessard* court recognized the extreme difficulty of this prediction: it discussed its distrust of predictions of future conduct, stating that commitments based upon such predictions must be "viewed with suspicion."[63] Thus, in light of the deprivations of liberty that are at hazard when such a prediction is made, the court mandated that the evidentiary inputs include recent and actual behavior.[64] In other words, unless an individual has recently committed, attempted to commit, or threatened to commit substantial harm to themselves or others, a prediction of future dangerousness is not warranted.

Erie R.R. v. Tompkins[65]

On July 27, 1934, Harry Tompkins was walking along railroads tracks in Pennsylvania when he was struck by something protruding from a moving railcar.[66] Tompkins filed suit in diversity in federal court in New York, as the Erie Railroad Company was a New York corporation.[67] In considering what duty of care was owed by railroads in situations such as the one in which Tompkins had found himself, federal common law (ordinary negligence) and Pennsylvania state law (wanton negligence) had different standards.[68] Thus, the underlying question concerned which standard ought to be applied: should federal courts sitting in diversity jurisdiction apply state or federal law?[69]

---

62 *See* Thomas K. Zander, *Civil Commitment in Wisconsin: The Impact of* Lessard v. Schmidt, 1976 WIS. L. REV. 503, 504.

63 Lessard, 349 F. Supp. at 1093.

64 *See id.*

65 304 U.S. 64 (1938).

66 *See id.* at 69.

67 *See id.*

68 *See id.* at 70.

69 *See id.* at 71.

A majority of the Court held that federal courts are not permitted to create their own common law for issues that might be construed as properly falling within state law.[70] Justice Louis Brandeis, who wrote for the majority, further argued that applying state substantive law in such situations would lead to greater efficiency, largely by making outcomes more predictable.[71] In all, the ruling was that, in diversity jurisdiction cases, federal courts must apply substantive state law and federal procedural law.[72]

But note well what Justice Brandeis wrote, quoting Justice Oliver Wendell Holmes, Jr.:

> [L]aw . . . does not exist without some definite authority behind it. The common law so far as it is enforced in a State, whether called common law or not, is not the common law generally, but the law of that State existing by the authority of that State without regard to what it may have been in England or anywhere else.[73]

Justice Brandeis even explicitly endorsed Justice Holmes, Jr.'s protestations against there being a transcendental body of law.[74] In essence, Justice Brandeis and the Court were saying that there is no law in an absolute sense that a federal court might justly divine and apply. What is the law in such a diversity action? It is what the highest state court would say it is.

State Farm Mutual Automobile Insurance Co. v. Armstrong[75]

In relaying this case, I begin with the Court's holding: "We predict that the Pennsylvania Supreme Court would not recognize an insurance company's common law right to rescind benefits payable to innocent third parties under an assigned risk policy on the basis of a fraud or material representation made by the policyholder."[76] In short, the federal court explicitly predicted what the state court would do were it faced with the matter.

---

70 *See id.* at 79–80.

71 *See id.* at 74 (noting that the previous doctrine in Swift v. Tyson had not led to uniformity or certainty in outcome).

72 *See id.* at 79–80.

73 *Id.* at 79 (quoting Black & White Taxicab & Transfer Co. v. Brown & Yellow Taxicab & Transfer Co., 276 U.S. 518, 533–35 (Holmes, J., dissenting)).

74 *See id.*

75 949 F.2d 99 (3d Cir. 1991).

76 *Id.* at 100 (emphasis added).

The background facts were as follows. A man (the "policyholder") using the name William J. O'Brian and/or William J. O'Brien had made material misrepresentations on his application for a State Farm Assigned Risk policy.[77] The policyholder, who had taken out the policy as a Pennsylvania driver, was subsequently in an auto accident, as a result of which a third party was injured.[78] This injured party attempted to recover from State Farm via the policyholder's policy.[79] The policyholder could not be found, and the district court entered a default judgment against him.[80] State Farm then sued, seeking declaratory judgment that: (1) the policy issued by State Farm was void; (2) the policyholder had made material misrepresentations that State Farm had relied upon to its detriment; and (3) State Farm did not have a legal obligation to any party related to the void policy.[81] State Farm lost in the lower court, where it was held that State Farm had no common law right to rescind third-party benefits under this automobile insurance policy.[82]

This question—whether Pennsylvania law recognizes the right of an insurer to rescind a fraudulently obtained insurance policy in an action brought by a third party against the fraudulent policyholder—had not been decided by the Pennsylvania Supreme Court, and thus the Third Circuit was tasked with resolving it without Pennsylvania precedent on which to rely.[83] To render this prediction, the Third Circuit court first looked to similar—although clearly distinct—matters decided by the Pennsylvania Supreme Court.[84] These matters contained plurality opinions, concurrences, and dissents, and so the Third Circuit looked at the individual justices on the Pennsylvania Supreme Court, and predicted, one by one, how they might rule in the present matter.[85] For example, the Third Circuit court wrote, "Given his plurality opinion in *Bonnie Beck* and his dissent in *Klopp*, we believe that Justice McDermott would also endorse the view that rescission is not an

---

77   *See id.* (noting that the name on the insurance application did not match the corresponding drivers license number on the application).

78   *See id.*

79   *See id.*

80   *See id.* at 100–01.

81   *See id.* at 100–01.

82   *See id.* at 101.

83   *See id.*

84   *See id.* at 102–03 (discussing two similar cases in which the Pennsylvania Supreme Court had interpreted statutes that governed the voluntary car insurance market).

85   *See id.* at 102–04.

available remedy to Assigned Risk Plan providers."[86] Finally, even admitting that the evidence it had marshalled was "only dicta," the court was confident in its predictive powers: "we find it unlikely that . . . the justices would vote otherwise in this case."[87]

Barnette v. West Virginia State Board of Education[88]

Two years before *Barnette* came before the district court for the Southern District of West Virginia, the Supreme Court of the United States decided *Minersville School District v. Gobitis*.[89] In *Gobitis*, the Court held that public schools could compel students, including Jehovah's Witnesses, who were parties to this matter, to salute the American fag and recite the Pledge of Allegiance, even when the students professed religious objections to these practices.[90] Virtually an identical issue was at issue in *Barnette*, and thus when the case came before the district court, the school board defended itself with a brief that relied heavily on *Gobitis* as controlling authority.[91]

However, Circuit Judge Parker, writing for the three-judge district court, launched into speculation about how the individual Justices on the Supreme Court might vote, aware as he was that the composition of the Supreme Court had changed since *Gobitis* had been decided: Justice Robert H. Jackson had joined the Court, as had Wiley Rutledge.

> Of the seven justices now members of the Supreme Court who participated in that decision, four have given public expression to the view that it is unsound, the present Chief Justice in his dissenting opinion rendered therein and three other justices in a special dissenting opinion in [. . .]. The majority of the court in Jones v. City of Opelika, moreover, thought it worth while to distinguish the decision in the Gobitis case, instead of relying upon it as supporting authority.[92]

As Professor Dorf has pointed out, Judge Parker was not treating the Court as a unified whole; rather, he was surmising why the Justices had voted how they had, inferring that at least one of the five Justices in the *Opelika*

---

86   *Id.* at 103.

87   *Id.* at 104.

88   47 F. Supp. 251 (S.D. W. Va. 1942), *aff'd*, 319 U.S. 624 (1943).

89   310 U.S. 586 (1940).

90   *Id.* at 599–600.

91   *Barnette*, 47 F. Supp. at 252.

92   *Id.* at 253.

majority must have disapproved of *Gobitis* and thus might join the other four Justices who had publicly expressed disapproval of the ruling.[93]

Relying on this prediction, the district court ruled in favor of the plaintiffs and invalidated the statute that made flag saluting compulsory in West Virginia public schools. The prediction proved correct: in its review of the *Barnette* decision, the Supreme Court of the United States overruled its earlier *Gobitis* decision.[94]

United States v. Girouard[95]

In *Girouard*, a Canadian sought U.S. citizenship but was unwilling to bear arms in the country's defense because he identified as a conscientious objector for religious reasons (he was a Seventh Day Adventist). Circuit Judge Mahoney, writing for a majority on the First Circuit Court of Appeals, denied the petition.[96] He cited multiple Supreme Court precedents that explicitly denied citizenship to similarly situated applicants and stated that Congress "has passed no legislation expressly changing the meaning [of the citizenship statute] as interpreted by the Supreme Court . . . ."[97] In short, Judge Mahoney pointed out that there was controlling precedent and that the ruling was thus straightforward to make.

Circuit Judge Woodbury dissented. Acknowledging his "duty as an inferior federal judge to accept and follow controlling decisions of the Supreme Court," Judge Woodbury still felt that his role was one of prediction.[98] He had a "duty to prophesy," since "[n]othing is to be gained by our deciding a question contrary to the way we think the Supreme Court would decide it."[99] In predicting how the Court would rule when faced with this matter, Judge Woodbury went Justice by Justice, detailing how they had decided other cases and especially what they had written as dicta in their dissents.[100] He concluded that "the prediction can be ventured that the

---

93   Dorf, *supra* note 3, at 663.
94   West Virginia State Bd. of Educ. v. Barnette, 319 U.S. 624, 642 (1943) (overruling *Gobitis*).
95   149 F.2d 760 (1st Cir. 1945), *rev'd*, 328 U.S. 61 (1946).
96   *Id.* at 761, 763.
97   *Id.* at 763.
98   *Id.* at 765.
99   *Id*.
100   *Id.* at 765–67.

above cases [i.e., the precedential cases] are no longer expressive of the law."[101]

What did the Supreme Court do? It granted certiorari,[102] and it reversed the judgment of the Circuit Court,[103] proving Judge Woodbury's prediction correct—although the Court did not comment on Judge Woodbury's dissent nor on whether Judge Woodbury was right to tack, as he did, against precedent and in favor of prediction.

In re Roche[104]

In this application for a stay of enforcement of an order finding a reporter who had refused to disclose the identities of his new sources to be in civil contempt, Justice Brennan served as Circuit Justice, tasked with deciding the matter on behalf of the other Justices. The opinion is a mere four pages, and it is worth reading a selection from it to see how Justice Brennan forecast the votes of his fellow Justices. Here is some of the evidence he garnered in powering his prediction: "[T]wo of my Brethren found the prospects for review by the full Court insufficient to warrant staying contempt proceedings against a New York Times reporter . . . ."[105] "Four dissenting Justices in *Branzburg* discerned at least some protection in the First Amendment for confidences garnered during the course of newsgathering."[106] "And Mr. Justice POWELL, who joined the Court in *Branzburg*, wrote separately to emphasize . . . ."[107] With insights such as these into the other individual Justices, Justice Brennan then made his prediction: "[I]t is reasonably probable that four of my Brothers will vote to grant certiorari . . . ."[108]

Rodriguez de Quijas v. Shearson/Am. Express, Inc.[109]

In *Rodriguez*, an investment dispute resulted in claims of unauthorized and fraudulent transactions, some of which related to the Securities Act of

---

101   *Id.* at 767.
102   326 U.S. 714 (1945).
103   Girouard v. United States, 328 U.S. 61, 70 (1946).
104   448 U.S. 1312 (1980).
105   *Id.* at 1315.
106   *Id.*
107   *Id.*
108   *Id.* at 1316.
109   490 U.S. 477 (1989).

1933.[110]   The District Court ordered that those claims that related to the Securities Act could not be submitted for arbitration but must proceed in the court action, as *Wilko v. Swan*[111] directly controlled the matter.[112]   The Court of Appeals, however, reversed.[113]   In reviewing post-*Wilko* decisions by the Supreme Court, the Court of Appeals determined that the Supreme Court had reduced *Wilko* to "obsolescence."[114]   Thus, the Court of Appeals was comfortable breaking with the precedent, as it believed the Supreme Court would overrule it if presented with the opportunity.

The court of appeals was correct in its prediction: on review, the Supreme Court overruled *Wilko*.[115]   But, Justice Kennedy, writing for the majority, explicitly admonished the court of appeals for ruling in accordance with its prediction, even though the prediction proved correct: "We do not suggest that the Court of Appeals on its own authority should have taken the step of renouncing *Wilko*."[116]

In a dissent from the holding, Justice Stevens nevertheless was of one mind with Justice Kennedy: "[T]he Court of Appeals . . . engaged in an indefensible brand of judicial activism."[117]   There is a striking paragraph that follows this pronouncement.   Justice Stevens goes on to give reasons as to why *Wilko* ought not to be overruled and why the argumentation in the present case represents one's beliefs about deference to congressional acts.[118]   "In the final analysis, a Justice's vote in a case like this depends more on his or her views about the respective lawmaking responsibilities of Congress and this Court than on conflicting policy interests."[119]   In other words, there is a sense in which Justice Stevens was telling the court of appeals that, while they may have correctly predicted how the Court would act, they did not correctly predict how Justice Stevens would act, and for that they must be taken to task for having the gall to predict in the first place.

---

110   *Id.* at 478.
111   346 U.S. 427 (1953).
112   *Rodriguez de Quijas*, 490 U.S. at 479.
113   *Id.*
114   *Id.*
115   *Id.* at 485.
116   *Id.* at 484.
117   *Id.* at 486 (Stevens, J., dissenting).
118   *Id.* at 487.
119   *Id.*

Planned Parenthood v. Casey[120]

In *Roe v. Wade*, the Supreme Court of the United States established that abortion is a fundamental right protected by the Due Process Clause of the Fourteenth Amendment and that women have a privacy interest protecting their right to abortion pursuant to the liberty clause of the Fourteenth Amendment.[121] At issue in *Planned Parenthood* was the Pennsylvania Abortion Control Act of 1982, five provisions of which were challenged by the plaintiffs as unconstitutional under *Roe*.[122] These provisions limited how and when individuals might receive an abortion. The District Court for the Eastern District of Pennsylvania held that all the provisions were unconstitutional, but the court of appeals reversed in part, holding that only one of the provisions was unconstitutional.[123] For the purposes of this article, the decision is notable for the extensive discussion of *stare decisis* that it incited. In short, we might wonder, how can the Supreme Court predict its own behavior, and when might it change its mind about its own past decisions?

In a plurality opinion joined by Justices Kennedy and Souter, Justice O'Connor wrote that, while there is an obligation to follow precedent, that obligation is not an inexorable command, and it is not without an outer limit.[124] That outer limit she defined as when "a prior judicial ruling should come to be seen so clearly as error that its enforcement was for that very reason doomed."[125] She went on to elucidate specific circumstances that the Court might weigh when considering whether to reverse itself: does the rule defy practical workability?[126] Would overruling it lead to inequity?[127] Have related legal principles so developed as to render the current case a vestige, an abandoned doctrine?[128] And has there been a significant change in facts such that the old rule does not apply or is not justified in its application?[129] Justice O'Connor concluded by arguing that "doctrinal disposition" alone is not enough: if the 1992 Court has a different doctrinal disposition than the

---

120   505 U.S. 833 (1992).

121   410 U.S. 113, 154–56 (1973).

122   *Planned Parenthood*, 505 U.S. at 844.

123   *Id.* at 845.

124   *Id*. at 854.

125   *Id*.

126   *Id*.

127   *Id*.

128   *Id*. at 855.

129   *Id*.

1973 Court, that is no reason to overturn a decision made by the 1973 Court.[130]

B. *Reflective Equilibrium*

What possible explanatory principles course through these ten cases? The main issues that I will discuss and from which principles will be drawn include the following. First, there is the issue of, as Professor Dorf put it, law as found vs. law as made.[131] Is there something absolute—the law itself— that transcends the decisions of legal actors? Moreover, if that question gets at a normative debate, then what is the practical aspect of it; that is, how are judges actually behaving? Second, there is the issue of what matters when prediction is undertaken—what factors come into play? Third, there is the question of when and where legal prediction is undertaken—who is predicting what?

I begin with the first issue. In *Barnette*, the District Court wrote, "It is true that decisions are but evidences of the law and not the law itself; but the decisions of the Supreme Court must be accepted by the lower courts as binding upon them if any orderly administration of justice is to be attained."[132] This almost perfectly mirrors the analogy that H. L. A. Hart proffered[133] and which I modernized and discussed above. In a pickup game of basketball, one in which there is no referee, the players abide by the rules. Once a referee is introduced, the referee interprets and applies the rules, but that does not mean that the rules should be reduced to whatever the referee says they are.

But consider the following. In Part II of this article, I present the results of an original empirical study in which currently sitting U.S. judges were the participants. One of the judges wrote to me with this thought:

> [I]t was curious that the survey made a distinction between 'the law' and the law as determined by the appellate courts. There is no law, except as determined by the appellate courts. [. . .] I once decided, in a criminal case, that a search was unconstitutional, and granted a motion to suppress. The case was heard by a three judge panel at the Court of Appeals, and then by the nine members of the state Supreme Court. Seven of the appellate court

---

130   *Id.* at 864.

131   Dorf, *supra* note 3, at 709.

132   Barnette v. West Virginia State Bd. of Educ., 47 F. Supp. 251, 253 (S.D.W. Va. 1942), *aff'd*, 319 U.S. 624 (1943).

133   Hart, *supra* note 4, at 138–44.

members said my decision was correct, and five said it was not. Was I correct on the law? No, because the five appellate judges who decided I was wrong were all on the Supreme Court.

Even if this approach seems reasonable, to adopt it is to hazard adopting an untenable position. Recall Circuit Judge Woodbury's dissent in *United States v. Girouard*.[134] There, he pointed out that, though the majority likely was following precedent and adhering to the demands of *stare decisis*, the Court had a duty to be, as my judicial participant stated, "correct," where to be correct is to be in accordance with how the Supreme Court—or one's highest state court—would rule. As we saw in *Girouard*:

> I conceive it to be our judicial duty to decide cases as we think they should be decided, but as an intermediate appellate court one of the factors, and a highly important one, for us to take into consideration in concluding how we should decide a case is the view which we think the Supreme Court would take on the question at issue before us.[135]

Such a predictive approach is only truly apposite when the Court has unequivocally not addressed a specific issue and a lower court is forced to consider an issue of first impression. Otherwise, we know that ruling in accordance with how one predicts the Supreme Court would rule is forbidden: let us remember the rebukes that resounded through the *Rodriguez de Quijas v. Shearson/Am. Express, Inc.* opinion, the Court angry at the appeals court for predicting correctly but in contravention of precedent.[136] After all, much if not all legal thought holds that lower courts must adhere to precedentwithout considering how higher courts may rule on appeal.[137] *Stare decisis* is the term for a judicially-created rule that courts must "abide by, or adhere to, decided cases."[138] "Let stand what is decided, and do not disturb what is settled"—this is the common law doctrine from which the term arises.[139] There is hierarchy to lack of precedentas well: when the Supreme Court of the United States has not issued a decision on a particular issue, the decisions of the court of appeals for a particular circuit are binding on the

---

134   149 F.2d 760, 764 (1st Cir. 1945) (Woodbury, J., dissenting).

135   *Id*. at 765.

136   Rodriguez de Quijas v. Shearson/Am. Express, Inc., 490 U.S. 477, 484 (1989) (overruling a past case and simultaneously admonishing the court of appeals for taking the same step).

137   *See, e.g.,* Caminker, *supra* note 28, at 5–6 (arguing this precise point).

138   *Stare decisis,* BLACK'S LAW DICTIONARY 1406 (6th ed. 1990).

139   "Stare decisis et non quieta movere." Coyne v. Westinghouse Corp. (In re Globe Illumination), 149 B.R. 614, 617 (Bankr. C.D. Cal. 1993), quoting Jeremy C. Moore et al., *Moore's Federal Practice* ¶ 0.402[1] (2d ed. 1992).

lower courts within that circuit.[140]   In fact, what many have failed to acknowledge is that there are two distinct prediction problems: rather than predict how higher courts would rule on the matter at hand if faced with it, lower courts might predict how a higher court will rule on the lower court's decision. This is still a predictive exercise, as the lower court is only interested in the absolute understanding of the law to the extent that it bears upon the decision-making of the higher court.

This issue of law as found vs. law as made is important because a judge who adheres to the prediction model also must discount the importance of law as found. Whether or not there is some "correct" interpretation of the law, the judge would only care about predicting what the higher court thinks the "correct" interpretation is, and such a determination is a far cry from embarking upon an independent search for the truth of the matter. Of course, as the *Barnette* Court pointed out, what the Supreme Court says is binding,[141] so there is a meaningful sense in which predicting what the Supreme Court will do is in fact representative of a search for the truth of the matter. Although we are not so naïve as to pretend that there is no space between the two conceptions, as Hart and as any astute commentary concedes.

Indeed, such space is most evident when we consider the task that the Supreme Court undertakes. There, especially with constitutional matters, *stare decisis* is not absolute. However, it is not an empty doctrine either. As we saw in *Planned Parenthood v. Casey*, the Court may overrule itself, but if it does, it will spill much ink making sure that practitioners and the lay public believe that truly exceptional circumstances (or truly exceptional *changes* in circumstances) have led to the overruling. And yet, even here, even at the Supreme Court, there is still doubt regarding whether there is law as found. Justice Brandeis, let us remember, in *Erie R.R. v. Tompkins*, endorsed Justice Oliver Wendell Holmes Jr.'s protestations against there being a transcendental body of law.[142]

---

[140]   Zuniga v. United Can Co., 812 F.2d 443, 450 (9th Cir. 1987) (emphasizing that District Courts are "bound by the law of their own circuit"); In re Globe Illumination, 149 Bankr. C.D. Cal at 617 (same); In re Taffi, 144 Bankr. C.D. Cal 105, 108 (same).

[141]   Barnette v. West Virginia State Bd. of Educ., 47 F. Supp. 251, 253 (S.D.W. Va. 1942), *aff'd*, 319 U.S. 624 (1943).

[142]   304 U.S. 64, 79 (1938).

Even though we have reached an impasse, if we turn to the practical matter of prediction, we see that judges think less about courts as unified wholes and more about individual justices with idiosyncratic beliefs. We saw this in *State Farm Mutual Automobile Insurance Co. v. Armstrong*, the diversity case in which the Third Circuit, in making a prediction, looked at individual state court justices and tallied votes.[143] We also saw it in *In re Roche*, the stay case in which Justice Brennan, in making his prediction, looked at individual Supreme Court justices and tallied votes.[144]

Given this, what might prediction involve that is missing from the ideal conception of the law? The answer is implicit in the above-described judicial behavior: individual personality is what is missing. In addition to conventional legal materials, a lower court judge seeking to replicate the result that would be reached if the Supreme Court were faced with the same set of facts and allegations likely would consider the views of the individual justices who sit on the high court as a basis for predicting how those particular justices would rule.[145]

Having addressed the normative debate and its practical manifestations, including the emphasis that has come to fall on individual judges rather than on courts, I now can turn to the third issue that is evident in our canon of legal prediction cases. Where and when does legal prediction emerge? The answer is that it tends to work up the legal hierarchy, just as one might climb up a ladder. A lawyer faced with a legal question might predict what the most likely legal or quasi-legal decision maker will decide (jurors, judges, insurance adjusters, or—in the case of plea bargaining—prosecutors, and so on), and he or she might predict what the highest (in the hierarchy) possible decision maker might rule (a state supreme court, or the Supreme Court of the United States). The intermediate legal decision makers (let us call them the "lower courts" for ease of discussion) might predict what their hierarchical counterparts will do—when faced with such a matter and/or

---

143   949 F.2d 99, 103–104 (3d Cir. 1991) (looking at individual Pennsylvania Supreme Court Justices to predict how they would decide an issue).

144   In re Roche, 448 U.S. 1312, 1315–1316 (1980) (predicting Supreme Court Justice votes and concluding "it is reasonably probable that four [Justices] will grant certiorari").

145   Dorf, *supra* note 3, at 654.

when faced with the decision of the lower (in the hierarchy) decision maker's decision.[146]

The lower courts also may predict down the ladder, so to speak, predicting the behavior of legal parties, such as whether defendants pose a danger to society or whether a party, if held liable, will adhere to the court's ruling. In my canon of legal prediction, we saw this play out in *Miller v. Alabama* (predicting future criminality), *Lessard v. Schmidt* (predicting dangerousness), and *State v. Loomis* (predicting down with the aid of automated prediction tools).

Lastly, the highest court might undertake the seemingly nonsensical predictive task of predicting what it might do itself, although we know it is not impossible, as we saw how, in stay cases, a member of the Court may be tasked with predicting the whole of the Court.[147] In other words, the Supreme Court might predict its own rulings, which of course is something akin to the process of following its own precedent, but I discuss that in more detail in Parts III and IV of this article. While this is the general flow of legal prediction, it does not fully capture the constraints under which each party might pursue such predictions. In the figure that follows, and in the succeeding paragraph, I outline both the predictive flow and the constraints on that flow.

---

[146] Note a curiosity that arises here, one that, in actuality, holds promise for addressing issues relating to efficiency, heavy docket loads, and the like. In a tort, if a plaintiff and a defendant both make the same prediction, then there would be a strong incentive to settle, to forgo adjudication. Likewise, if a defendant and a prosecutor both make the same prediction, there would be a strong incentive to forgo trial and reach a plea.

[147] In re Roche, 448 U.S. 1312, 1316 (1980) (predicting "there is a fair prospect that the court will reverse the decision below").
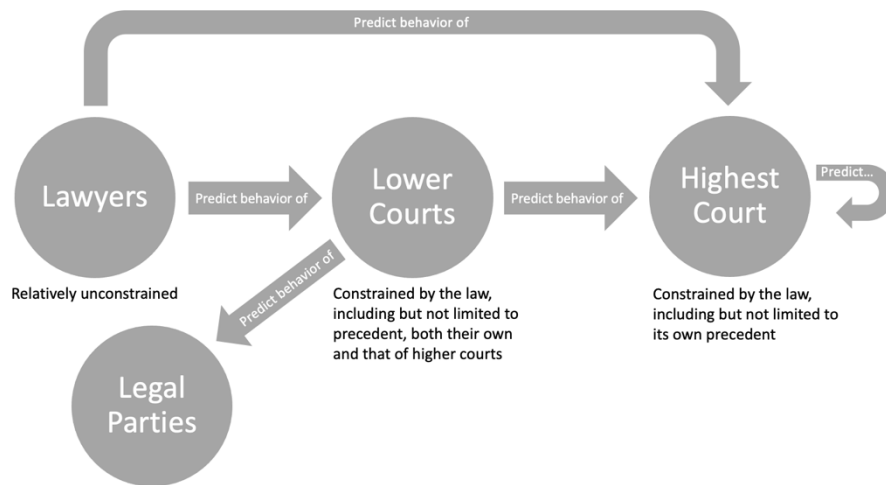
Figure 1. Flow chart of both predictive agents and objects and the forces that constrain the use of prediction-driven decision making. Note that the figure is just an overview and does not include all predictive objects; for instance, the Supreme Court predicts what officials would do in response to incentives proposed legal rules might create, how the public might respond to its rulings, and so on.

For lawyers, on a practical level, it makes sense to predict how the highest court that is likely to review the matter with which it is faced will rule. What the law means in the abstract is largely irrelevant beyond the extent to which a lawyer's reading of the law might sway the decision-making of the highest court that reviews the matter. But lawyers are still constrained in that their predictions must be so reasonable as not to trigger legal malpractice claims by their clients. That is, one might believe that one's interpretation of the law is in perfect accord with the law's true and correct meaning, but if no judge and nary another party agrees with that interpretation, then one arguably has failed at the job of lawyering.

For lower courts, there is some tension in whether they can distinguish the law "in absolute terms" from how higher court judges would rule, given that higher court precedent is binding and controlling. In a sense, as one of the judicial participants from Part II wrote, "Following the law means following precedent that is binding on me." Or, as another judicial participant wrote, "A judge should always look to what the law requires, but

'what the law requires' is defined in many instances by what higher courts have decided." But we can understand, as seen in the cases presented in Part I, that the law might still transcend precedent and higher court rulings, and judges have room to consider the law in and of itself, absent consideration of past or future higher court rulings. For the highest court (we might think of the Supreme Court of the United States), its justices also are constrained by precedent, although outside of diversity matters and other similar matters, that precedent is largely the Court's own (and also not always constraining). Thus, in the vast majority of cases, including cases of first impression, the Court considers the law in absolute terms, and any prediction that is possible would be prediction about how the Court, as then constituted, would rule on the matter at hand. This might seem like an inane or impossible exercise, but I argue that it is not. Indeed, it is precisely the exercise that lower court judges take up, in part, when considering a matter about which there is no clear precedent.[148] When there is clear precedent, the prediction should be easy enough, unless, as discussed in *Planned Parenthood*, there is some "special reason."[149]

II. EMPIRICAL RESEARCH ON LEGAL PREDICTION

In Part II, I present the results of original empirical research in which 181 currently sitting U.S. judges participated. The judges were asked about the ideas that were put forth in Part I of this article. In particular, there are three topic groupings, and the results are presented within these groupings. First, I broached the judges' beliefs regarding the conception of the law as transcending the higher courts vs. the law as virtually equivalent to higher court judicial rulings. In addition to questions about what is required of a lower court judge in terms of following precedent, this also includes considerations of whether, even if one does follow precedent, the likelihood of being overruled weighs upon judicial decision making. Second, I explored judges' beliefs regarding how well they could predict—both up and down the ladder, as described above and shown in Figure 1. In this part of the survey, I gathered important data regarding the feasibility of the legal rule imposed

---

148   And it is what, in stay cases, a single justice does on behalf of the whole Court, in a synecdoche-like version of the type of prediction I am discussing.

149   505 U.S. 833, 864 (1992).

by the *Miller v. Alabama* Court.[150] I also explored which factors the judges considered most important for predicting up, and whether risk assessments are useful to them when they predict down. Third, I queried the judges about their beliefs regarding how well others can predict their (the judges') decision making.

A. Methodology and Results

The survey was approved by Princeton University's Institutional Review Board. One hundred and eighty-one judges completed the full survey.[151] Mean age was 61 years old, almost 75% were male, and 91% were White/Caucasian. Seventy-eight percent were sitting on a state-level trial court, 7% on a state-level appeals court, 2% on a state-level highest court, 7% on a federal district court, 1% on a federal court of appeals, and 5% indicated other, which mostly meant a bankruptcy court. The survey consisted of a series of questions, which I now discuss concurrent to presenting the results.[152]

1. Idealist vs. realist; absolute vs. predictive

The first set of questions concerned the varying conceptions of the law— law as transcending the higher courts vs. the law as virtually equivalent to higher court judicial rulings. Participants were asked, "When a legal matter is being decided, which of the following should be the focus of the judges making the decision: (1) what result the law requires, or (2) how higher court judges would rule on the matter?"[153] As seen in Figure 2, most participants indicated that what result the law requires was relatively more important than how higher court judges would rule.

---

[150] 567 U.S. 460, 465 (2012) (holding "mandatory life [sentence] without parole for those under the age of 18 at the time of their crimes violates the Eighth Amendment's prohibition on 'cruel and unusual punishments'").

[151] The survey was sent via email to 2,155 U.S. judges. The emails were culled through THE AMERICAN BENCH (2018).

[152] For the complete survey questions, along with the order in which they were presented, see Appendix A. For data, see Appendix B.

[153] As is common in the social sciences, I relied on a dichotomy in this question, as it is the dichotomy that I think accurately represents the situation at-hand. Granted, there might be other choices that lurked in participants' minds, and these might be explored in future studies.
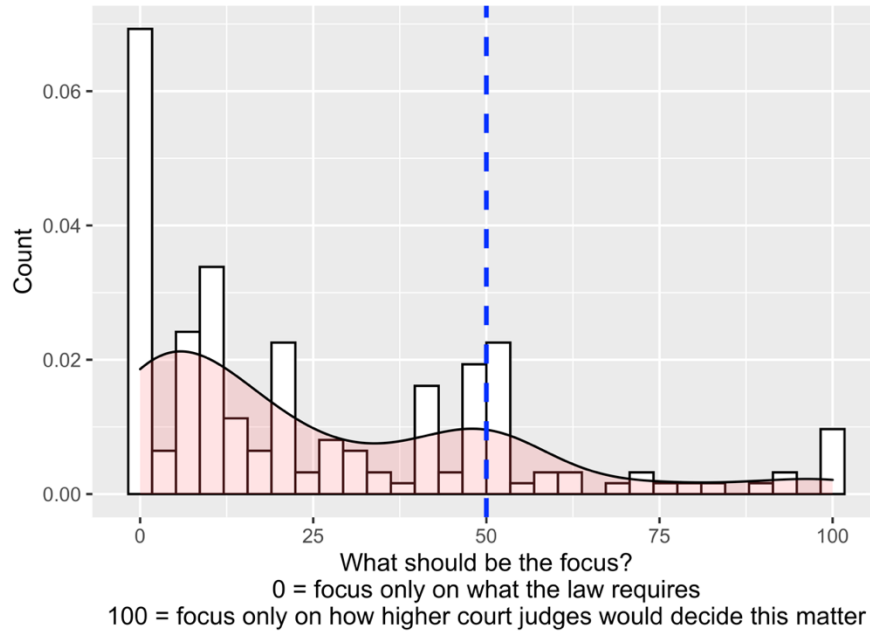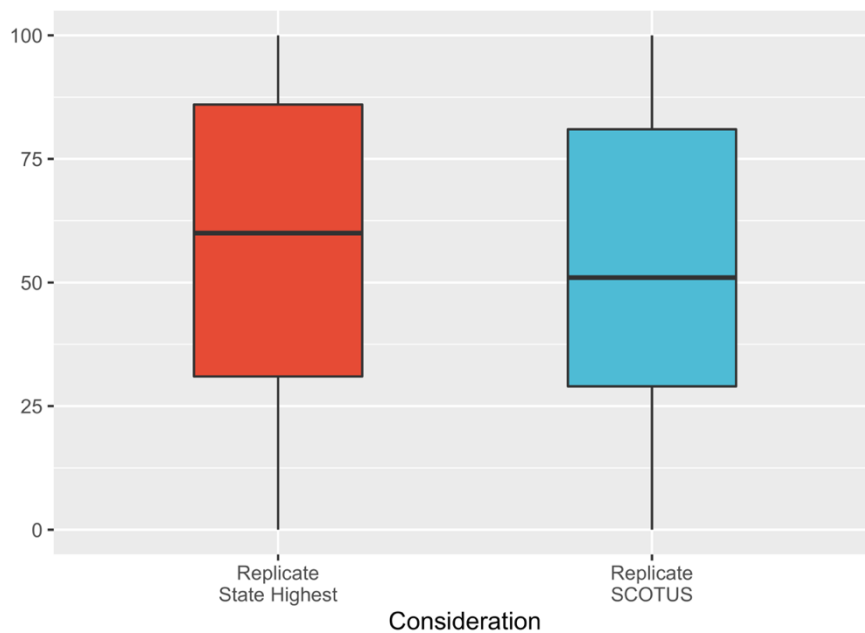
Figure 2. The judicial participants indicated whether judges' foci should be solely on what the law requires (score of 0), solely on how higher court judges would rule (score of 100), or equally focused on both (a score of 50, indicated by the blue dotted line). The y-axis ("Count") indicates the percentage selecting within each bin, and the red shaded region is a density plot, which shows a smoothed distribution of the points along the axis. The mean is 27.

As discussed in Part I, this question can only take us so far in our understanding. After all, what the law requires includes how higher court judges *have* ruled (if there is controlling precedent), and it does not necessarily include how they *would* rule, although the two certainly may overlap (they usually rule in accordance with how they have ruled—but not always). What is interesting, though, is the overall result: regardless of what they do, judges profess that they should shy away from a prediction model of ruling. They indicate that the focus should be more on what the law requires and less on what higher court judges would or would not do if faced with the same matter.

I dove deeper into both ends of this prompt. On the one end, I asked the judges to indicate their agreement with the following statements about predicting higher court decisions:

- "When deciding a legal matter, I attempt to replicate the result that would be reached if my state's highest court were faced with the same matter."

- "When deciding a legal matter, I attempt to replicate the result that would be reached if the Supreme Court of the United States were faced with the same matter."

The first question was analyzed only for state court trial and appeals judges, as it was not applicable to the other participants. The scale ran from 0 (not at all) to 100 (definitely). The results are depicted in Figure 3. In summary, the mean for replicating the state highest court was 56, and the mean for replicating the Supreme Court of the United States was 52. Thus, such replication/prediction does, it would seem, play a role in judicial decision making. While it may not be as important as the law itself (as we saw in Figure 2), it is impactful.

Figure 3. The judicial participants indicated their agreement with the statements "When deciding a legal matter, I attempt to replicate the result

that would be reached if my state's highest court were faced with the same matter" and "When deciding a legal matter, I attempt to replicate the result that would be reached if the Supreme Court of the United States were faced with the same matter," where 0 = not at all and 100 = definitely.
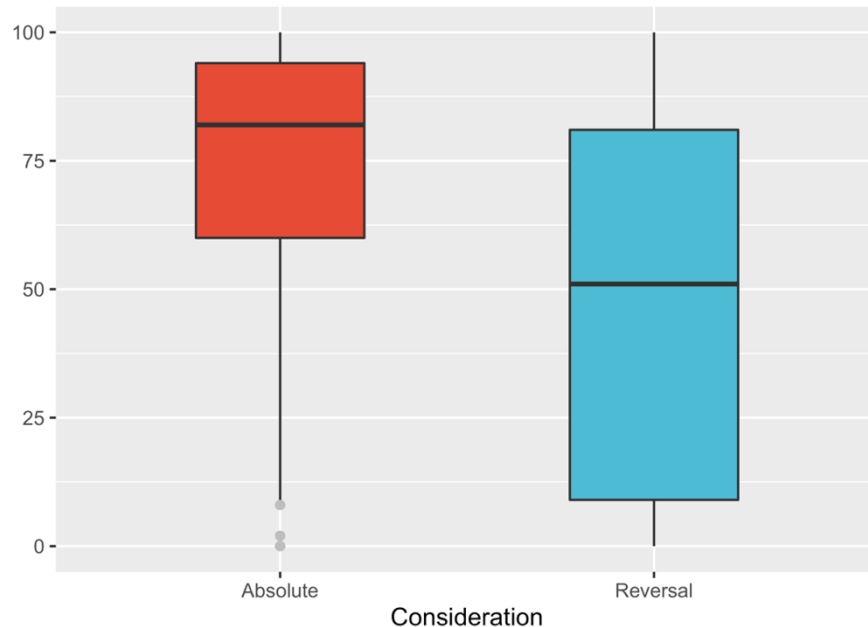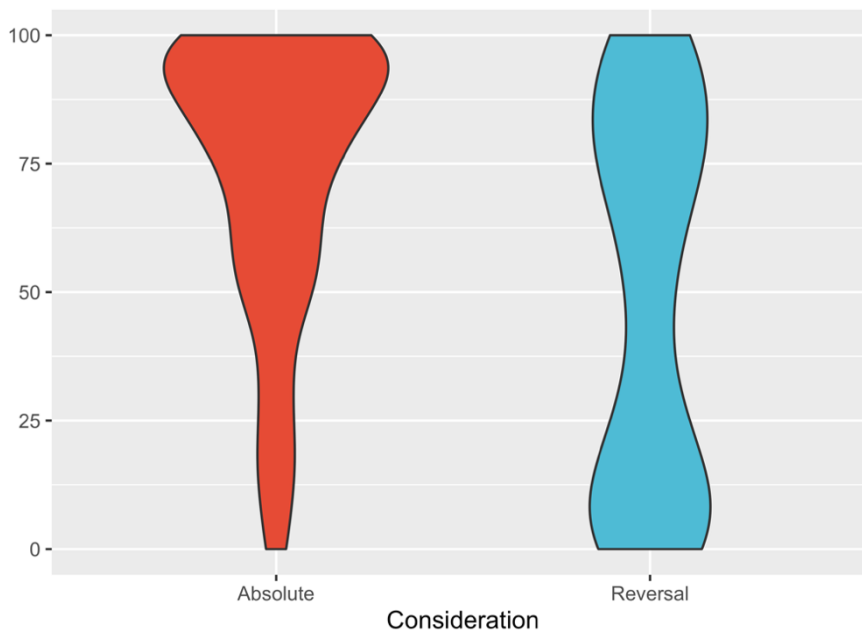
Finally, the judges were again asked about the importance of considering the law in an absolute sense: "When deciding a legal matter, I consider what the law requires in an absolute sense. (Please interpret 'in an absolute sense' as indicating the following: the law transcends the question of how higher court judges would rule if faced with the matter.)." They also were queried about an iteration of what is presented in Figures 2 and 3. Rather than whether they seek to replicate what higher court judges would do if faced with the same matter, they were asked whether they seek to avoid reversal: "When deciding a legal matter, I consider whether my decision would be reversed by a higher court." As before, the scale ranged from 0 (not at all) to 100 (definitely). For the law in an absolute sense prompt, the mean was 73 (with a standard deviation of 27), showing that this is serious concern of judges. For the reversal prompt, the mean was 48 (standard deviation of 37).

Figure 4. "Absolute" refers to "When deciding a legal matter, I consider what the law requires in an absolute sense. (Please interpret 'in an absolute sense' as indicating the following: the law transcends the question of how higher court judges would rule if faced with the matter.)." "Reversal" refers to "When deciding a legal matter, I consider whether my decision would be reversed by a higher court." The scale ranged from 0 (not at all) to 100 (definitely). These data are visualized in two forms. At top is a boxplot; at bottom is a violin plot. The advantage of the violin plot is that it shows the full distribution of the data. That is, it is similar to the boxplot, except that it also shows the probability density of the data at different values. It is included here in order to show the bimodal outcome. For the reversal prompt, the mean and median, as seen in the boxplot, are at about the scale midpoint. However, as seen in the violin plot, there are two clusters, one at the lower end and one at the higher end. In short, the judges either definitely do or definitely do not consider reversal; very few are in the middle. In contrast, for the prompt on the law as transcending the question of how higher court judges would rule if faced with the matter, nearly all judges indicated that this is definitely something they consider when making a ruling.

From this first line of questions, a few insights can be gleaned. Prediction plays a role in judicial considerations, as lower court judges think about both how higher court judges would rule if faced with the same matter and whether higher court judges will reverse them on appeal, although the story for this latter consideration is mixed: some judges deny spending much if any time pondering potential reversal. That said, even more than looking up the hierarchy, what judges care about most is what the law requires in an absolute sense. This may, of course, subsume some element of prediction, as properly applying precedent does require properly understanding how the higher court would rule if faced with the matter, but given the dyadic presentation of the question displayed in Figure 2, I feel confident that we can tease these apart and glimpse some light between legal idealism and legal realism, with the former receiving greater claimed allegiance from sitting judges.

2. *Judges predicting*

I broached the topic of predicting up in Part I. Moreover, in the survey results discussed above, we saw the extent to which predicting up influences judicial decision making. Now I asked a series of follow up questions: how good are lower court judges at predicting the decision-making of those higher up on the judicial hierarchy? In particular, I asked the participants the extent to which they believe they can accurately predict their judicial peers, their state's highest court, and the Supreme Court of the United States.
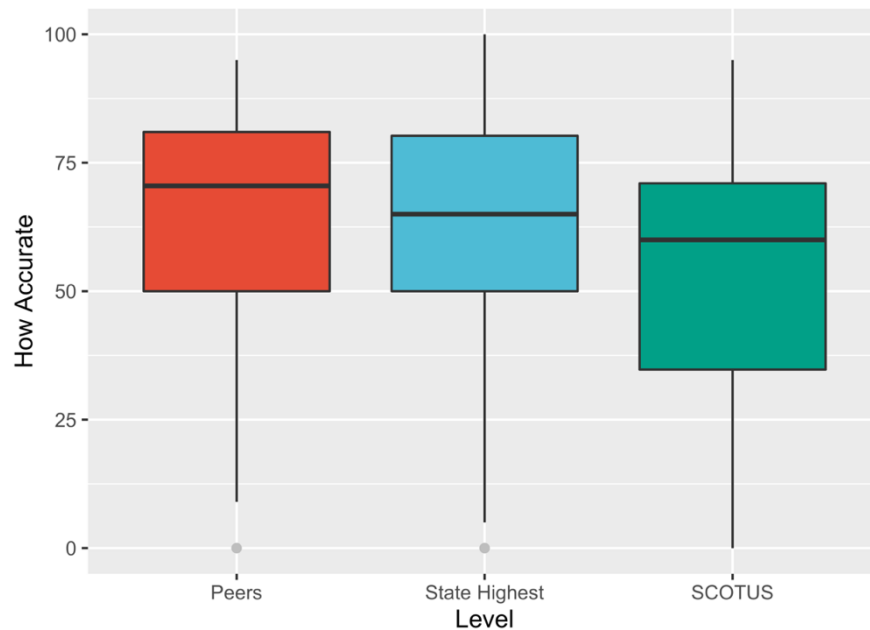
Figure 5. The judicial participants were asked to what extent they agreed with the following statements: "I can accurately predict how my judicial peers (specifically, those judicial peers whom I know well) will rule on specific legal matters." "I can accurately predict how my state's highest court will rule on specific legal matters." "I can accurately predict how the Supreme Court of the United States will rule on specific legal matters." The scale ranged from 0 (not at all) to 100 (definitely).

As Figure 5 shows, the judges were confident in their ability to predict up. This was especially true in their own states, where there was virtually no difference between their confidence in predicting their judicial peers (mean = 63) and predicting their state's highest court (mean = 59); the difference between the two was not significant, neither by a two-tailed t-test ($p = .21$) nor when the t-test was adjusted using Tukey's HSD ($p = .44$).

Embedded within such prediction is the question of what features judges rely upon when predicting up. To shed light on this, I asked the participants, "If you were attempting to predict the result that would be reached if the Supreme Court of the United States were faced with a specific legal matter, as a basis for predicting the ruling, to what extent would you consider the following: published opinions by the justices; the justices' non-judicial writings and speeches; the justices' general ideological commitments; your personal knowledge of and casual conversations with the justices, if any." As before, the scale ranged from 0 (not at all) to 100 (definitely).
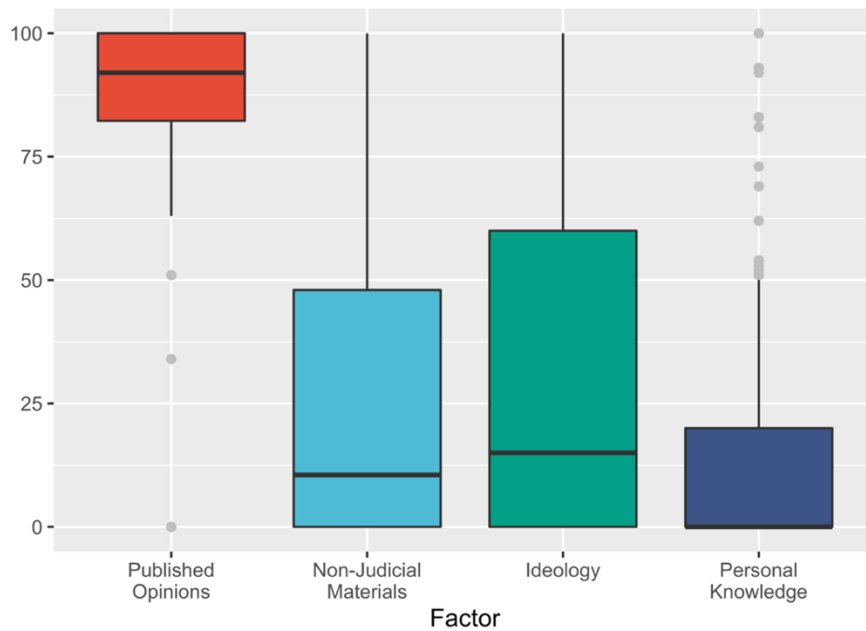
Figure 6. The judicial participants indicated how important these factors are when predicting Supreme Court rulings.

The judges hewed to the official story, with published opinions holding by far the greatest weight. Even ideology, which often is viewed by the public as a significant factor in Supreme Court decision making, came in quite low (mean of 31) compared to published opinions (mean of 89).

In addition to questions regarding predicting up, I also asked the judges about predicting down. Here, the prediction tasks are more particular and discrete, so I asked about two specific ones: future behavior of adult and juvenile defendants/offenders. Note, for these responses, the data analysis is limited to responses by those judges who indicated that they handle at least some criminal cases, a criterium that still included 91% of the participants.

For the first of these, I asked the judges to what extent they agreed with the following statement: "On the whole, I can accurately predict how criminal defendants will behave during the post-arrest and pretrial period (i.e., if released on bail, whether they will commit additional offenses, whether they will show up for subsequent court dates, etc.)." For the second item, I asked about a prediction task that was established by the Supreme Court in *Miller v. Alabama*.[154] As discussed above, in *Miller*, the Court banned mandatory life without parole sentences for juveniles and declared that only the "juvenile offender whose crime reflects irreparable corruption" should be made to spend the rest of his or her life in prison.[155] I asked the judges, "To what extent do you agree with this statement: I can predict whether a juvenile is 'irreparably corrupt'; that is, I can predict whether a juvenile will be a danger decades from now and after a long prison sentence."

---

[154] 567 U.S. 460 (2012).

[155] *Id.* at 479–80 (internal quotation marks omitted) (quoting Roper v. Simmons, 543 U.S. 551, 573 (2005)).
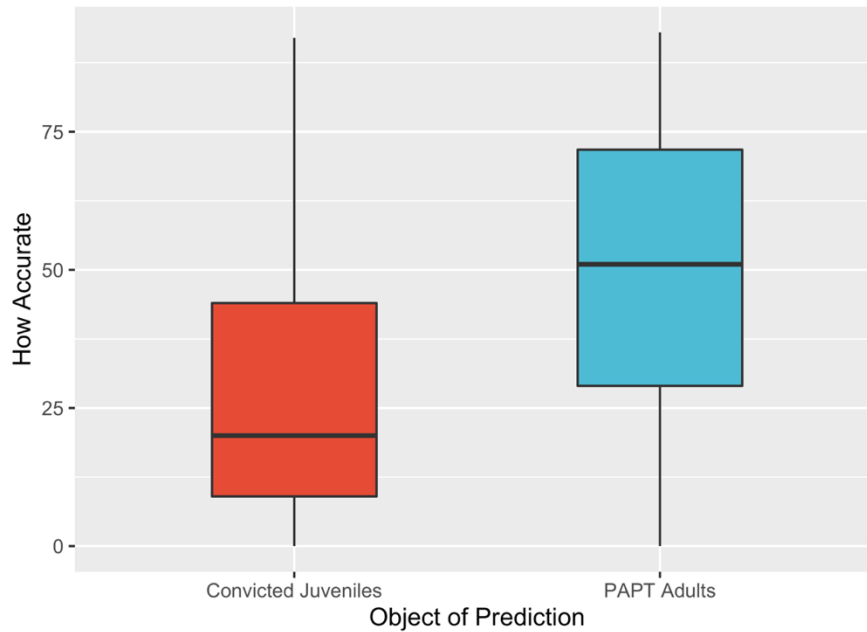
Figure 7. The judicial participants indicated how well they can predict whether a juvenile will be a danger in the distant future and whether a post-arrest and pretrial (PAPT) adult offender will behave appropriately if released on bail.

While the judges professed a moderate level of confidence in their bail decision making (mean of 50), they showed virtually no confidence in their ability to perform the prediction task that the *Miller* Court asked them to perform (mean of 27). A two-tailed t-test showed a significant difference in these confidence ratings, with p-value $< .001$ and Cohen's $d = .9$, which is a large effect size. In other words, while judges might have some confidence in their bail decision making, they have no confidence in their ability to do what the *Miller* Court has asked them to do.

Lastly, the judges were asked about a key issue in predicting down: are risk assessments helpful? Specifically, they were asked to what extent they agreed with the following: "When trying to predict how criminal defendants will behave during the post-arrest and pretrial period (i.e., if released on bail, whether they will commit additional offenses, whether they will show up for subsequent court dates, etc.), it is helpful for me to have access to data-driven predictions, such as risk assessment tools."

The results were overwhelming. Ninety-nine percent of judges indicated that risk assessment tools have at least some value (score of greater than 0). Even more, 83% indicated that risk assessment tools showed significant value (score greater than or equal to 50). The mean was above the midpoint: 65 (standard deviation of 25).
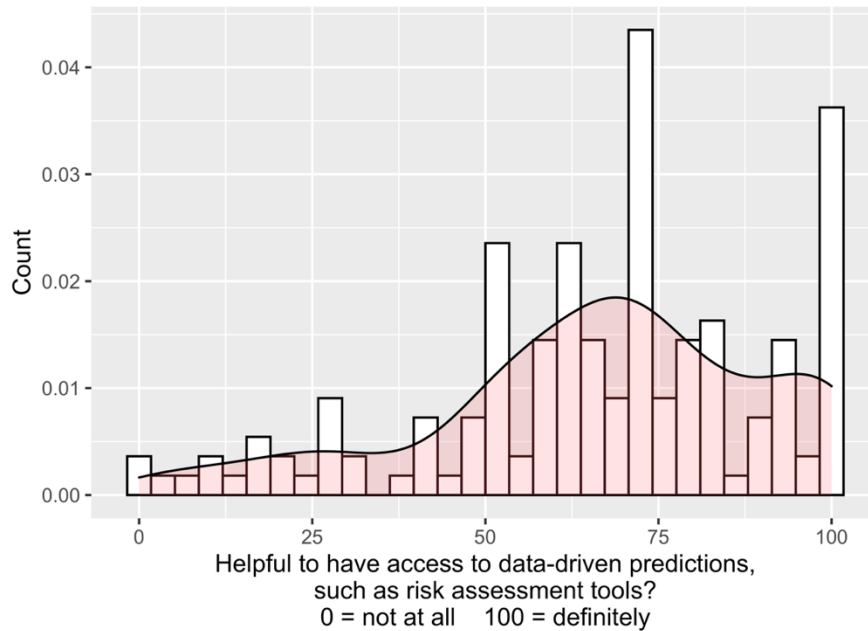
Figure 8. Do judges find risk assessment tools helpful? The answer was overwhelmingly in the affirmative.

3. Others predicting judges

Judges predicting is just one piece of the prediction model. There also are others who might be tasked with predicting judges: lawyers, for instance, as is discussed throughout this article; and machines or computational models, as in the Supreme Court Forecasting Project.[156] But are such outside predictors any good? I asked judges this very question. First, to provide a baseline—and to check it against the result reported in the preceding

---

[156]   Theodore W. Ruger, Pauline T. Kim, Andrew D. Martin & Kevin M. Quinn, *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking*, 104 COLUM. L. REV. 1150 (2004).

section—I asked to what extent the judicial participants agreed with this statement: "My judicial peers who know me well can predict with accuracy how I will rule on specific legal matters." The mean (69) was slightly higher than when the judges were asked how well they predict their judicial peers (63). In other words, the judicial participants had greater confidence that others could accurately predict their decision making than that they could accurately predict others' decision making. They had slightly more confidence that attorneys ("The average attorney who knows me well can predict with accuracy how I will rule on specific legal matters.") could predict their decisions (mean of 71), but this was not significantly greater than their confidence that judicial peers could predict their decisions ($p = .10$). Finally, they were asked to indicate their agreement with two statements regarding data-driven prediction tools. "It is currently possible to build a model (for example, a data-driven artificial intelligence) that can predict with accuracy how I will rule on specific legal matters." "Within the next 10 years, it will be possible to build a model (for example, a data-driven artificial intelligence) that can predict with accuracy how I will rule on specific legal matters." As is evident in Figure 9, they had significantly more confidence in the human predictors than in the computational ones, even when giving the AI ten more years during which to improve.
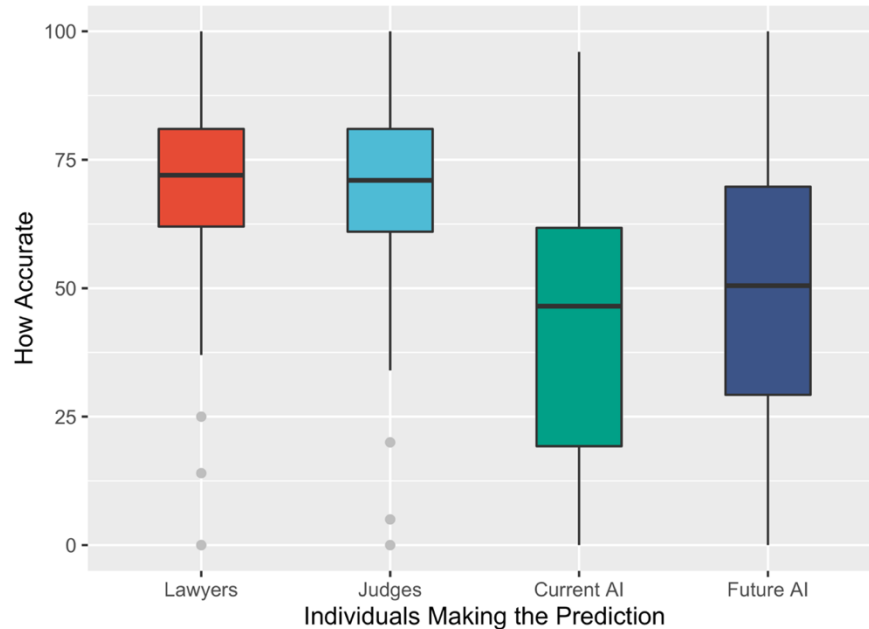
Figure 9. How accurate do judges think lawyers, judicial peers, and current and future (10 years from now) artificially intelligent tools are at predicting judicial decisions? The scale ranged from 0 (not at all accurate) to 100 (definitely accurate).

B. *Summary of Key Empirical Findings*

First, when it comes to a judicial notion of Justice Holmes's prediction model, it appears that prediction factors into judicial behavior but is less a factor than are concerns about what the law, in an absolute sense, demands.[157] In lay terms, judges try to rule in accordance with the law (that is, they try to "get it right," as one participant wrote to me), and that might include predicting how higher courts would rule on the same matter, which can be a question of *stare decisis*, of course. But this does not necessarily have much to do with predicting whether or not they will be subsequently reversed. When it comes to this subset of judicial legal prediction, there is a split in opinion: some judges care deeply about reversal, others do not, and very few are in the middle.

Moreover, on this same issue of predicting up, judges believe they can do it well, especially within their own states.[158] When making such predictions, judges rely most heavily on higher court judges' published opinions. In other words, they follow precedent and the principles of *stare decisis* in trying to ferret out what judges might do in the future. Two questions emerge from this. One, according to other research on judicial decision making, does this factor—published opinions—account for judicial rulings? Two, while lower court judges might profess to rely most heavily on higher court judges' published opinions, is this only true for the average, banal matter? That is, is it possible that published opinions matter much less when the legal issue is novel or the case a so-called "edge" case?

For this first question, consider the Supreme Court Forecasting Project in which researchers predicted the outcomes of cases heard during a single Supreme Court Term.[159] Did published opinions matter? In essence they did, as the model had to be trained on something, and past opinions are the bedrock of the law, the shifting sands which, taken together and taken across

---

157   *See* Figures 2–4.

158   *See* Figure 5.

159   Ruger et al., *supra* note 156.

time, make up a beach that is relatively consistent—or, at least, identifiable. But the modelers pulled out just a few factors from those opinions (issue area of the case, whether there was a constitutional claim), and what was as important—or more important—were factors like ideological direction of the lower court ruling, whether the state was a petitioner, and so on. In essence, we might conclude that the judges are right and are accurate in claiming the importance of published opinions (as we saw, *supra*, in *In re Roche*[160]), but we also can adumbrate additional factors that play a role.

For the second question, we have already hinted at the extent to which ideological direction and idiosyncrasies of individual justices matter for edge cases. Recall the flurry of petitions for rehearing that quickly followed on the heels of Clarence Thomas's 1991 Supreme Court confirmation.[161] Moreover, in Part I, I discussed *Barnette*, wherein a change in the composition of the Court was correctly perceived as *the* factor that would change the Court's decision making.[162] And general and lay opinion focuses on the importance of ideology and other factors for these edge cases, as is evident in the trenchant battles that ensue whenever a judge is up for Supreme Court confirmation.

The empirical work also provides important insights into judges predicting down. First, they are not overly confident in their ability to do so, especially in line with the *Miller* holding,[163] and they want data-driven tools, such as risk assessments, to help guide their decision making.[164] This latter conclusion may come as something of a surprise given the negative reputation that risk assessment tools have,[165] but it is worth noting and worth considering as policy and legal decisions are made concerning their use. This latter point also raises a curious inconsistency: while judges would like more computational guidance as they predict down, they have relatively little confidence in the ability of computational tools to predict up.[166] On the one hand, this inconsistency lends credence to arguments against the use of risk assessment tools—although I will discuss risk assessment tools in more detail

---

[160]   448 U.S. 1312 (1980).

[161]   Dorf, *supra* note 3, at 652.

[162]   47 F. Supp. 251, 252–53 (S.D. W. Va. 1942).

[163]   *See* Figure 7.

[164]   *See* Figure 8.

[165]   Gwen Van Eijk, *Socioeconomic Marginality in Sentencing: The Built-In Bias in Risk Assessment Tools and the Reproduction of Social Inequality*, 19 PUNISHMENT & SOCIETY 463 (2017).

[166]   *See* Figure 9.

in Parts III and IV. More importantly, I suspect it reflects general distaste individuals have for being the object of machine prediction, but this is a matter that must be empirically tested in the future.

On this same point, the results show how little insight judges have into the ability of statistical forecasters to predict their behavior—and the extent to which judges misvalue human (vs. machine) prediction. Almost two decades ago, researchers pitted a statistical model against legal experts in predicting the outcomes of every case argued during the 2002 Supreme Court Term.[167] The experts included luminaries from top law schools and law firms around the country.[168] The model did not even take into account information about the specific law or facts of the cases. Indeed, it only considered six factors: (1) circuit of origin; (2) issue area of the case; (3) type of petitioner (e.g., the United States, an employer, etc.); (4) type of respondent; (5) ideological direction (liberal or conservative) of the lower court ruling; and (6) whether the petitioner argued that a law or practice is unconstitutional.[169] Who performed better? The model, as it correctly predicted 75% of the Court's affirm/reverse decisions, while the human experts predicted only 59% correctly.[170] During the two decades since this study, computational tools have improved dramatically,[171] and it is shocking that judges still prize human forecasting over AI forecasting. In 2011, a team of researchers used the votes of any eight Supreme Court justices from 1953 to 2004 to predict the vote of the ninth justice in those same cases, and their accuracy was 83%.[172] In a 2017 study, a team of researchers used data from 1816 through 2015 to predict 28,000 Supreme Court decisions and 240,000 votes by the justices.[173] The algorithm correctly predicted 70% of the decisions and 72% of the votes.

---

[167]   Ruger et al., *supra* note 156.

[168]   *Id.* at 1206–07.

[169]   *Id.* at 1154 n.19.

[170]   *Id.* at 1171.

[171]   KEVIN D. ASHLEY, ARTIFICIAL INTELLIGENCE AND LEGAL ANALYTICS: NEW TOOLS FOR LAW PRACTICE IN THE DIGITAL AGE 107–125, 351–353 (2017).

[172]   Roger Guimerà & Marta Sales-Pardo, *Justice Blocks and Predictability of U.S. Supreme Court Votes*, 6 PLOS ONE e27188, 4 (2011).

[173]   Daniel Martin Katz, Michael J. Bommarito II & Josh Blackman, *A General Approach for Predicting the Behavior of the Supreme Court of the United States*, 12 PLOS ONE e0174698, 1 (2017).

III. A NEW FRAMEWORK FOR LEGAL PREDICTION

In Parts I and II of this article, we gained a vantage from which to view how legal prediction emerges in cases and controversies, and we were able to gather the results of an empirical study of judicial thinking on legal prediction. In short, we now have the raw material from which a more robust understanding of legal prediction can be formed and, more importantly, from which a new framework of legal prediction can be constructed. While this discussion is important for legal philosophy, that strand of thinking is taken up in Part IV, where we also begin to glimpse how computational advances might alter our philosophical understanding of the feasibility of Justice Holmes's prediction model. However, legal philosophy is not the aim of Part III, nor is it the aim of this new framework. Rather, the framework enables us, whenever a legal prediction problem or question is raised, to classify the prediction problem and immediately identify the issues, benefits, and risks attendant to it. This is important for issues as far ranging as individualization in sentencing and improper delegation of judicial authority, to name but two. Moreover, in the specific area of bias in the law, the failure to properly classify legal prediction problems has led to a lack of nuance in deciding when legal prediction might be harmful and lead to entrenched biases, and when it might result in the opposite effect, that is, decrease bias in legal outcomes. The framework takes steps towards resolving such error.

Let us return to the hypothetical presented in the introduction to this article. A startup founder is faced with a regulatory question. In seeking legal advice, does she want to predict the outcome of subsequent legal review of her behavior, or does she want something less practical: to know whether she has, in absolute or theoretical terms, violated applicable law? Whether there is a difference between these questions was discussed in Parts I and II of this article, and thus I set it aside here. That said, we can conclude, without major qualms, that the startup founder cares only about the former: has she done something that will be deemed illegal or improper by the relevant legal authority? It is a pure prediction problem, this much is clear. We should not stop here, however. It is a specific type of prediction problem, the kind found in only a subset of the canonical cases discussed in Part I. Specifically, it is a matter of predicting up. What does it mean to predict up? This is the more typical chain of prediction, indeed, the class of prediction that Justice Holmes

had in mind when he discussed his prediction model. Lawyers predict lower court judges, and lower court judges predict higher court judges.

The answer to the question of whether it is an up or down prediction problem hinges on the identity of the party whose behavior is being predicted. If that party is a decision maker (i.e., not the object of the decision), then it is a matter of predicting up. For example, lawyers might predict judicial behavior. This is a matter of predicting up. But they also might predict the behavior of jurors, and they might predict the behavior of prosecutors. These also are instances of predicting up.

In contrast, when the party whose behavior is being predicted is not a decision maker but rather is the object of the decision being made, then it is a matter of predicting down. Judges predict what criminal defendants will do when out on bail, and they predict what offenders will do post-sentencing. Supreme Court Justices predict how their legal rulings will be implemented by the impacted parties. These are instances of predicting down.

In recent years, predicting down problems have often been buttressed by computational predictors. For example, *Loomis* concerned judicial use of the COMPAS risk assessment tool.[174]   As another example, the Arnold Foundation created the Public Safety Assessment tool (PSA), which is in use around the country and functions by using nine factors to generate scores that predict three outcomes in criminal cases—failure to appear, new criminal arrest while on release, and new violent criminal arrest while on release.[175]

With one exception, this reliance on computation is less pronounced in prediction up problems. Lower court judges, for instance, do not rely on predictive analytics to gain insight into how the Supreme Court will handle a specific matter or whether they might get reversed. However, lawyers are

---

174   Brief of Defendant-Appellant at 8–10, State v. Loomis, Crim. A. No. 2015AP157, 2015 WL 1724741, at 9–11 (Wis. Ct. App. Sept. 17, 2015); *see also* State v. Loomis, 881 N.W.2d 749, 755–57 (Wis. 2016).

175   *About the Public Safety Assessment*, ADVANCING PRETRIAL POL'Y & RSCH. (2020), https://advancingpretrial.org/psa/factors/ [https://perma.cc/N6MF-KY67] (last visited Jan. 29, 2021).

increasingly using computational tools to predict the outcomes of motions[176] and even cases altogether.[177]

On this point, it is worth discussing that, as we have seen in Parts I and II of this article, legal prediction—both up and down—is taking place both with and without computational aid. Even without the aid of a risk assessment tool, a judge making a decision regarding pretrial release will make predictions regarding the defendant's future conduct. These predictions might be based on "objective" factors, such as past criminal convictions, but they also will be based on hunches, on intuition, or likely on factors not fully conscious to the judge herself. However, were the judge to use a risk assessment tool like the PSA, the factors going into the prediction would be explicit.[178] Moreover, whether such predictions are generated by a human expert (a psychologist or psychiatrist, as in *Lessard*[179]) or an algorithm (COMPAS; the PSA), both are using statistics (what, in the case of the human expert, we would call "experience") and both are subsequently used by judges to inform their own predictions. When an expert, in a case along the lines of *Lessard*,[180] considers the likelihood of a suicide attempt and states, "Harm is likely," is this any more transparent or reliable than a data-driven, machine-generated prediction?

While this classification system—predicting, up and down—is a slight change in our orientation to legal prediction, it is a significant one, as it draws into relief and clarifies issues over which scholars are currently wrestling. To start, a number of scholarly books have been written protesting the use of prediction, and these protestations are largely motivated by claims that such technology will lead to racially biased outcomes.[181] Entire organizations are devoted to this stance. The AI Now Institute at New York University focuses on identifying bias in artificial intelligence (AI) and its impact on human

---

176 *See* BLUE J, https://www.bluejlegal.com/ [https://perma.cc/B8DC-FL2Z] (featuring software that predicts the outcomes of tax motions) (last visited Oct. 27, 2021).

177 *See* CLAUDIUS LEGAL INTELLIGENCE, https://www.claudius.ai/ [https://perma.cc/2X6Q-7MUX] (featuring an artificial legal intelligence that predicts the outcomes of civil legal matters) (last visited Oct. 27, 2021).

178 ADVANCING PRETRIAL POL'Y & RSCH., *supra* note 175.

179 349 F. Supp. 1078, 1093 (E.D. Wis. 1972).

180 *Id.*

181 RUHA BENJAMIN, RACE AFTER TECHNOLOGY (2019); VIRGINIA EUBANKS, AUTOMATING INEQUALITY: HOW HIGH-TECH TOOLS PROFILE, POLICE, AND PUNISH THE POOR (2017); CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY (2016).

rights: the Institute has partnered with the American Civil Liberties Union, publishing papers expressing wariness of the increasing reliance on technology in the United States.[182]

Many of the issues that worry commentators are issues inherent to predicting down. Then-Attorney General Eric Holder famously spoke about the risk of penalizing individuals for immutable characteristics. Recent work,[183] including the aforementioned books by Ruha Benjamin and Virginia Eubanks, further this line of thought.[184] Michael Donahue's 2019 article focuses entirely on problems in what I call predicting down.[185] These scholars seem right to question such uses of prediction, but they must be careful to disambiguate predicting down from predicting up problems.

When it comes to predicting up, there is tremendous potential for correcting for racial bias *through* the use of predictive analytics.[186] Here, the prediction model can function like real-time statistics, flagging potential instances of bias and nudging for preemptive correction of those future decisions.[187] Predicting down, of course, is a different matter, and it should be understood as such when there are discussions about its propriety. But predicting up largely permits the benefits of what is called "algorithmic social

---

182   Meredith   Whittaker   et   al.,   *AI   Now   Report   2018*,   AI   NOW   INST., https://ainowinstitute.org/AI_Now_2018_Report.html   [https://perma.cc/2YS4-54B7]   (last visited Jan. 29, 2021).

183   *See, e.g.*, Gwen van Eijk, *Socioeconomic Marginality in Sentencing: The Built-In Bias in Risk Assessment Tools and the Reproduction of Social Inequality*, 19 PUNISHMENT & SOC'Y 463, 464 (2017) (analyzing "how including socioeconomic marginality as a risk factor in sentencing decision reproduces – and possibl[y] exacerbates – disparities in sentencing as well as social inequality more generally").

184   *See* Benjamin, *supra* note 181 (discussing how technology, even when it appears neutral, can advance racial discrimination); Eubanks, *supra* note 181 (investigating three data algorithms and predictive models that the government uses to provide public services and how those systems may contribute to inequality).

185   *See* Michael E. Donohue, *A Replacement for Justitia's Scales?: Machine Learning's Role in Sentencing*, 32 HARV. J.L. & TECH. 657, 660–66 (2019) (arguing that such tools lead to a lack of individualization, are opaque, and encompass only one philosophical conception of punishment).

186   *See* Joseph J. Avery & Joel Cooper, *Racial Bias in Post-Arrest and Pretrial Decision Making: The Problem and a Solution*, 29 CORNELL J.L. & PUB. POL'Y 257 (2020) (arguing that prosecutors should be given race-neutral baselines as a case progresses through the post-arrest and pretrial ("PAPT") period to increase fairness in decision making during this period).

187   *See id*. at 280–93 (proposing that prosecutors should be given race-neutral outcomes to use in their decision making and that those who deviate from the model's race-neutral recommendation have to document the reasons why they have deviated in order to monitor compliance with race-neutral mandates).

engineering," that is, using algorithmic modifications to nudge decision-makers toward preferred actions.[188]

Professors Cowgill and Stevenson outline a decisional equation that is useful for illustrating the difference between predicting up and predicting down. They posit that a decision maker's action selection is a maximizing function of both utility ("the highly subjective preferences about which there are no correct answers") and beliefs (p: "beliefs about the objective state of the world that could in principle be verified").[189] An example of p would be a judge's belief that a particular defendant will or will not recidivate if released on bail. This is what I call predicting down.

However, p does not exist in many decisional contexts, especially those within the criminal justice system. What sentence should a defendant receive? The answer will hinge on how one conceives of the purpose of criminal punishment. Is it primarily about rehabilitation? Or is it about incapacitation? Or perhaps specific deterrence, general deterrence, retribution, restitution, normative validation—or even something else? Various of these purposes have been embraced, oftentimes simultaneously, both historically[190] and at present.[191] The point is, for decisions such as this, there is no state of the world that can in principle be verified. There is not necessarily a right answer.

If a court decides to correct for racial bias in its decisions, it may design and implement a suggestive model. Let us assume this model is like the one discussed elsewhere: Black defendants are treated as if they are White defendants.[192] Now p, if we still choose to call it that, represents the extent to

---

[188] *See* Bo Cowgill & Megan T. Stevenson, *Algorithmic Social Engineering*, 110 AEA PAPERS & PROC. 96, 96 (2020) (discussing how the output of algorithms can be manipulated to embed a policymakers' preferences and thus push decision makers toward a preferred action).

[189] *Id*. at 97.

[190] *See* Henry Weinhofen, *The Purpose of Punishment*, 7 TENN. L. REV. 145 (1929) (discussing the origin of punishment and various theories of punishment).

[191] *See, e.g.*, Albert W. Alschuler, *The Changing Purposes of Criminal Punishment: A Retrospective on the Past Century and Some Thoughts about the Next*, 70 U. CHI. L. REV. 1 (2003) (discussing the emergence and later demise of rehabilitative goals of criminal punishment in the twentieth century and the recognition of retributive objectives); Marissa A. Booth, *The Road to Recovery: The Third Circuit Recognizes the Importance of Rehabilitative Needs during Sentencing in* United States v. Schonewolf, 64 VILL. L. REV. 569, 574–75 (2019) (analyzing the Third Circuit's holding in *United States v. Schonewolf* that courts may consider a defendant's rehabilitative needs at sentencing).

[192] *See* Crystal S. Yang & Will Dobbie, *Equal Protection Under Algorithms: A New Statistical and Legal Framework*, 119 MICH. L. REV. 291, 348–50 (2020) (describing a "minorities-as-whites" algorithm that would eliminate the effects of race when forming predictions).

which the decision-maker's action accords with the suggested model's output. Are White defendants treated how the machine predicts they should be treated? Are Black defendants treated how the machine predicts White defendants so situated should be treated? Importantly, the utility function would *not* be independent of this modified p; after all, the decision maker is audited and rated in terms of his or her performance in relation to the suggestive model's output. That is, doing what the machine suggests will result in greater utility for the decision-maker. Considered in this light, we see that, when predicting up, if a suggestive model is embedded within an appropriate framework, it will be a powerful force for constraining and guiding decision making. When the objective of the suggestive model and its attendant framework is racial equality, greater racial equality should be observed in actors' decisions.

With the classification system now clear, we can discuss benefits and risks, as well as legal issues, that attach to our two classes of prediction problems.[193] My aim here is not to exhaustively explore these. Rather, it is to adumbrate them while, more importantly, showing how different ones attach to the distinct classes of prediction problems: predicting up is importantly different from predicting down.

Whose behavior is being predicted?	**Classification**	**Benefits**	**Risks**	**Legal Issues**
Decision Makers (those not impacted by the decision): prosecutors, jurors, judges	-Predicting Up	-Bias reduction -Efficiency -Fairness	-Legal lock-in -Undermines the rule of law -Undermines lay connection to legal processes	-Improper delegation -Failure to adhere to *stare decisis*

---

193   *See* Table 1 for an overview.

Those impacted by the decision: defendants, plaintiffs, respondents	-Predicting Down	-Accuracy -Transpa-rency	-Enshrines historical bias -Error: judges predicting down, machines predicting down, experts predicting down	-Lack of individual -ization

Table 1. Cursory overview of the two-classification framework for legal prediction problems.

First, to continue the previous line of thought, when predicting up, there are a number of benefits that can emerge. One is bias reduction, which has been discussed in detail elsewhere,[194] as well as *supra*, so I will not belabor the point here. A second benefit is fairness, which dovetails with bias reduction but is distinct. Judges are human, and as discussed throughout this article, their individual personalities transcend the courts on which they sit. Indeed, the cumulative result is that courts are distinct from each other and circuits are distinct from each other, and for this reason, forum shopping is a robust practice.[195] Predicting up leads to convergence, as the predictions are focused on the highest court that is common to the lower courts. Thus, within the Circuit of Appeals system, so long as the same predictive tool is used, a plaintiff in the Fifth Circuit could expect the same verdict as a plaintiff in the Ninth Circuit. While this certainly is a form of fairness, we also can think of it as a form of efficiency. Dodging certain judges, forum shopping, crafting arguments specific to the judges one faces—these are lessened in the face of consistently applied prediction up.

---

[194] *See, e.g.*, Joseph Avery & Joel Cooper, *Technology in the Legal System: Uses and Abuses*, *in* BIAS IN THE LAW: A DEFINITIVE LOOK AT RACIAL PREJUDICE IN THE U.S. CRIMINAL JUSTICE SYSTEM 161, 161–81 (Joseph Avery & Joel Cooper eds., 2020) (discussing how to reduce the impact of biases by using data-driven algorithms and legal technology).

[195] *See* Scott E. Atkinson, Alan C. Marco, & John L. Turner, *The Economics of a Centralized Judiciary: Uniformity, Forum Shopping, and the Federal Circuit*, 52 J.L. & ECON. 411, 413 (2009) (finding evidence of forum shopping before the Court of Appeals for the Federal Circuit was created and given exclusive jurisdiction over appeals in patent cases initiated in federal district courts).

These benefits—bias reduction, fairness, efficiency—can be found at the different levels of predicting up. For example, an attorney in the Midwest indicated that he was interested in a specific prediction tool[196] because he knew that it could be used to show the insurance companies whether and when they were treating minority race plaintiffs differently than majority race plaintiffs. As for efficiency, if prediction is heavily relied upon, then prosecutors and defense attorneys might agree on what the likely outcome of a trial will be, thus saving time in case prosecution. Similarly, plaintiffs' and defense attorneys might agree on the likely settlement for a personal injury suit, thus avoiding costly and protracted litigation. But these are best left for discussion in another article, one focused on the feasibility and result of instantiating AI and prediction systems.

For the purposes of this article, I will, instead, highlight how the benefits of predicting down are importantly distinct from the benefits of predicting up. For one, predicting down is often thrust upon the decision maker, as in a *Miller v. Alabama* application,[197] so it might not make sense to talk about benefits outside of a few niche scenarios, such as a lawyer predicting how her client will perform in a deposition. Or, perhaps, a judge predicting dangerousness in a bail hearing when such a prediction is not truly warranted. That said, it does make sense to compare the benefits of machine vs. human predictions down. To wit, does machine prediction down lead to a decrease in bias? Most likely not, as a number of scholars have shown that predicting down suffers from problems stemming from biased datasets, suspect features (such as criminal arrests, which may themselves be reflective of bias), and so on.[198] Will it lead to fairness? Again, for similar reasons, probably not, although it could lead to greater consistency (i.e., the results will be consistently unfair). It may, however, result in greater transparency. For example, the PSA risk assessment tool uses nine factors, all of which are clearly explained and outlined.[199] When a judge operates in the absence of

---

196   The tool is Claudius Legal Intelligence, Inc. *See supra* note 177 (featuring an artificial legal intelligence that predicts the outcomes of civil legal matters).

197   *See* Miller v. Alabama, 567 U.S. 460, 465 (2012) (holding that "mandatory life without parole for those under the age of 18 at the time of their crimes violates the Eighth Amendment's prohibition on 'cruel and unusual punishments.'").

198   *See* Avery & Cooper, *supra* note 186, at 283 (describing an approach that would remove "suspect features" from statistical algorithms to facilitate in moving towards a system of racially equal treatment for criminal defendants).

199   *See* ADVANCING PRETRIAL POL'Y & RSCH., *supra* note 175 (showing the Public Safety Assessment Tool and the nine factors that it uses).

such a tool, does she use nine clearly explained and outlined factors? Of course not. She relies instead upon various factors that she may mention and a host of others that operate semi- or subconsciously. We also have seen, in Part II, that judges are wary of their ability to predict down, and rightly so. It is difficult to predict human behavior, and such predictions are always beset by biases that attach to the predictor.[200] Can machines do better? Perhaps, as reviews are starting to reveal that machines produce less biased and more accurate predictions than humans.[201]

Consideration of the risks of predicting up and down reveals the most evident differences. When predicting up, there is a serious possibility of technological-legal lock-in.[202] As Roscoe Pound wrote in 1907,

> The effect of all system is apt to be petrifaction of the subject systematized. Perfection of scientific system and exposition tends to cut off individual initiative in the future, to stifle independent consideration of new problems and of new phases of old problems, and to impose the ideas of one generation upon another. This is so in all departments of learning.[203]

Applying this insight to the case of widespread prediction up, the fear is clear: the system will become fixed, and so fixed will cease to have the malleable and changeable nature that we expect from our "living" system of law. Does this risk also attach to machine prediction down? Not really, and if it does, it is a slightly different risk. There, the risk is that the historical record, with its biases, will become enshrined through model training, not that the system itself will become fixed, since there will be new outcomes (did the defendant actually recidivate on release?) that will update the model moving forward.

What is a risk of predicting down? One risk is error. Whether the predictor is legal (a judge), expert (expert witness), or machine (risk assessment tool), one risk is always error. There is the risk of error in an individual case that might be an outlier, and there is the risk of consistent error due to mistreatment of certain demographic groups. Error is not such

---

200   *See* DANIEL KAHNEMAN, THINKING, FAST AND SLOW 185–255 (2011) (discussing how predictions contain various biases, such as hindsight and outcome bias).

201   *See, e.g.*, Alex P. Miller, *Want Less-Biased Decisions? Use Algorithms.*, HARV. BUS. REV. (July 26, 2018), https://hbr.org/2018/07/want-less-biased-decisions-use-algorithms [https://perma.cc/RG33-AYMK] (providing examples of research which shows that algorithms are less biased and more accurate than humans).

202   *See* Crootof, *supra* note 22, at 233, 246–50 (describing how hybrid human-AI judicial systems, once they are cemented in, may be difficult to change and not easily adaptable to new situations).

203   Pound, *supra* note 23, at 606.

a risk of predicting up. After all, the hierarchical nature of judicial review is designed to correct for error, including errors in predicting up, as we saw in the cases presented in Part I.

Even so, we know that predicting up would undermine the lay connection to legal processes. If such predictions were perfect, what need would there be for jurors? After all, we would be able to predict outcomes without having to provide space for lay deliberations. As Lord Herschell said, "important as it was that people should get justice, it was even more important that they should be made to feel and see that they were getting it."[204] Or, turning back to Roscoe Pound, we must acknowledge that the law "must not become so completely artificial that the public is led to regard it as wholly arbitrary."[205] This risk does not attach to predicting down, nor does the risk of undermining the rule of law. As Professor Dorf discussed, the prediction model, to the extent that it emphasizes the importance of the individual judges who comprise the highest court, undermines the ideal of the impartial judge, an ideal that is essential to individuals' faith in the rule of law.[206]

While there are a host of legal issues that attach to predicting up and down, I will mention just one for each, as these are the two most prominent—and they are quite different. For predicting up, the most obvious legal issue is what some commentators[207] call lower court defiance: accurately predicting up may lead a court, as we saw in *Rodriguez*, to defy Supreme Court precedence.[208] The Supreme Court has roundly rejected such behavior, and it typically will deal with it by issuing a summary reversal, as in *Shaun Michael*

---

[204]   Lord Herschell, *quoted in* 2 JAMES BERESFORD ATLAY, THE VICTORIAN CHANCELLORS 460 (1908).

[205]   Pound, *supra* note 23, at 606.

[206]   *See* Dorf, *supra* note 3, at 684–85 (discussing how the prediction model is inconsistent with central themes and practices of American law, such as the idea of the impartial adjudicator).

[207]   *See* Malia Reddick and Sara C. Benesh, *Norm Violation by the Lower Courts in the Treatment of Supreme Court Precedent: A Research Framework*, 21 JUST. SYS. J. 117, 119 (2000) (explaining that lower court defiance "refers to the rare event in which the eventually overruled precedent is not cited at all by the lower court, yet that precedent is overturned as the Supreme Court affirms the lower court").

[208]   *See* Rodriguez de Quijas v. Shearson/Am. Express, Inc., 490 U.S. 477, 484 (1989) (affirming a Fifth Circuit decision that renounced Supreme Court precedent, but chastising the Fifth Circuit for taking this action on its own authority).

Bosse v. Oklahoma.[209] Let us recall Justice Stevens, who decried such an "indefensible brand of judicial activism."[210]

For predicting down, the primary legal issue is quite different: lack of individualization is the concern, and it is potentially actionable. This was one of the issues in *Loomis*, where the claimant argued that the court's use of the COMPAS tool violated his right to individualization to the extent that it relied on group-level statistics.[211] Risk assessment always relies on group-level statistics, a fact that does threaten individualized treatment, something that has been guaranteed at least since *Furman v. Georgia*,[212] although the specifics of the guarantee are nuanced.[213] The *Loomis* court sidestepped the issue by concluding that, while the tool did not individualize, because it was not the sole basis for the decision, the sentencing procedure was valid.[214] This is not the final word on this issue, as it will undoubtedly come before courts—and higher courts—in the future.

IV. HOW THE MACHINE AGE IMPINGES THE LEGAL PHILOSOPHY UNDERGIRDING THE PREDICTION MODEL

In Parts I-III of this article, I garnered the raw materials (cases and controversies; empirical data from sitting judges) needed to understand legal prediction and generate a new framework for meaningfully engaging with legal prediction problems. In this final part of the article, I turn to an issue that is present in all three of the earlier parts and which finds its apotheosis in a discussion of computational law techniques. The debate between law as found and law as made mirrors the two dominant techniques for building

---

209   137 S. Ct. 1, 2 (2016) ("The Oklahoma Court of Criminal Appeals remains bound by *Booth*'s prohibition on characterizations and opinions from a victim's family members about the crime, the defendant, and the appropriate sentence unless this Court reconsiders that ban.").

210   Rodriguez de Quijas, 490 U.S. at 486 (Stevens, J., dissenting).

211   *See* State v. Loomis, 881 N.W.2d 749, 757 (Wis. 2016) (noting that the claimant was challenging the use of a COMPAS risk assessment at sentencing in part because he argued that "it violates a defendant's right to an individualized sentence.").

212   408 U.S. 238, 239–40 (1972) (per curiam) (holding that "the imposition and carrying out of the death penalty in these cases constitute cruel and unusual punishment in violation of the Eight and Fourteenth Amendments.").

213   *See* William W. Berry III, *Individualized Sentencing*, 76 WASH. & LEE L. REV. 13, 24–49 (2019) (discussing the development of the guarantee in criminal cases).

214   *See* Loomis, 881 N.W.2d at 773–74 (Roggensack, C.J., concurring) (noting that it was permissible to consider the COMPAS tool in sentencing, but a court cannot solely rely on COMPAS to impose a sentence).

computational legal prediction systems, and understanding of this divide is helpful for forming a full understanding of the prediction model.

In the nineteenth century, Rudolph Ritter von Jhering, a German jurist and legal scholar, argued for a shift from *Begriffsjurisprudenz* (a law of concepts) to *Wirklichkeitsjurisprudenz* (a law of outcomes/results).[215] In essence, Jhering was arguing that, instead of philosophically analyzing the field, or human nature, and forming rules, one should start with what one wishes to see in practice and build backwards to the rule. Thus, a law of outcomes, a law of results. If one wants a judge to reach a certain outcome when faced with a specific type of contention in a divorce case, then one should build a rule based on that outcome.

To drill deeper into the discussion, when thinking of the concept of *Begriffsjurisprudenz*, we might ask, what are the abstract legal problems, and what are the solutions to those problems? For Roscoe Pound, this would be the law of rules, the law of concepts, an attempt at building a body of law by rational inference from first principles.[216]   Most legal scholars, when confronted with this idea, immediately think of Langdell: "Law, considered as a science, consists of certain principles or doctrines. . . . If these doctrines could be so classified and arranged that each should be found in its proper place, and nowhere else, they would cease to be formidable from their number."[217] Langdell, of course, represents but one form of this: we also can think of the German Pandectists who worked on legal science derived from neo-Kantian first principles of justice.[218]   The problem with any such conception of the law is that, when courts deduce from concepts, they do not always inquire as to what the effect of such deductions will be, nor do they observe that the outcomes might be the reverse of even the spirit of the concepts.[219] As Pound put it, "[t]he nadir of mechanical jurisprudence is reached when conceptions are used, not as premises from which to reason, but as ultimate solutions."[220]

---

215   *See* Jhering, *supra* note 24, at 1–18 (discussing the Law of Purpose as opposed to the Law of Causality).

216   Thomas C. Grey, *Holmes and Legal Pragmatism*, 41 STAN. L. REV. 787, 809 (1989).

217   CHRISTOPHER COLUMBUS LANGDELL, A SELECTION OF CASES ON THE LAW OF CONTRACTS viii–ix (2d ed. 1879).

218   *See* OLIVER WENDELL HOLMES JR., THE COMMON LAW 163, 16667 (M. Howe ed. 1963) (originally published 1881) (critiquing Kantian principles of justice).

219   Pound, *supra* note 23, at 616.

220   *Id.* at 620–21.

Now, considering *Wirklichkeitsjurisprudenz*, we might ask a very different question: how can we characterize the efficient use made with the limited resources that we have? For Pound, this means that, instead of settling on rules and concepts, we instead ought to deduce decisions.[221] As he put it, instead of seeking rules that will last for years, we should seek the "intelligent application of the principle to a concrete cause, producing a workable and a just result."[222] Note that this is not in direct opposition to *Begriffsjurisprudenz* (direct opposition would be something more like the original strands of legal realism, where empiricism was set up to contrast with idealism[223]), but it is meaningfully the obverse of *Begriffsjurisprudenz* in that it is a pragmatic approach.[224] As an example of this approach, Pound cites Justice Marshall's work in yielding a living Constitution by judicial interpretation.[225] Broad as this example might be, we have enough of a conception to contour the debate, and we should see that it mirrors a specific divide in computational law.

Perhaps the oldest example of an algorithmic legal system (though the algorithm was "on paper" rather than instantiated in a machine) was the Valentinian "law of citations" from the fifth century that established a canon of materials that could be cited—and only these materials were allowed to be cited.[226] When faced with questions of law, judges were tasked with merely tallying which position had the most support from the precedents.[227] Let us also recall the flourishing of like-spirited movements towards widespread legal codification during the eighteenth century, as in the Code of Frederick the Great.[228]

---

221 *Id.* at 613.

222 *Id.* at 622.

223 *See* Morton J. Horwitz, *History of the Public/Private Distinction*, 130 U. PA. L. REV. 1423, 1426 (1981–1982) (discussing the Legal Realism Movement).

224 Grey, *supra* note 216, at 819.

225 Pound, *supra* note 23, at 615.

226 Alan Watson, *The Law of Citations and Classical Texts in the Post-Classical Period*, 34 TIJDSCHRIFT VOOR RECHTSGESCHIEDENIS 402, 405 (1966).

227 Pound, *supra* note 23, at 607.

228 *See, e.g.*, TIM BLANNING, FREDERICK THE GREAT: KING OF PRUSSIA 465–494 (2016) (discussing the development of the Code of Frederick the Great, in which at attempt was made to fulfill the Enlightenment ideal of systematization and codification of both public and private law).

If we fast forward to recent attempts to automate legal decision making, we see that the rule-based (or what some call "logic-based"[229]) approaches absorb the assumptions of *Begriffsjurisprudenz*. In the 1980s, a number of researchers implemented systems based on manually created logical representations of rules, including Sergot and colleagues' rule-based instantiation of the British Nationality Act[230] and Peterson and Waterman's attempt at rule-based product liability.[231] These attempts largely stalled, however, and serious progress at legal automation and, in turn, legal prediction did not flourish until rule-based approaches were discarded in favor of data-driven approaches.[232] These data-driven approaches have been facilitated by advances in computing power and developments of new computational techniques, such as deep learning.[233]

In his writing on legal prediction, Justice Holmes distinguished "general" factors from more specific ones that he deemed too singular or specific to be informative for prediction.[234] He gave the example of "the blandishments of the emperor's wife."[235] But Holmes was writing before social science and computer science had begun to flourish; today, a factor as singular as how recently a judge has eaten can be used to as a meaningful feature in a predictive function.[236] The result is that computational systems now can handle manifold information types and predict very well, better than humans in many instances.[237] But should we be alarmed? After all, such prediction is more closely rooted in *Wirklichkeitsjurisprudenz* than in the starry ideals of

---

[229] *See* Karl Branting, *Data-Centric and Logic-Based Models for Automated Legal Problem Solving*, 25 ARTIFICIAL INTELLIGENCE AND LAW 5 (2017) (discussing logic-based approaches of legal problem solving).

[230] *See* Marek Sergot, Fariba Sadri, Robert A. Kowalski, Frank Kriwaczek, Peter Hammond, & H. Terese Cory, *The British Nationality Act as a Logic Program*, 29 COMMUNICATIONS OF THE ACM 370, 370–72 (1986) (explaining that much of the British Nationality Act of 1981 was translated into simple logic).

[231] Donald A. Waterman & Mark Peterson, *Rule-Based Models of Legal Expertise*, ASSOCIATION FOR THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE 272 (1980).

[232] *See* Branting, *supra* note 229, at 6 (discussing the recent rise of data-centric approaches).

[233] Ashley, *supra* note 171, at 107–125 (explaining and illustrating machine learning techniques).

[234] Holmes, *supra* note 7, at 168–170.

[235] *Id.*

[236] *See* Shai Danziger, Jonathan Levav, & Liora Avnaim-Pesso, *Extraneous Factors in Judicial Decisions*, 108 PROC. OF THE NAT'L ACAD. OF SCIS. 6889, 6892 (2011) (finding that "when judges tend to make repeated rulings, they show an increased tendency to rule in favor of the status quo," which can all be "overcome by taking a break to eat a meal").

[237] *See* Ashley, *supra* note 171, at 107–125, 351–53 (explaining how complex computer case-based reasoning can help predict new problems).

Begriffsjurisprudenz. Whether or not we are alarmed, we ought not be surprised. In 1989, in an article on Justice Holmes's thought, Professor Grey wrote, "the primacy of experience over logic, still seems to me the central, if obscure, truth of American legal thought."[238]

In essence, the dilemma is: do we predict as if we are drawing from "a rigid scheme of deductions from *a priori* conceptions,"[239] or do we rely instead on a Poundian notion of a law of principles that are closely tied to outcomes?[240] We can come to this question more fully by traveling down a rather steep yet well-maintained path. David Marr, the British neuroscientist and physiologist, developed a framework for analyzing complex information processing systems. In particular, in considering the brain, Marr can be understood as proposing three levels of analysis.[241] These levels are the computational, algorithmic, and implementational ones.[242] Before explaining the levels, I must note that the "computational" level does *not* necessarily refer to computers. Readers should be careful to avoid this confusion. Computation may refer more broadly to complex mathematical principles, such as Fluid Dynamics, which I show below.

Let us assume there is a system in place, a system that is engaged in information processing. At the computational level, we must ask: what task is the system carrying out? Said slightly differently, when it comes to computation, we are interested in the goal and its attendant strategies. At the algorithmic level, we must ask, what method does the system use? In other words, how should we represent the input and output. Even more, what is the algorithm that can transform the input into the output? Lastly, there is the implementational level. Given the hardware at the system's disposal, how is the algorithm carried out? Said again, how can we physically realize the algorithmic process and representation?

From a broader perspective, we can reconceive of these levels as negotiating between establishing the problem (computational level) to determining how to solve the problem (algorithmic level) to actualizing the

---

238 Grey, *supra* note 216, at 792.

239 Pound, *supra* note 23, at 608.

240 *See Id.* at 613 (arguing that the American courts should adopt a German code which lays down the idea of using principles over rules in order "to make rules fit cases instead of making cases fit rules").

241 DAVID MARR, VISION (1982) (giving an in-depth investigation into how humans represent and process information).

242 John P. Frisby & James V. Stone, *Marr: An Appreciation*, 41 PERCEPTION 1040 (2012).

solution (implementational level). These levels can be illustrated by recourse to an example that Marr provided:

> [T]rying to understand perception by studying only neurons is like trying to understand bird flight by studying only feathers: It just cannot be done. In order to understand bird flight, we have to understand aerodynamics; only then do the structure of feathers and the different shapes of birds' wings make sense.[243]

The example Marr is using is that of flight. Feathers, the flapping of wings: these might seem like computational issues, but they are not. The computational-level problem is how to fly; the computational problem then is one of Fluid Dynamics.[244] Some birds, some insects, and all working airplanes are able to solve this computational problem. Their unique solutions, however, represent different algorithms. Most birds and all planes take a gliding approach, generating lift through downward movement. Insects, such as bees, and some birds, such as hummingbirds, generate lift through both upward and downward wing movement—and lack the ability to glide.[245] These different methods all are representational of the algorithmic level. They are solutions to the computational problem of how to fly.

Then what happens at the implementational level? Very simple. The wings might be made of cartilage and feathers, or wood, or aerospace grade aluminum.[246] These are all different implementations of the algorithmic solution, each of which works to solve the problem rooted in Fluid Dynamics.

Applying these levels to law helps us to better understand the stakes of the current push towards computational law and legal prediction. At the computational level, we obviously must think of the initial computational law forays into rule-based systems. There, the essence of the values and tensions are distilled, and the primary task is asking the fundamental questions about which the law is concerned (indeed, many rule-based efforts stalled because

---

243   Marr, *supra* note 241, at 27.

244   *See* NIGEL DUFFIELD, REFLECTIONS ON PSYCHOLINGUISTIC THEORIES 38 (2018) (explaining that Marr believes that what distinguishes planes, their pilots, and birds from rocks, platypuses and other projectiles is that the former are able to satisfy the computational principles, "which include 'Level 1' principles of fluid dynamics").

245   *Id.* at 39.

246   *Id.* at 40.

the fundamental questions were underspecified[247]). Granted, this is a centuries long process, one at which legal scholars have been and are laboring, but if it can be rendered into a computational level, then what is the value of the algorithmic level? Most importantly, it would help us to understand systematic deviations from the optimal models. For example, when asked the question, how long is Mars's orbit around the sun, people always give distances that are shorter than the correct answer because they anchor on earth's orbit.[248] But this error is easy to identify: there is anchoring (on earth's orbit) and insufficient adjustment away from the anchor, and this is the bias. When we are operating at the algorithmic level, we can simply apply the correct algorithm for correcting for such an error; in this case, it would be something like Markov Chain Monte Carlo via the Metropolis-Hastings Algorithm.[249] And thus we can quantify the cost of implementing the algorithm: the longer the adjustment period, the more resources one has spent and the less time one has to pursue other things in life. So, bias in this example case is resource-rational.

Knowing that bias might be resource-rational, understanding law and human decision-making at the algorithmic level helps to inform predictions. To name just a few ways: anchoring adjustment will be reduced when there is cognitive load, time pressure;[250] it also will be reduced when there is uncertainty.[251] Both of these apply to legal decision-making and especially to judicial decision-making, and we can see how such an algorithmic approach could then be used both to predict and to proactively inform interventions in the direction of desired legal goals (e.g., that similarly-situated defendants be treated similarly). By knowing which algorithms are

---

247 *See* Ashley, *supra* note 171, at 48–52, 73–106, 171–208 (discussing the process and challenges of translating statutes into programs, modeling legal reasoning, and representing legal concepts with different analytical tools).

248 *See* Nicholas Epley and Thomas Gilovich, *The Anchoring-and-Adjustment Heuristic: Why the Adjustments Are Insufficient*, 17 PSYCHOLOGICAL SCIENCE 311 (2006) (giving examples of what 'anchoring' is).

249 *See* Falk Lieder, Thomas L. Griffiths, and Noah D. Goodman, *Burn-in, Bias, and the Rationality of Anchoring*, In NIPS 2699–2707 (2012) (explaining that the process to generate a single perfect sample from a posterior distribution involves the Markov chain Monte Carlo algorithm).

250 *See* Epley and Gilovich at 312, *supra* note 246 ("[P]eople under cognitive load are less able than those who are not to modify their initial dispositional inferences in light of subsequently considered situational constraints.").

251 *See* Karen E. Jacowitz and Daniel Kahneman, *Measures of Anchoring in Estimation Tasks*, 21 PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN 1161, 1164–65 (1995) (showing that sometimes responses that are most strongly affected by an anchor are made with relatively low confidence).

being used, we can reconstruct the environment so that inputs produce the outputs that are wanted. As discussed in Part III, predicting up becomes something quite powerful and desirable.

Thus, we have come full circle. The future of prediction might lie in thoroughly inhabiting the algorithmic level, but the algorithmic level might very well include neuro-symbolic AI, where both rule-based and data-driven processes are deftly combined and permitted to function in tandem. The result will be both better prediction and better insight into when and why predictions—and behavior—deviate from what is foreseen, wanted, desired. Such developments render self-prediction eminently feasible, which should come as some surprise to those legal scholars who believed that self-prediction was infeasible if not impossible.[252] A court of last resort may sensibly predict how it will rule.

What is truly remarkable is that this computational development, this manifesting of prediction up and down systems, this machine-future, is in no way divorced from the fundamental debate with which this article began. Transcendentalism and functionalism,[253] positivism and realism,[254] varying forms of legal realism—all are encapsulated by the ideas that occupy researchers at the nexus of law, computer science, and prediction.

CONCLUSION

This article made a number of advances in understanding legal prediction, advances that impinge upon criminal law (risk assessment tools), civil law and especially torts (judge and jury prediction systems), administrative law (the role of automated decision tools), and technology law. There are three advances, in particular, that are worthy of recapitulation in this conclusion. First, the article provided an intuitive framework for classifying and handling legal prediction problems. For here on out, a threshold question is to ask whether one is predicting up or predicting down. I showed that this classification is useful for thinking through legal prediction problems. For instance, in the specific area of bias in the law, the failure to

---

252	*See* Dorf, *supra* note 3, at 659 (explaining how Hart thought that the prediction model is incoherent when applied to courts).

253	*See* Cohen, *supra* note 1 (discussing the differences between transcendentalism and functionalism).

254	*See* Leiter, *supra* note 2 (challenging two widespread views about the relationship between Legal Realism and Legal Positivism).

properly classify legal prediction problems has led to a lack of nuance in deciding when legal prediction might be harmful and lead to entrenched biases, and when it might result in the opposite effect, that is, the lessening of bias in legal outcomes.

In Part III, I put flesh upon the bones of this new prediction framework, and I took steps towards clarifying the distinction and why it matters in practice. This discussion was, of course, a starting point, and I propose that it prompt research into the contours and implications of pulling apart legal prediction in the way the framework does. For instance, considering only predicting up, if it were mechanized and ubiquitous, what legal issues would be raised? What recourse would claimants and respondents have? In addition, there should be research into the extent to which specific predicting up problems require input from predicting down problems. That is, when modeling how a judge will rule on a particular defendant, that prediction (up) likely includes predictions about what the judge predicts the defendant will do (down). Disentangling these is key to unpacking the nuances of the matter.

Second, the article's empirical findings on judicial thinking regarding legal prediction are important for academic, theory-driven reasons (understanding what the law is and what it is that judges believe they are doing when they are applying the law; unpacking the tension between law as found and law as made) and also for applied reasons: the research provides insights into risk assessment tools and the workability of extant legal rules, such as that promulgated by the *Miller v. Alabama* Court.[255] In particular, I found that while judges overwhelmingly favored the use of risk assessment tools, a finding that may be much to the chagrin of legal commentators, they also admitted that certain prediction tasks, such as predicting future violence and distinguishing between "state and trait" in minor offenders (i.e., which offenders are "irreparably corrupt"), are beyond their perceived abilities. What this means for such standards, that is, what it means for the tasks that currently befall our judges, is a good question and one worth exploring further.

Lastly, not only does the article situate legal prediction both in history (legal philosophy) and in current and future developments (computational law), but it also fills a scholarly lacuna by providing a canon of legal prediction cases. In all, the article enables us to see how legal prediction is

---

[255]   567 U.S. 460 (2012).

waxing in importance. What was once a largely legal philosophical concern has become a pressing matter, one the impacts the functioning of the legal system and will shape its future form.

APPENDIX A

This appendix contains the survey questions from Part II, presented in the same order in which they were presented to the participants.

- When deciding a legal matter, I consider what the law requires in an absolute sense. (Please interpret "in an absolute sense" as indicating the following: the law transcends the question of how higher court judges would rule if faced with the matter.)
- When deciding a legal matter, I consider whether my decision would be reversed by a higher court.
- When deciding a legal matter, I attempt to replicate the result that would be reached if my state's highest court were faced with the same matter.
    - o Limit analysis to only those attorneys who are state-level judges.
- When deciding a legal matter, I attempt to replicate the result that would be reached if the Supreme Court of the United States were faced with the same matter.
- When a legal matter is being decided, which of the following should be the focus of the judges making the decision: (1) what result the law requires, or (2) how higher court judges would rule on the matter?
- If you were attempting to predict the result that would be reached if the Supreme Court of the United States were faced with a specific legal matter, as a basis for predicting the ruling, to what extent would you consider the following:
    - o published opinions by the justices
    - o the justices' non-judicial writings and speeches
    - o the justices' general ideological commitments
    - o your personal knowledge of and casual conversations with the justices, if any
- I can accurately predict how my judicial peers (specifically, those judicial peers whom I know well) will rule on specific legal matters.
- I can accurately predict how my state's highest court will rule on specific legal matters.
- I can accurately predict how the Supreme Court of the United States will rule on specific legal matters.

- The average attorney who knows me well can predict with accuracy how I will rule on specific legal matters.
- My judicial peers who know me well can predict with accuracy how I will rule on specific legal matters.
- It is currently possible to build a model (for example, a data-driven artificial intelligence) that can predict with accuracy how I will rule on specific legal matters.
- Within the next 10 years, it will be possible to build a model (for example, a data-driven artificial intelligence) that can predict with accuracy how I will rule on specific legal matters.
- On the whole, I can accurately predict how criminal defendants will behave during the post-arrest and pretrial period (i.e., if released on bail, whether they will commit additional offenses, whether they will show up for subsequent court dates, etc.).
- When trying to predict how criminal defendants will behave during the post-arrest and pretrial period (i.e., if released on bail, whether they will commit additional offenses, whether they will show up for subsequent court dates, etc.), it is helpful for me to have access to data-driven predictions, such as risk assessment tools.
- To what extent do you agree with this statement: I can predict whether a juvenile is "irreparably corrupt"; that is, I can predict whether a juvenile will be a danger decades from now and after a long prison sentence.

APPENDIX B

All data, with the exception of demographic information, which have been withheld in order to protect the privacy of our judicial participants, are stored on the Open Science Foundation portal at this location: https://osf.io/7tbx3/