Lawyer as Soothsayer: Exploring the Important Role of Outcome Prediction in the Practice of Law

Mark K. Osbeck*

ABSTRACT

Outcome prediction has always been an important part of practicing law. Clients rely heavily on their attorneys to provide accurate assessments of the potential legal consequences they face when making important decisions (such as whether to accept a plea bargain, or risk a conviction on a much more serious offense at trial). And yet, notwithstanding its enormous importance to the practice of law (and notwithstanding the handsome legal fees it commands), outcome prediction in the law remains a very imprecise endeavor.

The reason for this inaccuracy is that the three principal tools lawyers have traditionally relied on to facilitate outcome predictions—legal analysis, lawyerly experience, and the use of certain types of empirical information (e.g., jury verdict reporters)—are all subject to significant problems and limitations. This article examines in detail the reasons for these problems and limitations, concluding that they are essentially intractable. Thus, there is little hope that the traditional tools of outcome prediction on their own can ever enable consistently accurate assessments of potential legal outcomes.

Fortunately, however, recent advances in data science offer some grounds for optimism. Already, these advances are beginning to alter the way law firms operate, and there are good reasons to believe that data science (or more specifically, predictive analytics) will soon enable more accurate outcome predictions as well. Of course, predictive analytics is not a panacea: significant challenges remain if it is going to enable accurate outcome predictions on its own. And so it is doubtful that

* Clinical Professor of Law, University of Michigan Law School. I wish to thank the following colleagues for their very helpful comments on earlier drafts of this paper: Debra Austin, Ted Becker, Howard Bromberg, Michael Gilliland, Derek Kiernan-Johnson, David Moran, Nantiya Ruan, Gabrielle Marks Stafford, Todd Stafford, and David Thomson. I also wish to thank the Association of Legal Writing Directors and the Legal Writing Institute for their generous support of this project.

predictive analytics will supplant the traditional tools of outcome prediction in the foreseeable future. Rather, predictive analytics is likely to complement the traditional tools in order to power more accurate outcome predictions. However, even that modest change is likely to have a significant effect on the way lawyers practice law, and it should also come as very welcome news to their clients.

Table of Contents

I. INTRODUCTION

The practice of law requires lawyers to assume various roles.[1] The most celebrated of these roles is lawyer as advocate: where is, the lawyer stands in the client's stead, promoting the client's interests.[2] Less celebrated, but equally important, is the role of the lawyer as advisor.[3] In this role, "a lawyer serving as adviser primarily assists his client in determining the course of future conduct and relationships."[4] One of the most important tasks lawyers undertake in furtherance of this advisory role is outcome prediction: that is, advising the client as to the likely outcome of various legal proceedings.[5] In undertaking this vital task, the lawyer is required to analyze the various options and advise the client regarding the likely outcome of each so that the client can make an informed decision.[6]

Outcome prediction, therefore, is an essential lawyering skill. Lawyers, particularly litigators, cannot provide effective counsel to clients if they cannot accurately assess the potential outcomes of

1. *See* MODEL CODE OF PROF'L RESPONSIBILITY Preamble (AM. BAR ASS'N 1980) ("In fulfilling his professional responsibilities, a lawyer necessarily assumes various roles that require the performance of many difficult tasks.").

2. *See* MODEL RULES OF PROF'L CONDUCT rs. 3.1–3.9 (AM. BAR ASS'N 2016) (discussing professional responsibilities attendant to the lawyer's role as advocate).

3. *Id.* at rs. 2.1–2.4 (discussing professional responsibilities attendant to the lawyer's role as advisor). *See also* KARL N. LLEWELLYN, THE BRAMBLE BUSH 17 (Oxford Univ. Press 2008) (noting that when acting as business counsel, "the lawyer is interested . . . in anticipating what the court might do and in shaping his client's conduct to his client's desires in view of that anticipation.").

4. *See* MODEL CODE OF PROF'L RESPONSIBILITY EC 7-3.

5. *Id*. at EC 7-5 ("A lawyer as adviser furthers the interest of his client by giving his professional opinion as to what he believes would likely be the ultimate decision of the courts on the matter at hand and by informing his client of the practical effect of such decision.").

6. *Id.* at EC 7-8 ("A lawyer should advise his client of the possible effect of each legal alternative.").

litigation and other legal matters and advise their clients accordingly.[7] Outcome prediction is basically the legal equivalent of prognosis in medicine: an attempt to forecast the consequences of various courses of action so that the lawyer can help the client make informed decisions about matters of significant consequence to the client. It pervades the practice of law,[8] just as it does the practice of medicine. Every time a criminal defense attorney advises a client whether to accept a plea agreement; every time a civil litigator advises a prospective plaintiff whether to initiate a lawsuit, or to settle a lawsuit; every time a tax lawyer advises a client whether to take an aggressive deduction on the client's return—in all of these circumstances and many more, the lawyer is called upon to serve as a prognosticator as part of the lawyer's role as advisor.

Yet in spite of the enormous importance of outcome prediction to the practice of law, the academic legal literature is lacking any thoroughgoing analysis of how outcome prediction in the law actually works, and how it might be improved upon.[9] What literature there is concerning outcome prediction in the law has mainly been generated by scholars in cognate disciplines, such as artificial intelligence and political science, and it is primarily concerned with generating predictive models.[10] What remains wanting is a thorough understanding of the actual tools lawyers use to formulate outcome predictions and a critical assessment of their effectiveness.

7. *See* Oliver Wendell Holmes, *The Path of the Law*, 10 HARV. L. REV. 457, 457 (1897) (arguing that law is a profession precisely because people are willing to pay lawyers to advocate on their behalf and to advise them as to possible legal consequences they may face).

8. *See* MODEL CODE OF PROF'L RESPONSIBILITY EC 7-5.

9. This lack of attention in the literature may well be because, to put it simply, predictions are difficult, and lawyers have traditionally been less than stellar at making outcome predictions. *See* NANCY L. SCHULTZ & LOUIS J. SIRICO, JR., LEGAL WRITING AND OTHER LAWYERING SKILLS 184 (6th ed. 2014) ("[T]rying to predict what parties, witnesses, judges, and juries are likely to do is often little more than an educated guessing game."); *see also infra* Part IV.

10. *See, e.g.*, Lee Loevinger, *Jurimetrics: The Next Step Forward*, 33 MINN. L. REV. 455 (1949); Fred Kort, *Predicting Supreme Court Decisions Mathematically: A Quantitative Analysis of the "Right to Counsel" Cases*, 51 AM. POL. SCI. REV. 1 (1957); Glendon Schubert, *A Psychometric Model of the Supreme Court*, AM. BEHAV. SCIENTIST, Nov. 1961, at 14; Franklin M. Fisher, *The Mathematical Analysis of Supreme Court Decisions: The Use and Abuse of Quantitative Methods*, 52 AM. POL. SCI. REV. 321 (1958).

This Article attempts to fill the void in the legal academic literature concerning outcome prediction. It does so by first examining how outcome prediction has traditionally functioned in the practice of law, whereby lawyers have relied on three principal tools: (1) legal analysis (of a particular sort this article refers to as "element-focused analysis"); (2) lawyerly experience; and (3) the use of certain types of empirical information.[11] The Article then evaluates critically the effectiveness of these traditional tools, focusing on a variety of issues that impede accurate predictions.[12] The Article gives particular attention to problems that inherently afflict the "element-focused analysis" that lawyers have long relied upon to inform outcome predictions, as this topic has received almost no attention in the scholarly literature.[13] Lastly, the Article discusses how outcome prediction in the practice of law might be improved upon going forward, thanks to recent advances in data science.[14] It concludes that while the traditional tools that lawyers use to make outcome predictions (particularly the element-focused analysis) have a number of shortcomings that lead to significant inaccuracy, the new tools that rely upon predictive analytics offer a glimmer of hope that lawyers going forward will be better at making outcome predictions than they traditionally have been.[15]

Part II of this Article discusses in detail the reasons why outcome prediction is a vital part of practicing law. Part III looks at the tools lawyers have traditionally used to make outcome predictions. It examines in detail each of the three principal tools and the ways in which lawyers use them in practice. Part IV examines the problems and limitations that afflict the traditional tools of outcome prediction, again with a particular focus on the element-focused analysis that has traditionally played a vital role in outcome prediction, but which is largely ignored in the literature. And finally, Part V of this Article discusses the prospects going forward for employing predictive analytics to help lawyers make more accurate outcome predictions.

11. *See infra* Part III.
12. *See infra* Part IV.
13. *See infra* Section IV.A.
14. *See infra* Part V.
15. *See infra* Part VI.

II. THE IMPORTANCE OF OUTCOME PREDICTION TO THE PRACTICE OF LAW

The legal profession is not unique in its need for accurate outcome predictions. In a number of professional fields, practitioners need to be able to assess the likelihood of potential outcomes. In the field of medicine, for example, doctors need to make prognoses to properly assess treatment options; in the field of investment advising, stockbrokers strive to provide their clients an accurate assessment of a stock's likely prospects in the market; and in the field of sports, prognosticators are valued not only for the assistance they can provide gamblers but also for such things as evaluating the potential success of prospective players.

In the practice of law, lawyers need to assess the likely outcome of litigation matters for several important reasons. First, the decision whether to originate a litigation matter requires a reasonable balancing of costs versus expected benefits, and a significant component of this calculation is an estimation of the client's likelihood of success. Second, deciding whether to accept a settlement offer, whether in the criminal or civil context, depends upon a reasonable assessment of the likely outcome in the absence of a settlement. And third, outside of a litigation context, transactional lawyers often need to assess the likely outcomes of the various decisions confronting their business clients (for example, the prospects of litigation arising from a proposed business decision), and this too requires a reasonable prediction as to what is likely to happen if the client proceeds in a certain manner. For each of these reasons, as well as some other more minor reasons, outcome prediction forms an important part of a lawyer's role when the lawyer acts as an advisor rather than an advocate.[16]

A. The Importance of Outcome Prediction in Case Selection

First, outcome prediction is vital to efficient case selection. When a civil litigator or prosecutor is evaluating whether to initiate an action, the

16. *See* Daniel Martin Katz, *Quantitative Legal Prediction – Or – How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 EMORY L.J. 909, 912 (2013) ("[P]rediction is a core component of the guidance that many lawyers offer. Indeed, it is by generating informed answers to these types of questions that many lawyers earn their respective wages.").

lawyer needs to first assess the merits of the prospective case,[17] which in turn requires the lawyer to evaluate the likelihood of success.[18]

The lawyer has an ethical obligation not to pursue a spurious action[19] (or, in the case of a prosecutor, to refrain from prosecuting an action the lawyer knows is unsupported by probable cause[20]), and this requires an assessment as to the likelihood of winning. If there is little or no chance of success, then the lawyer needs to evaluate the lawyer's ethical obligations carefully.

However, even assuming these ethical obligations are satisfied, the lawyer must still make an outcome prediction to properly assess the case. First, the lawyer has a fiduciary obligation to act in accordance with the client's interests,[21] and this requires (among other things) a risk-benefit analysis balancing the costs of litigation against the possible recovery.[22]

Second, from the perspective of the lawyer's own pecuniary interests, outcome prediction is often important in determining whether the action is worth pursuing from the point of view of the lawyer or the lawyer's firm, particularly in contingency citations, when the lawyer has a stake in the litigation.[23] If, for example, a plaintiff's lawyer overestimates either the likelihood of success or the likely amount of the recovery, the client is not going to be happy with the result because the recovery is less than expected, and the lawyer's law firm will also be unhappy, if the firm has a contingent interest in the litigation and the ultimate recovery does not justify the firm's expenditure on the matter.

17. *See* George L. Priest & Benjamin Kline, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1, 4 (1984); *see also* THOMAS A. MAUET & DAVID MARCUS, PRETRIAL 86–91 (9th ed. 2015).

18. *See* Edie Greene & Brian H. Bornstein, *Cloudy Forecasts*, TRIAL, Apr. 2011, at 28, 29 ("When evaluating a case's potential, the lawyer weighs the costs and benefits based upon an educated guess as to the case's outcome.").

19. *See* MODEL RULES OF PROF'L CONDUCT r. 3.1 (AM. BAR ASS'N 2016).

20. *Id*. at r. 3.8.

21. *See* RESTATEMENT (THIRD) OF THE LAW GOVERNING LAWYERS §§ 49–50 (AM. LAW INST. 2000).

22. *See* Dru Stevenson & Nicholas J. Wagoner, *Bargaining in the Shadow of Big Data*, 67 FLA. L. REV. 1337, 1370–71 (2015) (describing research indicating that lawyers undertake this type of analysis in assessing the validity of prospective actions); *see also* Greene & Bornstein, *supra* note 18, at 29.

23. In a contingency action, the plaintiff's lawyer's fee is dependent on the recovery. Typically, such lawyers get a percentage of the final recovery (capped by state rules), with the percentage amount depending on whether and when the matter is settled, or whether it proceeds to trial and/or appeal. *See* MAUET & MARCUS, *supra* note 17, at 97–99.

Likewise, if a prosecutor overestimates the likelihood of success on a criminal action, this increases the risk of an acquittal that might otherwise have resulted in a plea bargain. And even if a plea is ultimately entered, the costs required to obtain that plea will likely exceed what they would have been had the prosecution made a more reasonable offer early in the litigation process.

Thus, in both the civil and criminal contexts, outcome prediction is an important part of the initial case assessment that takes place before an action is originated.

B. The Importance of Outcome Prediction in Making Settlement Decisions

Outcome prediction is perhaps even more important in the context of settlement negotiations, given that a critical component of rational negotiation is a reasonable assessment of the likely outcome of the case in the absence of a negotiated settlement agreement.[24] Imagine, for example, that you come home from work one night and waiting for you in the shadows are several police officers. They arrest you, charge you with a crime, and cart you off to jail. Hopefully, you are able to obtain bail and gain your release. And now it's time to begin working on your defense. At that point, you earnestly want a lawyer to assist you.

So, what would you look for in that lawyer? You might be inclined to look for a Perry Mason type—a brilliant trial lawyer who could prove your innocence at trial.[25] However, that may not be your wisest option. For in the vast majority of both criminal and civil cases, the outcome is determined not through jury trials, but rather through negotiation and plea bargaining.[26] And so you would probably be better served by

24. *See* George Loewenstein et al., *Self-Serving Assessments of Fairness and Pretrial Bargaining*, 22 J. LEG. STUD. 135, 136–37 (1993) (describing the generally accepted model of settlement, whereby cases that fail to settle are those in which the plaintiff overstates and/or the defendant underestimates the expected value of going to trial).

25. **Perry Mason** is a fictional American trial lawyer. He was the main character in numerous novels written by Erle Gardner. The character was also portrayed by actor Raymond Burr in a popular television series that originally ran from 1957–1966. *See Perry Mason*, ENCYCLOPEDIA BRITANNICA, www.britannica.com/topic/Perry-Mason (last visited July 30, 2018).

26. Laura A. Kaster, *Cognitive Barriers to Valuing Your Case for Settlement or Mediation: Improving Your Risk Assessment*, 269 N.J. LAW. 43, 43 (noting that "[o]ver 95% of litigated cases are settled.").

focusing on a lawyer who was skilled at negotiation and at providing you with sage advice as to the desirability of accepting whatever plea bargain that the prosecution may ultimately offer.[27]

In order to provide such counsel, your lawyer will have to properly assess your case, which involves a risk-benefit analysis. Specifically, your lawyer must balance the prospect of a sure adverse result (e.g., a one-year prison term) against a potentially worse adverse result (e.g., a 20-year prison term), *if* your defense fails at trial, and you are convicted.[28] And that, in turn, requires your lawyer to forecast both the likelihood of losing at trial, should you reject the prosecutor's plea bargain, and the length of the sentence you are likely to receive if you are convicted at trial.[29]

Reasonable outcome prediction is also essential to making wise decisions regarding settlement prospects in the civil context. In order to provide sage counsel as to the desirability of accepting any given settlement offer, a lawyer must be able to properly assess the odds of winning at trial and the potential ramifications of losing at trial.[30] Suppose, for example, that the lawyer is defending a company in a breach of contract action. The plaintiff seeks $1,000,000 in damages for the breach, and the defendant has made an offer of $100,000. In that situation, whether the $100,000 settlement offer is reasonable (from the defendant's perspective) depends, at least as a starting point, upon the likelihood of a plaintiff's verdict, the likely amount of any verdict, and the anticipated costs (primarily attorney's fees) of proceeding to trial.[31]

27. These two skills, however, are interrelated. A lawyer with a widespread reputation for strong trial skills is likely to have an advantage in terms of settlement clout over a less skilled trial lawyer because opposing counsel will be less inclined to take their chances at trial against a skilled trial lawyer.

28. *See* Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2463, 2496–527 (2004).

29. *See* Holmes, *supra* note 7, at 457 ("People want to know under what circumstances and how far they will run the risk of coming against what is so much stronger than themselves, and hence it becomes a business to find out when this danger is to be feared. The object of our study, then, is prediction, the prediction of the incidence of the public force through the instrumentality of the courts.").

30. *See* Robert H. Mnookin & Lewis Kornhauser, *Bargaining in the Shadow of the Law: The Case of Divorce*, 88 YALE L.J. 950, 968, 997 (1979); *see also* Stevenson & Wagoner, *supra* note 22, at 1375–77.

31. *See, e.g.*, Richard A. Posner, *An Economic Approach to Legal Procedure and Judicial Administration*, 2 J. LEGAL STUD. 399 (1973); William M. Landes, *An Economic Analysis of the Courts*, 14 J. L. & ECON. 61 (1971); *see also* Mnookin & Kornhauser, *supra* note 30, at 968.

The more inaccurate the defendant's assessment of the case, the more money the defendant is likely to lose, either by paying too much in the settlement, or by taking an unjustified risk at trial.[32] And the same is true of the plaintiff. The more inaccurate the plaintiff's assessment, the less the plaintiff is likely to recover, either by accepting too little in the settlement, or by taking an unjustified risk at trial.[33] So accurate outcome prediction is essential to a lawyer's ability to provide sound advice to the client regarding settlement prospects.

Furthermore, the ability of lawyers to effectively advise clients as to the likely outcome of litigation matters affects more than just individual clients. For unless both lawyers in a litigation matter properly assess the likelihood of a particular outcome, the efficiency of the settlement process itself suffers. Presumably, the goal of settlement is to avoid

32. In mathematical terms, the settlement offer ("S") is reasonable from the defendant's perspective if: $S \leq (P \times V) + C$ (where "P" is the probability of a plaintiff's verdict, "V" is the likely size of the verdict, and "C" is the cost of proceeding to trial). Thus, if the defense lawyer estimates the probability of a plaintiff's verdict (P) at 15%, the potential verdict amount (V) at $1 million, and the costs of additional litigation to trial (C) at $25,000, then the $100,000 settlement amount is reasonable, given that the value of $(P \times V) + C$ is $175,000 in that hypothetical. Conversely, if the plaintiff refuses to settle the case for $175,000, then it would be prudent for the defendant to take its chances at trial, assuming it isn't risk averse for some other rational reason (e.g., the company is uninsured and would be bankrupted by a $1 million verdict, but could absorb some lesser amount, such as $200,000). *See generally* Priest & Kline, *supra* note 17, at 12–13; MAUET & MARCUS, *supra* note 17, at 408–13 (discussing in depth this type of settlement calculation). This formula is overly simplified, of course. There are other potential costs, such as tax consequences, bad publicity, etc., that must be factored into the equation as well. Furthermore, it may sometimes be reasonable for a defendant (or its insurer) to offer less in settlement than the formula would indicate is reasonable, in hopes of deterring future litigation. On the simple formula set out above, even if the likelihood of a successful recovery is zero (because P is zero), it would still be rational for the defendant to settle for an amount greater than zero but less than the cost of going to trial, which is generally not insignificant. But if the defendant settles right away for nuisance value, as the simple formula would prescribe, there is little to prevent unscrupulous plaintiffs from filing frivolous lawsuits in hopes of scoring a quick settlement. The defendant has an incentive to require the plaintiff to prove up its case, which helps to eliminate spurious cases, and also requires the plaintiff to incur costs that would serve as a deterrent to filing a lawsuit. The formula is also overly simplified with respect to determining the expected value of V, the potential verdict. While this amount might be fairly definite in some types of actions, such as breach of contract actions in which plaintiffs seeks a set amount, in other types of actions, such as negligence actions, estimating the potential recovery is more difficult. In those types of actions, there will a range of potential verdicts (e.g., the defense lawyer may reasonably assess the expected verdict at anywhere from $100,000 to $1 million). *See* MAUET & MARCUS, *supra* note 17, at 408–09.

33. *See* Charles J. Snyder, *Moneyball Lawyering*, 65 ARK. L. REV. 837, 854 (2012).

uncertainty and wasted resources (in addition to mental stress) by short-circuiting the litigation process, so that roughly the same result is reached in settlement that would have been reached at trial, without the negatives of protracted litigation.[34] However, if either or both lawyers have unrealistic expectations of their client's likelihood of success at trial, then an efficient settlement won't be reached.[35] If, for example, the plaintiff's lawyer overestimates the likelihood of success, then some cases that shouldn't proceed to trial will; conversely, if the plaintiff's lawyer underestimates the likelihood of success and decides not to pursue a meritorious action vigorously, then some deserving plaintiffs will settle for less than fair compensation.[36]

Thus, the ability to make reasonably accurate predictions regarding litigation outcomes is key to the efficiency of our litigation system as a whole. And the same is true with respect to the efficiency of our criminal justice system. If prosecutors make errant judgments about the likely result of potential prosecutions, then the system will misallocate judicial and prosecutorial resources, resulting in fewer convictions of those who deserve to be convicted, and a greater waste of time and resources trying to convict those who merit lesser plea bargains.

C. The Importance of Outcome Prediction in Transactional Practice

While outcome prediction is most clearly at issue in litigation matters, the importance of outcome prediction is not confined to litigation. Transactional lawyers also need to assess potential outcomes to properly counsel their clients. A tax attorney advising a client whether to take a certain deduction, for example, must analyze both the likelihood

34. *See* MAUET & MARCUS, *supra* note 17, at 393 ("The law encourages settlement, and clients often prefer settlement over the increased expenses and uncertainties of a trial.").

35. *See* Jane Goodman-Delahunty et al., *Insightful or Wishful: Lawyers' Ability to Predict Case Outcomes*, 16 PSYCHOL. PUB. POL'Y & L. 133, 134 (2010) ("At the end of the day, it is the accurate predictions of the lawyer that enable the justice system to function smoothly . . . ."). A similar cost/benefit analysis is required for efficient litigation strategy. To decide whether a particular motion is warranted, for example, or whether it is worthwhile to pursue a certain type of evidence in support of a claim or defense, requires a balancing of the costs versus the anticipated likelihood of success. *See* Snyder, *supra* note 33, at 854.

36. Goodman-Delahunty et al., *supra* note 35, at 134–35; Kaster, *supra* note 26, at 43, 46.

of an audit as well as the likely outcome should the IRS decide to pursue an audit.

Often, potential litigation is an important contingency that individuals and companies have to consider when they enter into business transactions. If there is a possibility that the contemplated transaction will involve litigation, then both the potential consequences of that litigation and the expected costs of that litigation must be factored into the cost-benefit analysis as the client determines whether to go forward with the deal. And in order to properly assess the costs and benefits of the prospective transaction, the transactional lawyer (often in conjunction with a litigator) must make a prediction with respect to the likely outcome of the potential litigation.[37]

Suppose, for example, a group of investors is considering purchasing a tract of real estate in order to develop it into a private golf and ski resort. The client is interested in investing tens of millions of dollars to purchase a large tract of undeveloped land, which the client thinks could ultimately be developed into parcels worth several hundred million dollars in the aggregate. There is, however, one not-so-little hitch: title to the land is in dispute, and that title is the subject of a pending lawsuit. Both the client (i.e., the prospective buyer), and the seller believe that the seller's pending quiet-title action is likely to succeed, and yet it is crucially important to get a sense of just *how* likely the odds of success are, so that the client can make a rational decision whether to invest in the property at all, and if so, how much to invest. The difference between a 5 percent likelihood that the quiet title action would fail versus a 20 percent likelihood that it would fail will make a significant difference in the amount the client is willing to pay for the property. In this type of situation, the transactional lawyer's job is to provide (probably in conjunction with firm's litigation attorneys) the best possible outcome prediction as to the quiet title action so that the client can make an informed choice whether to proceed with the purchase.[38]

The transactional client's decision whether to proceed with a deal, and if so, how much to invest in it, may also hinge on outcome

37. *See* Peter A. Antonucci & Lianne S. Pinchuk, *The Importance of Product Liability Risk Assessment in Business Valuation and Acquisitions*, THE METROPOLITAN CORPORATE COUNSEL, Sept. 13, 2017, at 1 (arguing that transactional lawyers should enlist experienced litigators in assessing product liability risks in connection with business valuations and acquisitions).

38. *See id.*

predictions regarding other types of proceedings, such as the prospects for proposed legislation or regulatory action.[39] For example, a company's decision whether to build a new plant in a particular location may hinge in large part on a proposed regulation that affects potential liability for environmental concerns. In that instance, the transactional lawyer may turn to the firm's regulatory lawyers for guidance as to the likely prospects for agency approval.

There are other contingencies transactional lawyers need to consider as well. They need to make predictions as to uncertainties, such as whether necessary licenses and permits can be obtained, and if so, how quickly; whether the U.S. Department of Justice will approve a proposed merger; or whether adequate financing will be available to fund the transaction. All of these require the transactional lawyer to engage in some degree of prognostication.[40]

Thus, outcome prediction is a vital component of client counseling. The client will have difficulty making important decisions unless the client has confidence in the lawyer's ability to make accurate outcome predictions. While this article focuses mostly upon outcome prediction in the context of litigation, the importance of outcome prediction is not limited to that area of the law, and much of what is discussed below applies in the transactional context as well.

III. THE TRADITIONAL TOOLS OF OUTCOME PREDICTION

The principal tools lawyers have traditionally used to predict case outcomes are: (1) an element-focused analysis of each asserted cause of action and defense in the case, looking to prior decisional law to determine whether these elements are met; (2) lawyerly experience; and (3) certain types of empirical information that may provide insight into how a prospective judge or jury would decide the instant matter. This Part examines the nature of these tools, and Part IV examines their shortcomings.

39. Electronic legal research providers are now providing tools to assist with this type of outcome prediction (e.g., the LexisNexis *Legislative Outlook* tool). *See* Press Release, *LexisNexis Debuts Legislative Outlook and Moves Extensive News Archive to Lexis Advance*, LEXISNEXIS, https://bit.ly/2NYqq2q (last visited July 30, 2018).

40. *See* LINDA H. EDWARDS, LEGAL WRITING & ANALYSIS 3–4 (3d ed. 2011).

A. The Element-Focused Analysis

The foundational tool lawyers have traditionally used to assess cases, particularly in the early stages of a dispute, is an "element-focused analysis" of the causes of action and defenses. In undertaking this type of analysis, the lawyer anticipates the process the trier of fact will need to follow in its assessment of the claim by analytically breaking down the cause of action or defense into its constituent elements, then determining for each element whether it applies in light of the known facts in order to predict the likely outcome.[41] Occasionally, courts will employ other types of tests (e.g., a factor test[42]) in reaching their determinations, but the most fundamental type of analysis in judicial decision-making revolves around an analysis of the elements of the various causes-of-action and defenses. The lawyer then goes through the same process with respect to each potential defense. In other words, when assessing the viability of a claim or potential claim, the lawyer examines, for each cause of action that has or might be alleged, the various elements and defenses that are applicable, and then makes a separate assessment as to the viability of each such element or defense. The lawyer can then assess the likely outcome of the cause of action overall, since (by definition) a cause of action fails if any element is not met, or if any (complete) defense applies.[43]

41. *See* Stevenson & Wagoner, *supra* note 22, at 1342; *see also* RICHARD K. NEUMANN, JR. ET AL., LEGAL REASONING AND LEGAL WRITING 9–16 (8th ed. 2017) [hereinafter NEUMANN ET AL., LEGAL REASONING].

42. In an equitable matter, such as a child custody determination, for example, courts balance various *factors* such as the parents' employment status, or a parents' drug/alcohol abuse, in deciding what custody arrangement is in the best interests of the child. Presumably, however, judicial decisions involving such factor tests are even more difficult to predict than those involving element focused analyses, since they involve a complicated weighing and a balancing of the various factors to determine whether the standard applies, and not just a "checking off" of those requirements that need to be met to fall within the scope of a legal rule. *See generally* Kathleen M. Sullivan, *The Justices of Rules and Standards*, 106 HARV. L. REV. 22, 58–59 (1992); Kevin H. Smith, *Practical Jurisprudence: Deconstructing and Synthesizing the Art and Science of Thinking Like a Lawyer*, 29 U. MEM. L. REV. 1, 58–61 (1998) (discussing how standards operate in legal analyses and arguing that legal standards are less constraining than rules).

43. *See* Smith, *supra* note 42, at 47–57 (discussing how elements operate in applying legal rules to facts, and arguing that rules are essentially conditional (i.e., "if/then" statements) comprised of triggering conditions in the form of elements); *see also* Stevenson & Wagoner, *supra* note 22, at 1342 ("This application of law to facts would yield an estimate about probabilities: that is, a prediction of the likelihood that a given rule would govern a given scenario.").

Suppose, for example, that a potential client wants to bring a cause of action for the tort of intentional infliction of emotional distress. In most states, a cause of action for the tort of intentional infliction of emotional distress (IIED) requires the plaintiff to establish four elements: (1) that the conduct was intentional or reckless, (2) that the behavior in question caused the emotional distress, (3) that the defendant's conduct was "outrageous," and (4) that the resulting emotional distress was severe.[44] The traditional element-focused analysis proceeds by evaluating each of these elements in turn, trying to assess whether, in light of the known facts, each one would be deemed applicable, were the court or jury to evaluate the potential case. Then, having analyzed each element (and any applicable defenses, such as the defense of privilege), the lawyer tries to make a projection as to the likelihood of success of the cause of action as a whole, based on the likelihood that each component element is met.

In determining whether the elements of a cause of action are met, the traditional analysis evaluates each element primarily in light of the case precedents interpreting that element.[45] Thus, the traditional analysis relies heavily upon legal research to find precedents that can be compared and contrasted on their facts with the instant case to determine whether the element is met.[46] Suppose, for example, that the potential defendant in our hypothetical IIED case is a teacher who yelled and cursed at the potential plaintiff, who was a 12-year-old student. Does that behavior constitute "outrageous conduct" for purposes of establishing the third element of the tort of intentional infliction of emotional distress? To answer this question, the lawyer traditionally starts by researching case law in the appropriate jurisdiction (or assigning the research to an associate), looking for IIED cases that shed light on the meaning of "outrageous conduct."[47]

44. Most states follow the RESTATEMENT (SECOND) OF TORTS § 46 (AM. LAW INST. 1965), which lays out these four elements of the tort.

45. The elements themselves may derive either from a textual source (e.g., a statute) or from the common law (e.g., a cause of action in tort, such as IIED). *See* Smith, *supra* note 42, at 49.

46. *See* Frederick Schauer, *Precedent*, 39 STAN. L. REV. 571, 576–79 (1987) (describing in depth how precedent functions at a theoretical level) [hereinafter Schauer, *Precedent*]; *see also* Stevenson & Wagoner, *supra* note 22, at 1371–72.

47. Case-law precedents are not the exclusive interpretive tools, however; other interpretive aids, such as scholarly commentary, may also be used to shed light on the meaning and applicability of the various elements. With respect to IIED, for example, the *comments* and *illustrations* that accompany RESTATEMENT (SECOND) OF TORTS § 46 are

Having located relevant precedents on this issue, the lawyer then looks at two things: (1) whether the courts promulgate any rules, factors, or principles (explicit or implicit) that outline the boundaries of that element (e.g., rules that define "outrageous conduct"); and (2) whether the defendant's behavior in the instant case is similar to the behavior of the defendants in those cases in which courts have found the element to be met.[48]

As regards the first criterion, the law in most states follows the Restatement in requiring a very high threshold for outrageous conduct: it must be conduct "so outrageous in character, and so extreme in degree, as to go beyond all possible bounds of decency, and to be regarded as atrocious, and utterly intolerable in a civilized community."[49] The second criterion—factual similarity between the precedent case and the instant case—is key for fleshing out the applicability of abstract rules, such as the one quoted above. If the factual circumstances are sufficiently and relevantly similar, then the lawyer concludes that the element of outrageous conduct is likely satisfied, and will move on to the next element; if the behavior is not on par, then the element is not met.[50] Once a determination is made with respect to each element, the cause of action as a whole can be evaluated because a cause of action exists only if every element is established.[51]

Of course, the degree of confidence the lawyer has in a particular outcome prediction depends upon how confident the lawyer is with respect to each element. While lawyers typically don't assign percentages to the individual elements (e.g., a 60 percent chance the jury will find the element of causation is met), they do tend to qualify their determinations broadly (e.g., it is "highly likely" or just "more likely than not" that the jury will find the conduct to be outrageous).[52] And this, of course, affects their assessment of the cause of action as a whole. Thus, if the lawyer feels that it is highly likely the jury will find each

generally considered important persuasive authorities for interpreting the elements set out in the model rule.

48. *See* Smith, *supra* note 42, at 40–46 (discussing the synthesis of legal rules from precedents and their application to facts).

49. *See* RESTATEMENT (SECOND) OF TORTS § 46 cmt. d.

50. *See* Smith, *supra* note 42, at 45 ("[T]he doctrine of stare decisis or precedent is quite complex, but it can be reasonably captured by a single phrase: similar facts, same law, same result.").

51. *See* NEUMANN ET AL., LEGAL REASONING, *supra* note 41, at 121.

52. *Id.* at 164.

element of IIED to be met, then the lawyer can be quite confident in predicting that a cause of action for IIED will succeed.[53] Conversely, if the lawyer determines that it is just slightly more likely than not that one or more of the elements is met, then the lawyer will make a less confident prediction.

Research memoranda (a.k.a. "legal memoranda" or "formal office memoranda") have traditionally been the vehicles through which lawyers record and convey their outcome predictions.[54] Further, for many years, researching and preparing such memoranda occupied the lion's share of a typical junior attorney's time.[55] Traditionally in these memoranda, the lawyer started with a question presented and a short answer to the question presented. This was followed by a summary of the facts, and then a detailed element-by-element analysis of one or more causes of action and/or defenses, followed by a brief conclusion that assessed the viability of the overall action.[56] Some experienced lawyers, however, may prefer that the lawyer undertaking a research project confine the scope of the memorandum to explaining the legal requirements for each individual element or defense, leaving the overall analysis as to the viability of the overall cause of action to the senior lawyer. Frequently, the research memorandum will then form the basis for an advice letter to the client, through which the senior lawyer can convey the results of the

53. *But see infra* Section IV.A.6.

54. *See* NEUMANN ET AL., LEGAL REASONING, *supra* note 41, at 159 ("An office memorandum predicts how the law will treat the client."); JOHN C. DERNBACH ET AL., A PRACTICAL GUIDE TO LEGAL WRITING & LEGAL METHOD 259–60 (4th ed. 2010); EDWARDS, *supra* note 40, at 131 ("Making an accurate prediction, then, is the function of an office memo.").

55. Kirsten K. Davis, *"The Reports of My Death Have Been Greatly Exaggerated": Reading and Writing Objective Legal Memoranda in a Mobile Computing Age*, 92 OR. L. REV. 471, 472–74, 482–83 (2013) (summarizing the traditional use of the office memorandum and describing a survey of law school graduates and their continued use of traditional legal memoranda, as well as more contemporary alternatives, such as short email memos).

56. RICHARD K. NEUMANN, JR. ET AL., LEGAL WRITING 123–26 (3d ed. 2015). Of course, legal memoranda have other possible uses as well. For example, a legal memorandum can be used merely to summarize the law on a particular topic, without applying that law to the facts in question. Similarly, it can be used to merely make the best arguments the clients can make in light of the law and the facts, without necessarily trying to predict a likely outcome. However, the main use of a legal memorandum traditionally has been to assess the client's case and to predict the likely outcome. NEUMANN ET AL., LEGAL REASONING, *supra* note 41, at 159.

element-focused analysis to the client, and advise the client accordingly.[57]

Most often, this analysis takes place at the beginning of the litigation process, where the lawyers for the parties are trying to ascertain how they should respond to a potential litigation matter (e.g., whether to file a claim if they represent the plaintiff, or whether to make an early settlement offer if they represent the defendant).[58] As discussed in Part II above, the plaintiff's lawyer needs to evaluate the potential amount and the likelihood of a potential jury verdict in order to determine whether filing the action is justified, and the defendant's lawyer needs to assess the client's potential exposure in order to evaluate early settlement options and in order to set a strategy and budget for case management purposes.[59]

In recent years, such formal legal memoranda have been used less frequently by lawyers, mainly due to the cost of preparing them.[60] As clients have become more cost-conscious, lawyers have tried to become more cost-efficient, relying less on formal memoranda and more on shorter, informal memoranda, email memoranda, and oral research reports.[61] So while it may still make financial sense for a law firm to have an associate prepare a traditional office memorandum in a high-stakes matter, where cost-containment is not a pressing consideration, it may not make sense in a more mundane litigation matter.

Still, regardless of the vehicle through which lawyers convey their analyses, the element-focused analysis, based on legal research, has formed the backbone of the traditional approach to outcome prediction and remains an important predictive tool. It is still one of the principal tools lawyers use to assess cases, particularly at the beginning of a litigation matter.[62] It is also one that can at least partially be delegated to

57. EDWARDS, *supra* note 40, at 4, 131; CHARLES R. CALLEROS, LEGAL METHOD & WRITING 207, 210–11 (7th ed. 2014); DERNBACH ET AL., *supra* note 54, at 259.

58. *See* MAUET & MARCUS, *supra* note 17, at 5–15 (discussing the role of element-focused analysis in the initial case-evaluation process).

59. *See supra* notes 25–35 and accompanying text.

60. *See* Kristen Konrad Robbins-Tiscione, *From Snail Mail to Email: The Traditional Legal Memorandum in the Twenty-First Century*, 58 J. LEGAL EDUC. 32, 32–36 (2008) (describing survey results showing a reduced use of formal legal memoranda in the practice of law).

61. *Id*. at 32–36, 41–42.

62. NEUMANN ET AL., LEGAL REASONING, *supra* note 41, at 40–45 (discussing the steps involved in an element-focused analysis and describing it as the principal tool of predictive writing).

more junior lawyers, as it takes full advantage of the legal research, writing, and analysis skills that law students develop in law school.[63]

B. Lawyerly Experience

Another important resource that lawyers rely on when making outcome predictions is *lawyerly experience*. Seasoned lawyers instinctively temper the predictive analysis of an associate's legal memorandum with their own experience in assessing the likely outcome of cases.[64] An experienced plaintiffs' lawyer, for example, may know from past experience that plaintiffs' verdicts for a cause of action such as IIED are relatively uncommon, and the experienced lawyer will temper accordingly the tendency of junior lawyers to skew the analysis in favor of the client.[65] Furthermore, the element-focused analysis contained in a typical research memorandum often sheds more light on how likely an action is to survive a motion to dismiss or a summary judgment motion, rather than the likelihood of a plaintiff's verdict at trial. In the case of IIED, for example, most of the reported cases are appeals from dismissals for failure to state a cause of action or appeals from summary judgment orders. Thus, they provide little guidance for how a jury is likely to resolve a matter that survives a dispositive motion and proceeds to trial.

An experienced lawyer may also consider other factors, besides the doctrinal considerations that are analyzed in a traditional element-focused analysis, in trying to predict the likely outcome of a litigation matter. For example, an experienced lawyer may take into account the background and perceived predilections of the individual judge(s) involved in the case particularly if the lawyer has personal experiences to

63. For this reason, learning to draft memoranda is still one of the principal topics taught in nearly all first-year legal writing classes, and it occupies an important place in nearly all first-year legal writing textbooks. *See, e.g.*, CALLEROS, *supra* note 57, at 189–343; DERNBACH ET AL., *supra* note 54, at 259–91; EDWARDS, *supra* note 40, at 131–45; NEUMANN ET AL., LEGAL REASONING, *supra* note 41, at 159–67.

64. *See* CLARENCE MORRIS, HOW LAWYERS THINK 11–19 (2d prtg. 1938) (discussing generally a lawyer's reliance on intuition).

65. *See* Suzanne E. Rowe, *Legal Research, Legal Writing, and Legal Analysis: Putting Law School Into Practice*, 29 STETSON L. REV. 1193, 1198 (discussing the tendency of law students to "skew their research or analysis to reach the answer they think the supervising attorney or the client wants"); Amanda Smith, *Preparing for Practice Beyond the Bench: Opinion Writing as the "Heart and Soul" of the First Semester of Legal Writing*, 18 J. LEG. WRITING INST. 263, 282 (2012) (arguing that most inexperienced researchers skew the analysis in favor of the client).

draw on with respect to these variables.[66] The experienced lawyer may also factor in non-doctrinal considerations such as the equities of the lawsuit, the sympathetic or not-so-sympathetic nature of the parties, the reputation of the opposing counsel, etc.[67]

In drawing upon experience to inform outcome prediction, the lawyer is not necessarily confined to the lawyer's own personal experience. Rather, the lawyer may draw upon the opinion of more seasoned lawyers (or expert consultants), in much the same way that a physician may draw upon the experience of more seasoned physicians in making a diagnosis.[68] In either case, anecdotal evidence tempers the purely legal emphasis of the element-focused analysis.

This anecdotal evidence thus enables the experienced lawyer to take a more holistic approach to outcome prediction.[69] The lawyer relies not just on a legalistic examination of the constituent parts of a particular cause of action, but also on a more "gestalt" view of the case (based on the lawyer's intuition) that takes into account a broader range of potentially relevant considerations.[70] Again, the parallel to medicine presents itself. Just as an experienced physician may rely not only on a checklist of symptoms in making a diagnosis but also on the physician's intuitive sense regarding the patient's overall presentation, so too the experienced lawyer may rely at least in part on whether the case "feels" like a winner, drawing on the lawyer's experience-based intuitions about the strength of the case.[71] This may well track (at least in part) the manner in which judges and juries reach decisions.[72] There is significant

66. *See* Daniel M. Schneider, *Assessing and Predicting Who Wins Federal Tax Trial Decisions*, 37 WAKE FOREST L. REV. 473, 509–14 (2002) (arguing that a judge's background traits, such as gender, education, and past work experience, are highly predictive of case outcomes); Stevenson & Wagoner, *supra* note 22, at 1342–43 (experiential knowledge supplements, and is sometimes more important than, the legal rules for purposes of outcome prediction).

67. *See* MAUET & MARCUS, *supra* note 17, at 406.

68. *See* Greene & Bornstein, *supra* note 18, at 31–32 (discussing research showing that lawyers who consult with experienced colleagues regarding their outcome predictions make more accurate predictions).

69. *See* Gerd Gigerenzer & Henry Brighton, *Can Hunches Be Rational?*, 4 J.L. ECON. & POL'Y 155, 172 (2007).

70. *Cf.* RICHARD A. POSNER, HOW JUDGES THINK 106-111 (2008) (describing the role of intuition and emotion in judicial decision making).

71. Stevenson & Wagoner, *supra* note 22, at 1346 nn. 24–25 (discussing research indicating that lawyers draw heavily upon their own experiences in making decisions).

72. POSNER, *supra* note 70, at 108 ("Thus, the more experienced the judge, the more confidence he is apt to repose in his intuitive reactions . . . ."); Joseph C. Hutcheson, Jr.,

anecdotal and other evidence that judges and juries do not decide cases merely by analyzing the individual elements and defenses; rather, they balance that analysis against their intuitive sense of what justice demands in an individual case.[73] If so, then balancing the element-focused analysis with the lawyer's intuitive sense as to the likely outcome of a case, where this intuition is ultimately derived from experience, would reasonably be expected to improve the accuracy of the lawyer's outcome prediction.[74]

C. Empirical Information

The third principal tool that lawyers have traditionally used to forecast case outcomes is *empirical information* about cases. While this tool has not historically been as widely used as the first two tools discussed above, empirical information is likely to become increasingly important in this age of data analytics, as discussed in Part V.

The empirical information that has traditionally been used to guide outcome predictions has been derived from several sources. The oldest and most widely used source is the *jury verdict reporter*, which summarizes jury verdicts by subject matter so that lawyers can see how similar cases have been resolved in the past and can gain an understanding of the expected verdict range in similar cases.[75] Jury verdict reporters are published in most jurisdictions. Jury verdict reporters are published in most jurisdictions and are prepared by private entities.[76] The information in these reporters comes from publicly available court records, as well as from the attorneys that were involved

The Judgment Intuitive: The Function of the "Hunch" in Judicial Decision, 14 CORNELL L. REV. 274, 284 (1929); *see generally* Mark C. Modak-Truran, *A Pragmatic Justification of the Judicial Hunch*, 35 U. RICH. L. REV. 55 (2001) (discussing the philosophical underpinnings of the "hunch" theory of judicial decision making).

73. *See* POSNER, *supra* note 70, at 110.

74. There is reason to believe that this type of reflective balancing between intuition derived from experience, on the one hand, and more deliberate analytical processes, on the other, is not confined to legal reasoning, but may instead be a fundamental feature of human cognition. *See* Gigerenzer & Brighton, *supra* note 69, at 156 ("Simple heuristics that ignore information can be better—faster, more frugal, and more accurate—than complex strategies that use all available information."). *But see* Davis, *supra* note 55, at 494–499 (describing some cognitive biases that may creep in when lawyers rely on intuition to evaluate potential case outcomes).

75. *See* MAUET & MARCUS, *supra* note 17, at 406.

76. Two of the largest commercial jury verdict reporters are VerdictSearch and the National Association of State Jury Verdict Publishers. *See* VERDICTSEARCH, www.verdictsearch.com (last visited July 30, 2018); NAT'L ASS'N STATE JURY VERDICT PUBLISHERS, www.juryverdicts.com (last visited July 30, 2018).

in the cases.[77] Generally, the reporters organize the case summaries by type of claim, type of injury, jurisdiction, amount, plaintiff's demographics, the insurer's settlement history, etc. Originally, jury verdict reporters were published in print in a newsletter format. Most of them are now available online as well, and many of them provide access to large online databases of case information that the lawyer can search by category.[78] Some of them offer research services as well, whereby a staff researcher will search for cases in the database that are on par with the case the lawyer is working on.[79]

The purpose of jury verdict reporters is to provide lawyers with information about how cases that are similar to the cases they are working on have been resolved. Lawyers can then use this information to make reasonable predictions as to the range of expected jury verdicts in similar cases.[80] Thus, the emphasis is different from the element-focused analysis discussed above in that jury verdict reporters provide guidance as to what is likely to happen after a case makes it to trial. The traditional element-focused analysis, by contrast, is geared more toward predicting whether a case will make it to trial at all, as it focuses primarily on the decisions of judges and not the actions of juries. Jury verdict reporters, therefore, provide an additional outcome-prediction tool that lawyers can use to assess both the likelihood of a plaintiff's verdict and the potential size of such a verdict in the event a case proceeds to trial.

Another source of empirical information that is more limited in terms of its accessibility is confidential settlement data. Attorneys who work with insurance companies, for example, have the benefit of the insurer's collected settlement and jury-verdict data from earlier cases the insurer has litigated to help inform their outcome predictions. In one respect, this information is narrower than that found in jury verdict reporters, insofar as it is limited to the cases handled by that insurer (although some insurers may elect to pool such information for their

77. *See, e.g.*, *Online Verdict Search Tool*, VERDICTSEARCH, http://verdictsearch.com/online-verdict-search-tool/ (last visited July 30, 2018).

78. Online commercial research providers like Westlaw and Lexis also offer verdict research tools. *See, e.g.*, *Westlaw Jury Verdicts*, THOMSON REUTERS LEGAL SOLUTIONS, https://tmsnrt.rs/2AqwJKt (last visited July 30, 2018); *LexisNexis Verdict & Settlement Analyzer*, LEXISNEXIS, https://www.lexisnexis.com/en-us/products/verdict-and-settlement-analyzer.page (last visited July 30, 2018).

79. *See, e.g.*, *Custom Verdicts and Settlements Research on Call*, VERDICTSEARCH, http://verdictsearch.com/custom-research/ (last visited July 30, 2018).

80. MAUET & MARCUS, *supra* note 17, at 406.

collective use). However, in other respects it is broader: for one thing, the insurers are generally going to have more extensive information about the facts, and for another, they have information concerning *settlements*, in addition to verdicts. This gives the lawyer a more comprehensive picture of possible case outcomes, as most cases settle prior to trial.[81] Jury verdict reporters, by contrast, are unable to obtain information about most settlements, due to the confidentiality clauses contained in most settlement agreements.[82]

The third source of empirical information that is used to inform outcome predictions is *jury research*. Lawyers can hire jury-research firms to consult with them on cases,[83] and these firms provide lawyers empirical information in two principal ways. First, they can provide information about jury behavior generally, based upon their own research. Second, they can empanel mock juries that sit through practice trials and evaluate the dispute in question firsthand. This allows the lawyers to try out different arguments and strategies and see how effective they are with the mock jury, and it also provides them information about how an actual jury is likely to resolve the dispute.[84] The information derived from a mock jury is of a different nature from that obtained from either jury verdict reports or from compilations of settlement data, insofar as the latter two sources focus on actual results from past cases, whereas a mock jury verdict relies on a hypothetical assessment of facts that are identical to the prospective case. The

81. *See* Kaster, *supra* note 26, at 43 & n.1.

82. Laurie Kratky Doré, *Secrecy by Consent: The Use and Limits of Confidentiality in the Pursuit of Settlement*, 74 NOTRE DAME L. REV. 283, 285 (1999) (observing that confidentiality agreements are frequently used to hide from public view the terms of settlements and the underlying facts); Scott A. Moss, *Illuminating Secrecy: A New Economic Analysis of Confidential Settlements*, 105 MICH. L. REV. 867, 867 (2007) ("Even the most hotly contested lawsuits typically end in a confidential settlement . . . .").

83. *See generally* Robert F. Ruckman et al., *Focusing Your Case Through Jury Research: Mock Trials and Other Tools*, THE BRIEF, Spring 2017, at 58 (describing the basic tools of jury research consultants).

84. *See* Jeh Charles Johnson, *Mock Juries: Why Use Them?*, LITIGATION, Winter 2009, at 32; Jerry W. Thomas, *Mock Juries*, DECISION ANALYST (1993), https://www.decisionanalyst.com/media/downloads/MockJuries.pdf. For a good summary of how mock juries are used and how the process proceeds, see Mary A. Bedikian & Jerome D. Hill, *The Ultimate Power of Persuasion: Using the Mock Trial to Enhance Litigation Strategy*, 72 MICH. B.J. 1046 (1993).

drawback to jury research is that it is very expensive, and therefore it is of limited availability to practitioners in many cases.[85]

The principal tools of case forecasting, then, are: (1) the traditional element-focused analysis based on legal research, (2) the experience of seasoned lawyers, and (3) empirical information about how similar cases have been resolved in the past, compiled from jury verdict reporters, and, in certain cases, from compilations of settlement data and jury research information. The following section of this article examines how effective these tools are in terms of predicting likely litigation outcomes, looking with particular focus at the traditional element-focused analysis.

IV. SHORTCOMINGS OF THE TRADITIONAL TOOLS

The tools lawyers have traditionally used to predict case outcomes have a number of limitations. As a result, outcome prediction—notwithstanding its major importance to the practice of law—has always been a rough science, its accuracy leaving much to be desired.[86] This section examines the limitations of the traditional tools of outcome prediction, with a particular focus on the element-focused analysis.[87]

85. *See generally* Thomas, *supra* note 84, at 1. Some research information regarding jury behavior in a generic sense is published in academic journals. *See, e.g.*, Robert J. MacCoun, *Experimental Research on Jury Decision-Making*, 244 SCIENCE 1046 (1989) (offering an example of an early research piece on jury behavior). But generic information of this type is not as valuable for purposes of outcome prediction as case-specific information.

86. *See* Kaster, *supra* note 26, at 44–45 (discussing research on the degree to which attorneys value cases incorrectly, and assessing reasons for the shortcoming); Goodman-Delahanty et al., *supra* note 35, at 133.

87. Notwithstanding the shortcomings of the traditional tools, clients continue to pay handsomely for lawyerly advice, indicating at least a market belief that lawyers' prognosticative skills have value. Brian Leiter, *Rethinking Legal Realism: Toward a Naturalized Jurisprudence*, 76 TEX. L. REV. 267, 312 (1997). But then again, people pay good money for tarot-card and palm readings as well, which probably says more about many people's strong desire to know the future than about the actual success of traditional case forecasting. In theory, the accuracy of lawyer outcome predictions could be tested. In fact, a recent study is enlightening as to the accuracy of the traditional tools. *See* Daniel Martin Katz et al., Predicting the Behavior of the Supreme Court of the United States: A General Approach 1 (July 27, 2014) (unpublished manuscript), https://arxiv.org/pdf/1407.6333.pdf (describing a forecasting model that correctly predicts with approximately 70% accuracy the outcomes of Supreme Court decisions, which is approximately the same success rate as expert Supreme Court watchers). Presumably, a similar study could be undertaken of trial court outcomes, using focus groups to test the accuracy of lawyers' predictions. But currently no such data are publicly available.

A. The Element-Focused Analysis

As a predictive tool, the traditional element-focused analysis lawyers use to forecast case results has a number of shortcomings. This is primarily because its accuracy depends upon an overly simplified view of how legal analysis works. In order to accurately predict how a prospective case will come out using the traditional element-focused analysis, the lawyer making the prediction must be able to rely on the consistent applicability of legal rules to known facts. In other words, it must be the case that the rules can be clearly ascertained, that the facts are known, and that relevantly similar factual contexts can be compared so as to determine the applicability of the rules. And it must be the case that the rules will be consistently applied in future cases, in the same way they were in past cases.[88]

There are several inherent problems with this approach, however, that hinder its reliability as a predictive tool. These include: (1) uncertainty as to the precise facts that should be applied to the analysis; (2) uncertainty as to the precise scope of the legal rules that should be applied to the analysis; (3) the difficulty in assessing the legal significance of certain facts; (4) the difficulty in accounting for non-doctrinal considerations that may affect the outcome of the case; (5) limitations in the types of information that can be derived from published opinions; and (6) the difficulty in making probability assessments in any precise way using the element-focused analysis. Most of these problems have been widely recognized—though not necessarily in the context of outcome prediction. The final factor, however, has not received significant attention in the scholarly literature and merits a more detailed examination. The remainder of this section discusses each factor in turn.

1. Factual Uncertainty

The first problem with the traditional element-focused analysis is that it relies on accurate factual comparisons between the prospective case and case precedents,[89] and yet there is frequently uncertainty as to the facts in a prospective (or even ongoing) case.[90] This is particularly a problem at the beginning of a case when the element-focused analysis is

88. Smith, *supra* note 42, at 15–16; Stevenson & Wagoner, *supra* note 22, at 1342.

89. *See* Smith, *supra* note 42, at 13–16.

90. JEROME FRANK, LAW AND THE MODERN MIND xii–xiii, xix (Peter Smith Publishing 1970) (1930).

often used to assess the viability of a particular cause of action. At this stage of the proceeding, the lawyer must rely primarily upon the factual account provided by the client, together with any additional information the lawyer can glean from any documents provided by the client, and any independent initial fact investigation the lawyer undertakes.[91]

The problem with relying on the client's account alone, of course, is that the client's account may be biased; thus, the facts that ultimately emerge at trial may not be in keeping with the story that the client reported to the lawyer during the initial client interview.[92] Furthermore, the lawyer's subsequent factual investigation, along with the discovery process, may reveal surprises. Unanticipated documents may turn up, and witnesses may provide somewhat different accounts of the facts than the lawyer may have anticipated at the beginning of the case. Additionally, the trial itself is often unpredictable. Witness credibility and likeability are important factors in the jury's assessment of the facts,[93] and it is difficult to work this information into an element-focused analysis, even if credibility can be accurately assessed pre-trial. Also, the trier of fact may not weigh the evidence the way the lawyer initially thought they would, and the judge may exclude or limit the use of certain evidence at trial that the lawyer was intending to rely on to build the case.[94]

Therefore, the difficulty of knowing in advance just how the finder of fact will weave the evidence into a particular narrative makes the application of the legal rules to the facts more difficult for the traditional element-focused analysis than might be apparent at first blush.

2. Legal Uncertainty

This difficulty is frequently compounded by uncertainty as to the legal rules. The traditional element-focused analysis used to assist predictions depends upon the ability of the lawyer to ascertain the controlling legal rules and apply them to the facts of the prospective

91. MAUET & MARCUS, *supra* note 17, at 89–90.

92. *See* MICHAEL E. TIGAR, NINE PRINCIPLES OF LITIGATION AND LIFE 240 (2009) ("The client may not level with you about the documents. The client may shade the truth.").

93. MAUET & MARCUS, *supra* note 17, at 405–06.

94. *See* Smith, *supra* note 42, at 23 ("[T]he court may not draw the same inference from the historical facts as you, your client, a witness, or the opposing party did.").

case.[95] Yet it is often unclear what exactly the parameters of the rules are and how exactly they apply to the prospective case.

For one thing, it is often difficult to synthesize a cogent legal rule from disparate cases.[96] The intellectual exercise of distilling a rule out of multiple cases that promulgate somewhat different and nuanced rules is not a determinative endeavor; rather, it is often possible to connect the dots in more than one way (that is, to formulate a synthesizing rule in different ways), and it is not always apparent in advance how a court will do so.[97]

Furthermore, the legal rules themselves can be vague or ambiguous. Hart's well-known hypothetical about a statute prohibiting "vehicles" in a park is a classic example of a vague textual rule; as Hart argued, it is not at all obvious from the mere meanings of the words whether a bicycle (or say a bicycle with a supplemental electric engine) is a "vehicle."[98] Thus, there is often uncertainty as to how a court will construe the "penumbra" around the core of a rule.[99] Furthermore, legal rules that are derived from cases, rather than textual sources, can be even more vague and indeterminate.[100] Such rules are generally highly dependent upon the particular factual context in which they arise and are subject to refinement and modification if the facts in subsequent cases are significantly different.[101]

To be sure, comparisons of the facts of precedents to the facts of the prospective case can often shed light on vague terms. For example, one

95. *See id*. at 10–23; *see also* Antonin Scalia, *The Rule of Law as a Law of Rules*, 56 U. CHI. L. REV. 1175, 1179 (1989) (discussing the link between clear legal rules and predictability).

96. NEUMANN ET AL., LEGAL REASONING,' *supra* note 41, at 18–19. For a classic legal-realist formulation of this argument, *see* FRANK, *supra* note 90, at 159–71 (arguing that precedents are inherently indeterminate in terms of possible rule syntheses).

97. FRANK, *supra* note 90, at 163 ("Every lawyer of experience comes to know (more or less unconsciously) that in the great majority of cases, the precedents are none too good as bases of prediction.").

98. H. L. A. Hart, *Positivism & the Separation of Law and Morals*, 71 HARV. L. REV. 593, 606–15 (1958).

99. *Id*. at 607; *see also* DERNBACH, *supra* note 54, at 44.

100. This makes it easier for advocates, who can employ some creativity in formulating the applicable rule, but it makes formulating a rule more challenging for outcome prediction because the lawyer essentially has to make a prediction as to how the reviewing court will formulate the rule. *See* NEUMANN ET AL., LEGAL REASONING, *supra* note 41, at 93–96, 104.

101. *See id*. at 73–74 (discussing how there are sometimes gaps in the law due to a lack of sufficient precedents on point).

can compare and contrast cases that have granted summary judgment to the defendant in IIED cases based upon the severity of the emotional distress. If a certain level of emotional distress was deemed not to be severe enough to satisfy the severity element in a particular precedent, then the lawyer can conclude that a similar or lesser level of emotional distress will not suffice in a prospective case. However, the precedents may still leave significant room for doubt. Because the number of precedents is frequently quite limited, there is often a significant gray area remaining into which the facts of the prospective case may fall. Legal rules as articulated by the courts are thus, by their nature, open-ended with respect to their potential applicability, and there frequently is not enough precedent to provide meaningful guidance as to how they will be applied in new circumstances.[102]

Finally, legal rules are not entirely static, and the governing decisional law interpreting a rule may evolve during the course of the case. What looked like a solid case at the beginning of a litigation matter may look less certain if an ensuing legal precedent reshapes the governing rule, and this too can affect outcome prediction.

Therefore, for these reasons, determining exactly what "the law" is that should be applied to a given factual scenario is often less straightforward than it might seem at first blush, further limiting the effectiveness of the element-focused analysis as a predictive tool.

3. The Difficulty in Assessing the Legal Significance of Certain Facts

A related difficulty arises from the challenge of trying to ascertain from reported court opinions exactly which facts are legally significant to a particular holding. This too adds an element of uncertainty to the element-focused analysis because it means that determining which potential precedents are really on point can sometimes be difficult.[103]

102. *Id*. at 19 ("[O]nce a rule has been formulated, situations will inevitably crop up that the rulemaking or did not anticipate her could not have been expected to contemplate."); *see also* Schauer, *Precedent*, *supra* note 46, at 576–79 (arguing that the ability of precedents to constrain future decision-making depends upon the ability of the decision-maker to determine that the legally relevant facts are similar).

103. *See* Schauer, *Precedent*, *supra* note 46, at 577–88 (discussing the factors that determine which facts are legally relevant).

Consider, as an example, the development of the tort of IIED in Florida case law. In *Slocum v. Fair Food Store of Florida Inc.*,[104] which was the first appellate case to consider whether such a tort existed in Florida, the Florida Supreme Court held that the tort, even if valid in Florida, would not apply to the factual context in which the defendant caused the emotional distress to the plaintiff merely by stating "you stink to me."[105] However, in the next case to consider the tort, *Korbin v. Berlin*,[106] the Florida Court of Appeals found that a cause of action *was* stated where the defendant told a six-year-old girl things like: "do you know that your mother took a man away from his wife"; "do you know that God is going to punish them"; and "do you know that a man is sleeping in your mother's room."[107] In *Korbin*, the court held that a cause of action was stated because a reasonable jury could find that these statements were "calculated to cause the child 'severe emotional distress.'"[108] The court concluded that "[t]he alleged statements and the manner and circumstances under which they were communicated municated [sic] to the child leave little room to doubt they were made with a purpose and intent to shame her, and to shock the sensibilities of this child of tender years."[109] Additionally, the court noted that "[r]elating, as they did, to the child's mother, the content and import of the statements" were sufficient to state a cause of action.[110]

From the Court of Appeal's holding, it is apparent that the age of the defendant (young Ms. Korbin was just six years old) was a legally significant factor to be taken into account in future cases.[111] In other words, a lawyer interpreting *Korbin* would reasonably know from the language of the court opinion that the bar for IIED is going to be somewhat lower for outrageous verbal statements involving children than it would be for adults. What is less clear, however, is what exactly the court means when it refers to the "content and import" of the statements as being a decisive factor.[112] For example, should a lawyer comparing a future case to *Korbin* on its facts also take into account the fact that the

104. Slocum v. Food Fair Stores of Fla., Inc., 100 So. 2d 396 (Fla. 1958).
105. *Id.* at 398.
106. Korbin v. Berlin, 177 So. 2d 551 (Fla. Dist. Ct. App. 1965).
107. *Id.* at 552–53.
108. *Id*. at 553.
109. *Id.*
110. *Id*.
111. *Id.* at 552–53.
112. *See id.* at 553.

defendant in *Korbin* alleged sexual impropriety on the part of the child's mother? Or is that fact immaterial? Would the holding in the *Slocum* case perhaps have been different if, instead of telling the defendant "you stink to me," the defendant had impugned the plaintiff's virtue, as in *Korbin*? A reasonable argument could be made (by analogy to defamation law, for example) that sexual accusations are more likely to be actionable than accusations of body odor. However, it is not apparent from the court's holding in *Korbin* that this is so, and thus, it is not apparent that the sexual nature of the statements is a legally significant fact that could be applied to a new set of facts.[113] In other words, the judicial opinion may not reveal all of the factual considerations the judge actually relied upon in reaching the decision.[114]

Thus, in addition to problems of factual and legal uncertainty, the element-focused analysis is hindered by the challenge in certain cases of determining precisely which facts have legal significance to a particular holding.

4. The Difficulty in Assessing the Significance of Non-Doctrinal Considerations

A further uncertainty in the traditional element-focused analysis stems from its difficulty in accounting for certain non-doctrinal considerations that may affect an outcome prediction.

First, there are economic and psychological factors that can skew the analysis by leading lawyers to take an overly optimistic view of the client's case. Most obviously, there is often an economic incentive to favor the client's position.[115] Thus, plaintiff's lawyers will have a tendency to view the cause of action (and, of course, the lawyer's own

113. *See* Emily Sherwin, *Judges as Rulemakers*, 73 U. CHI. L. REV. 919, 920 (2006) (describing how judges' reliance on simplifying heuristics leads decision-makers to "focus on facts that come readily to mind at the expense" of less apparent, but equally important, background factors).

114. *See* Kevin D. Ashley & Stefanie Brüninghaus, *Computer Models for Legal Prediction*, 46 JURIMETRICS J. 309, 315–16 (2006) ("Judges may not have disclosed the features that influenced their decision or stated their rationales accurately or completely."); *see also* Llewellyn, *supra* note 3, at 35 (arguing that insofar as "facts or factors not shown in the [judge's] report are at work . . . the opinion gives us a misleading picture of what happened, and therefore, misleading basis for prophecy of what will happen in the future").

115. This is sometimes called "optimism bias." *See* Katz, *supra* note 16, at 929; *see also* Kaster, *supra* note 26, at 45; Oren Bar-Gill, *The Origin and Persistence of Optimism in Litigation*, 22 OXFORD J. L. ECON. & ORG. 490, 491 (2006).

abilities)[116] through rose-colored glasses. This is particularly problematic in cases where the potential plaintiff will be paying an hourly fee; but even in more traditional contingent-fee cases, it can still play a role. Similarly, defense lawyers paid on an hourly basis have an obvious economic incentive to prolong a litigation matter, even if an early settlement might be in the client's interests.

In addition, several other cognitive biases may skew a lawyer's predictions.[117] For example, clients tend not to favor lawyers who are pessimistic or perceived to be overly sympathetic to the opponent's position.[118] Taken together, these factors create a tendency to unrealistically assess the client's prospects,[119] which can lead lawyers to reject as unreasonable settlement offers that may, in fact, be reasonable. Theoretically, these psychological considerations could be taken into account in tempering the conclusions of an element-focused analysis. However, they are difficult to tease out and nearly impossible to quantify.

The traditional element-focused analysis also fails to account for certain non-doctrinal considerations that may influence the decision-maker(s) (i.e., the court or the jury), and thereby affect the outcome of the case. These include the types of considerations that legal realists discussed at length a number of years ago, such as the personal biases of the judge or jury.[120] Judges and juries are not machines, and they cannot be counted on to apply legal rules to the facts in a purely mechanical

116. *See* Green & Bornstein, *supra* note 18, at 31 (arguing that "lawyers may express heightened confidence in their abilities in order to attract and maintain a clientele").

117. *See* Snyder, *supra* note 33, at 841–54 (discussing various cognitive factors that can distort predictions); *see also* Davis, *supra* note 55, at 495–98 (discussing different cognitive biases that affect case evaluations).

118. *See* Stevenson & Wagoner, *supra* note 22, at 1347 n.27 (discussing research indicating that lawyerly overconfidence may be a "necessary trait to attract and retain clients"); *see also* Kaster, *supra* note 26, at 45 ("If the client communicates the expectation of hearing only positive views, and the ability to go elsewhere if unsatisfied, client think is even more likely.").

119. *See generally* George Loewenstein et al., *Self-Serving Assessments of Fairness and Pretrial Bargaining*, 22 J. LEGAL STUD. 135 (1993); *see also* Goodman-Delahunty et al., *supra* note 35, at 139–43 (discussing research showing that lawyers are overly optimistic in predicting trial outcomes).

120. *See, e.g.*, FRANK, *supra* note 90, at xii–xiii (discussing the hidden, unconscious biases of trial judges that affect case outcomes). For a contemporary formulation of this view, see POSNER, *supra* note 70, at 10–11 (discussing various personal attributes that affect judicial decision-making).

manner.[121] Accordingly, the traditional-element analysis is hindered by its inability to accurately account for factors such as the likability and credibility of the parties that can affect the outcome.[122] It also neglects factors such as the reputation and success rate of the attorneys, the historical tendencies of the individual judge(s) assigned to the case, differences in the predilections of different courts (e.g., in different localities), etc.[123] Again, this is not an intractable problem for the traditional element-focused analysis, since the traditional analysis can be balanced against and adjusted for these non-doctrinal factors. However, it is very difficult to weigh such factors, and this makes it difficult to factor them into the traditional element-focused analysis.

5. Limitations on the Amount and Type of Information Available from Published Opinions

In addition to the issues discussed above, there is an inherent limitation on the usefulness of the element-focused analysis due to the nature of its source material: i.e., published judicial opinions. Since these opinions (particularly those responding to pre-trial motions) focus their attention primarily on the proper interpretations of the law, rather than factual applications and determinations of damages, their usefulness is primarily confined to determining whether a cause of action may exist, rather than whether the plaintiff may succeed at trial and the amount of potential recovery. In the IIED example discussed above, for instance, most of the published opinions address motions for summary judgment or to dismiss, and they are primarily concerned with whether the plaintiff properly states a cause of action for IIED. One seminal Florida Supreme Court case, *Metropolitan Life Insurance Company v. McCarson*,[124] addresses an appeal from a jury verdict.[125] But even that opinion is primarily concerned with whether IIED is a recognized tort in the state of

121. POSNER, *supra* note 70, at 8 ("Empirical scholars have found that many judicial decisions, by no means limited to the Supreme Court, are strongly influenced by a judge's political preferences, or by other extralegal factors. . . .").

122. *See* MAUET & MARCUS, *supra* note 17, at 405–06.

123. *See* Leiter, *supra* note 87, at 312 (arguing that lawyers use informal "psychological, political, and cultural knowledge" to help predict judicial outcomes). Also see the discussion in Part V of this Article regarding Professor Schauer's discussion of outcome prediction based on non-doctrinal factors. *See infra* notes 187–93 and accompanying text.

124. Metro. Life Ins. Co. v. McCarson, 467 So. 2d 277 (Fla. 1985).

125. *Id.* at 278.

Florida, and whether the plaintiff properly states a cause of action for that tort.[126] Thus, it provides only limited guidance as to what an IIED plaintiff can expect to happen if the case proceeds to trial.

The upshot is that the element-focused analysis is primarily useful as a tool for predicting whether the plaintiff's claim will survive the pretrial motion stage of the litigation and be allowed to proceed to trial. Lawyers have to look to other predictive tools, e.g., jury verdict reporters and settlement data, to formulate outcome predictions as to their clients' chances of success at trial, and the amount of any potential recovery.

6. The Difficulty in Making Probability Assessments in an Element-Focused Analysis

Finally, there is one other problem that is intrinsic to the element-focused analysis that has not received significant attention in the scholarly literature. That is the difficulty of assigning probabilities to the applicability of the individual elements in the traditional element-focused analysis, and the difficulty in making an overall probability assessment based upon these individual assessments.

As to the first issue, one of the primary tenets of the element-focused analysis is that outcome prediction is facilitated by breaking down a cause of action or defense into its constituent elements. In other words, the traditional analysis assumes that the viability of the cause of action or defense as a whole can best be determined by assessing the viability of each individual element. The viability of each individual element can, in turn, be assessed by looking at various IIED precedents discussing that element, and an overall assessment of the cause of action or defense will then flow from these individual assessments.[127] For example, a lawyer examining all the Florida precedents on the severity element of IIED and applying them to a new factual scenario tries to determine the likelihood that the court or jury in the present case would find that the severity element was satisfied. Then, after doing the same with respect to each of the other three IIED elements, the lawyer will be in a good position to judge whether the cause of action as a whole will succeed. Thus, the more contingent elements the plaintiff has to prove,

126. *Id*. at 278–79.
127. *See* MAUET & MARCUS, *supra* note 17, at 5–7.

the lower the likelihood, other things being equal, that the defendant will prevail.[128]

For the reasons discussed in the preceding subsections, however, determining whether an element is satisfied is often not something that can be ascertained with any reasonable degree of certainty. For example, as discussed above, merely by looking at the various IIED precedents in which the severity element was found to be applicable or not, the lawyer cannot always assess the likelihood that the trier of fact will find the severity element to be met in the prospective case. Determining the likelihood that a particular court or jury will find a particular element to be met is far from a precise science.[129]

But even assuming that lawyers can make reasonably accurate probability assessments with respect to the individual elements in a cause of action or defense, there remains a further obstacle: determining the likelihood that the cause of action or defense as a whole is likely to succeed, based upon the likelihood that the finder of fact will determine that each individual element is met. This task raises some difficult theoretical issues that seem, practically speaking, to be intractable.

To see why, consider again the IIED example discussed above. To establish a cause of action for IIED, a plaintiff must convince the finder of fact that *each* of the four required elements is met: (1) the action was intentional; (2) the action *caused* emotional distress; (3) the emotional distress was severe; and (4) the action giving rise to the distress was "outrageous," as that term has come to be defined through decisional law.[130] If the defendant succeeds in convincing the finder of fact that any one of these elements is not met, then the defendant prevails. Now assume that the plaintiff's lawyer believes that there is an 80 percent likelihood of persuading the finder of fact with respect to each of the four elements. Is it likely that the action as a whole will succeed? And if so, what is the probability?

At first blush, it may seem that the answer to the first question is clearly "yes," even if the answer to the second question is not obvious. But in fact, even the answer to the first question is complicated. To assess the likelihood that the cause of action for IIED will succeed, the

128. *See* David A. Moran, *Jury Uncertainty, Elemental Independence and the Conjunction Paradox: A Response to Allen and Jehl*, 2003 MICH. ST. L. REV. 945, 950 (2003).

129. *See* Goodman-Delahunty et al., *supra* note 35, at 149–50.

130. *See* RESTATEMENT (SECOND) OF TORTS § 46 (AM. LAW INST. 1965).

starting point in a probability analysis would be to consider each separate element as an independent variable, just as one would with a series of coin tosses. For example, the probability of getting four "heads" in a row flipping a coin is only 1/16, even though the probability of getting "heads" with respect to each individual toss is 1/2. In other words, the basic rule in determining the probability that a *series* of (independent) events will occur is determined by multiplying the odds that each individual event will occur.[131] Accordingly, in the IIED example given above, the starting point for determining the probability that the action as a whole will succeed is determined by multiplying the likelihood that each individual element will be satisfied. Thus, for this calculation, the probability that the trier of fact will find all four elements to be met is $0.8 \times 0.8 \times 0.8 \times 0.8 = 0.41$ or 41%, which means that the plaintiff is actually more likely to lose than win.

It is tempting to conclude from this that there is a tendency for plaintiffs' lawyers to make overly optimistic outcome predictions, since what at first blush looked intuitively likely—that the cause of action as a whole would probably succeed—is in fact unlikely.[132] However, even that conclusion is questionable because it is based upon an assumption, as noted above, that the individual variables (i.e., the elements of the tort) are *independent* of each other, when often there is probably some degree of interdependence between them, due to the holistic nature of judicial decision-making.[133] In other words, the factors that influence the court's

131. *See* IAN HACKING, AN INTRODUCTION TO PROBABILITY AND INDUCTIVE LOGIC 41–43 (2001).

132. This is related to an evidentiary quandary that has sometimes been referred to as the "conjunction problem" or the "conjunction fallacy." *See* Saul Levmore, *Conjunction and Aggregation*, 99 MICH. L. REV. 723, 724 (2001).

133. *See* Moran, *supra* note 128, at 946. Furthermore, there is a counter-veiling factor at work when a plaintiff brings multiple claims for recovery. The plaintiff in a civil matter can (and often does), allege more than one cause of action in a lawsuit. *See* MAUET & MARCUS, *supra* note 17, at 61–64 (discussing the pros and cons of asserting multiple claims). Suppose, for example, that the IIED claim is one of only three causes of action the plaintiff has alleged in an action, and suppose further that each cause of action is determined to have a 41% chance of success. Then the odds that the plaintiff will *fail* on each count is accordingly 59% (1 – 0.41), and thus the odds that plaintiff will fail on all three counts is, per the multiplication rule discussed above, $0.59 \times 0.59 \times 0.59$, which comes out to a probability of 0.205, or 20.5%. *See* HACKING, *supra* note 131, at 42 (discussing the multiplication rule of probability). Similarly, the odds that the plaintiff will *succeed* on at least one of its three causes of action alleged is accordingly 0.795, or 79.5% (1 – 0.205). *Id.* Therefore, even though plaintiff is likely to fail on each of the three causes of action considered independently, it follows from the multiplication rule of probability that the plaintiff is quite likely (almost 80%) to succeed on at least *one* of the

resolution of one element may well affect its assessment of the other elements too.[134] However, multiplying the individual probabilities of the variables in a series to determine the probability that the series as a whole will occur is appropriate only where the variables are independent, i.e., where the resolution of each variable has no effect of the resolution of the other variables.[135] For example, the probability of getting four "heads" in a row in a series of coin tosses is determined by multiplying the odds of each individual coin toss (1/2 or 50%) because no one coin toss affects any other coin toss.[136] However, where one event may affect the likelihood that another will occur, a simple multiplication of the individual probabilities will not prove accurate, and a more complicated calculation is necessary.[137]

Suppose, for example, that two evenly matched teams are in a World Series. Prior to the first game, statisticians would reasonably assign a 50% likelihood that a given team would win with respect to each individual game. If Team A wins the first three games, it is reasonable to assume that the odds of Team A winning the fourth game will be deemed higher than 50%, in order to take into account factors such as momentum and demoralization of the opponent. But how much higher would the odds be? The answer to that question cannot be determined purely mathematically, though empirical data about how prior teams performed when down 3-0 in a World Series could shed some light on it.

Bringing this back to the traditional element-focused analysis, there is reason to believe that the individual variables in the element focused analysis (i.e., the likelihood that the trier of fact will find each element to be satisfied) are similarly interdependent. In other words, the trier of facts' resolution of one element may well affect the resolution of one or more other elements.[138] As discussed above in Section III.B, there is reason to think that experienced lawyers tend to analyze cases in a more holistic way, rather than merely parsing the likelihood that each element

plaintiff's causes of action. Again, however, this calculation assumes the independence of each individual cause of action, which is probably not accurate for the same reason the individual elements are often not independent. *See* Moran, *supra* note 128, at 946.

134. *See* Moran, *supra* note 128, at 946.

135. *See* HACKING, *supra* note 131, at 42.

136. JOHN W. FOREMAN, DATA SMART: USING DATA SCIENCE TO TRANSFORM INFORMATION INTO INSIGHT 81 (2014).

137. *See id*. at 80–82 (discussing conditional probabilities and Bayes' Rule).

138. *See* Moran, *supra* note 128, at 950–52.

of the cause of action will be satisfied.[139] So too there is evidence that judges and juries decide cases in a more holistic way, and that they balance a more intuitive view as to which party should prevail in a case with a strict analysis of the separate elements.[140] If so, then the trier of fact's resolution of one element would tend to affect the trier of fact's assessment of the other elements, and this interdependence would have to be factored into the outcome prediction, just as it would in the World Series example discussed above.

However, whereas statisticians can in that example determine (with some modicum of accuracy) the degree of interdependence of the variables by looking at past World Series results,[141] no such general data is currently available with respect to judicial decision-making. Thus, it will not be possible, practically speaking, for a lawyer to assess with accuracy the likelihood that a cause of action will succeed, even if it were possible to assign probabilities to each individual element. For while it can safely be assumed that the odds of success are higher than merely multiplying out the individual probabilities of the elements would lead us to believe (due to the likely interdependence of the elements),[142] there is no publicly available information as to how much higher. Thus, one of the fundamental axioms of the element-focused analysis is inherently flawed.

B. Limitations on the Other Traditional Tools of Outcome Prediction

The previous section discussed a variety of problems afflicting the traditional element-focused analysis that lawyers use to make outcome predictions. Fortunately, this is not the only predictive tool available to lawyers. Rather, as discussed above in Part III, lawyers, particularly

139. *See* Gigerenzer & Brighton, *supra* note 69, at 171–72.

140. *See* POSNER, *supra* note 70, at 107–09; *see also supra* notes 65–66 and accompanying text.

141. For example, statistical records may show that historically only one team in twenty-five (4%) that was down three games to none has ever won a World Series, whereas the odds that an evenly matched team would win four straight is normally 6.3% (*1/2 × 1/2 × 1/2 × 1/2 = 1/16* or 0.0625), which would provide some evidence of the momentum effect. The sample set is so small, however, that one could not rely with any degree of confidence on the empirical data. *See generally* HACKING, *supra* note 16, at 229–45 (discussing issues with statistical inferences based on small or unrepresentative samples). If, for example, a team overcame a 3–0 deficit the following year and won the World Series, the odds would change significantly based on the then-cumulative data—from 1/25 or 4%, to 2/26 or 7.6%—leading to the opposite conclusion.

142. *See* Moran, *supra* note 128, at 950–52.

seasoned lawyers, also rely heavily on experience—their own and that of other lawyers—in assessing potential outcomes, and they also have available to them certain empirical resources, such as jury verdict reporters, jury research, and, in some cases, settlement data. However, while these additional tools provide useful supplements to the traditional element-focused analysis, they are not without their own significant limitations.

1. Lawyerly Experience

The first supplemental tool available to lawyers in making outcome predictions is lawyerly experience—their own, as well as that of other lawyers they consult. For seasoned lawyers, in particular, personal experience is a very valuable tool, just as it is for seasoned physicians in making diagnoses and in predicting the course of various diseases. As discussed above in Section III.B, experience enables a lawyer to broaden the scope of the analysis, bringing in non-doctrinal considerations such as the lawyer's knowledge of a particular judge's propensities, or the tendencies of juries in particular localities to favor plaintiffs or defendants in certain types of cases.[143] Furthermore, experience enables lawyers to take a more holistic approach to outcome prediction, evaluating the big picture by relying on intuitions about likely outcomes that are honed from past cases.[144]

While experience is undoubtedly a valuable tool in the lawyer's arsenal, it is certainly not without its limitations.[145] For one thing, experience is obviously developed over time, so more junior lawyers will be quite limited in this respect. However, even for more seasoned lawyers, information derived from experience tends to be impressionistic, as it relies upon the accuracy of the lawyer's memory, and is a filtered interpretation of past events that may be influenced by the lawyer's own beliefs and biases about the law, and about people and institutions.[146] Additionally, personal experience is, by its nature,

143. *See supra* notes 115–23 and accompanying text.

144. *See* Gigerenzer & Brighton, *supra* note 69, at 171–72. For a discussion of the advantages and disadvantages of this reliance on lawyerly intuition, see Davis, *supra* note 55, at 494–99.

145. In fact, one study indicates that lawyers' accuracy in predicting outcomes is not significantly enhanced by experience. *See* Goodman-Delahunty et al., *supra* note 35, at 133.

146. *See* Stevenson & Wagoner, *supra* note 22, at 1340 n.12 (citing research indicating that lawyers "operate with beliefs and biases that can cloud their judgment");

limited.[147] Even in fields such as securities trading, while a broker's experience with how other clients have fared with respect to certain types of investments is certainly valuable, it is not a substitute for actual experimental data on stock performance that transcends the broker's personal experience.[148] And in law, personal experience is even more limited because lawyers have less to go on, given the relatively few clients most of them have as compared to, say, stock-brokers and physicians.

Furthermore, lawyerly experience as a predictive tool is subject to several of the same problems that afflict the element-focused analysis. Experience does not provide an end-around to the challenges of factual uncertainty and legal uncertainty. Lawyers are still reliant on limited sources of information, such as the client's account of facts, when initially assessing a case.[149] Likewise, the lawyer's past cases will never be entirely on par factually with a prospective matter, so how a court will apply a vague legal rule to a new legal situation remains subject to doubt.[150] Furthermore, the lawyer will not easily be able to get around the problem of identifying exactly which facts are legally significant, because a lawyer generally has only limited information about what facts really influenced the court or jury's decision.[151]

Thus, while experience is certainly a helpful guide to the lawyer, it is not by itself a particularly accurate source of outcome prediction in individual cases.

2. Empirical Information

The other principal predictive tool discussed in Part III was empirical information available to the lawyer, in particular, jury verdict

see also Kaster, *supra* note 26, at 44–45 (discussing the factors that can cloud a lawyer's judgment); Frederick Schauer, *Do Cases Make Bad Law?*, 73 U. CHI. L. REV. 883, 897–901 (2006) (discussing distortions that arise in judicial decision-making due to judges' reliance on simplifying heuristics that are based in part on cognitive biases).

147. *See* Snyder, *supra* note 33, at 849.

148. As a result, medical advice is heavily dependent on experimental data. This is true in a number of other fields as well, such as the stock brokerage industry. *See* Katz,*supra* note 16, at 948–49 (discussing the increasing reliance on data-driven decision-making in the stock-brokerage industry).

149. *See* MAUET & MARCUS, *supra* note 17, at 89–90.

150. *See supra* notes 97–102 and accompanying text.

151. *See* FRANK, *supra* note 90, at 119–24.

reporters, jury research, and, in certain cases, settlement data. While these too are helpful tools, they too have some significant limitations.

With respect to jury verdict reporters, they are limited in terms of the types of information they provide (and also, to some extent, in terms of their reliability, given that some of the information comes from the lawyers involved in the case). Often, for example, they provide only cursory factual summaries, which exacerbates the problems created by factual uncertainty and legal uncertainty.[152] As discussed in Section IV.A above, it is difficult to compare cases on their facts when factual information about the prospective case is of limited reliability (due to factual uncertainty), and the information about the precedent case is very limited in scope.[153] Furthermore, since the factual information available in jury verdict reporters is not generally presented in a very detailed manner, the ability of lawyers to draw generalizations about specific jury findings is limited.

As discussed in Section III.C above, some lawyers are able to draw upon case-specific jury research tools to supplement jury verdict reporters.[154] These have the advantage of providing very detailed and case-specific information about a prospective case and its likely outcome, and they also allow lawyers to experiment with different adversarial approaches.[155] However, jury research tools are also limited insofar as they do not address whether a cause of action is likely to survive a motion to dismiss or summary judgment; rather, they only shed light on the likely outcome in the event of a trial. In that respect, jury research tools suffer from a limitation that is the reverse of the limitation discussed above with respect to the element-focused analysis.[156] In addition, mock trials are one-sided as to the nature of the evidence, in that they do not allow the mock jury to hear the other side's actual case.[157] Mock trials are also limited insofar as they tend to be shorter than actual trials, relying on truncated evidentiary presentations.[158]

152. *See supra* note 77 and accompanying text.
153. *See supra* notes 89–102 and accompanying text.
154. *See supra* notes 83–85 and accompanying text.
155. *See* Johnson, *supra* note 84, at 32–33.
156. *See supra* Section IV.A.5.
157. DAVID TABAK, NERA ECONOMIC CONSULTING, SETTLEMENT REASONABLENESS FROM NEGOTIATION TO SETTLEMENT DISPUTES 4 (2012), http://www.nera.com/content/dam/nera/publications/archive2/PUB_Settlement_Reasonableness_0212.pdf.
158. *Id.*

Furthermore, as discussed in Part III, mock trials are very expensive, and thus they are generally limited to only high-dollar cases.[159] Even testing a case on a single mock jury or focus group requires a substantial expenditure, and to ensure greater accuracy, it would be necessary to try the prospective case to multiple mock juries, to eliminate possible idiosyncrasies of a single panel. However, the costs associated with doing that narrows the usefulness of mock juries as a predictive tool to a very small subset of cases.

Finally, as discussed in Section III.C above, settlement data can be used to assist lawyers in making outcome predictions. Such data is often valuable in assessing a potential client's exposure or expected recovery, particularly since the great majority of cases ultimately settle prior to trial. To be sure, settlement data will not provide direct information about the prospects for a case's surviving a dispositive motion, such as a summary judgment motion, nor will it provide direct information about the expected outcome at trial. Still, it is a valuable predictive tool that certain lawyers rely on heavily in assessing cases.[160] Relying on past cases, for example, an insurance defense lawyer can make a reasonably accurate assessment of the settlement value of a case and can make appropriate settlement decisions accordingly.

The main drawback to the use of settlement data as a predictive tool (as in the case with jury research) is that it is not readily available to most lawyers, since only certain clients, such as insurance companies and large corporations, face repeated litigation and are thus in a position to acquire large quantities of useful settlement data. Additionally, since the majority of settlements are confidential and not available to the public,[161] most lawyers have no way of tapping into this pool of information.

In sum, empirical information serves as a valuable predictive tool for lawyers, but its value is limited insofar as much of the most useful information is unavailable to the majority of lawyers who make outcome predictions.

V. USING DATA SCIENCE TO IMPROVE OUTCOME PREDICTION

Part IV of this article discussed the principal reasons why lawyers struggle with outcome prediction, using the traditional tools available to

159. *See* Thomas, *supra* note 84, at 1.
160. *See supra* notes 75–84 and accompanying test.
161. *See supra* note 82.

them. While those tools certainly have some predictive value, they are also subject to significant shortcomings and limitations that hamper their ability to provide helpful guidance to their clients concerning potential or pending legal matters. In the past few years, however, a potentially powerful new tool has received significant attention: the prospect of using data science to help lawyers make better outcome predictions. This part of the article discusses the potential for data science to provide lawyers with an additional tool to improve their outcome predictions.

A. Data Science and Prediction

"Data science" and "data analytics" are fairly vague terms, encompassing a number of different techniques analysts use to drive information from large sets of data. As one prominent analyst defines it, "[d]ata science is the transformation of data using mathematics and statistics into valuable insights, decisions, and products."[162] Data science includes traditional analytics techniques such as optimization, forecasting, and simulation, along with more recent innovations such as data mining, artificial intelligence clustering, machine-learning, and detection of outliers.[163]

The use of such tools to make predictions is often referred to as "predictive analytics" or "outcome analysis."[164] Predictive analytics has been successfully employed in a variety of contexts. In the realm of politics, for example, analysts such as Nate Silver have used predictive analytics with some degree of success to anticipate election results.[165] In the area of medicine, predictive analytics has shown promise in predicting disease outbreaks, helping physicians diagnose diseases, and in advancing genomics research.[166] In the area of sports, predictive

162. FOREMAN, *supra* note 136, at xiv.

163. *Id*. Foreman's book provides a good overview of these various predictive techniques and how they work. *See id*. at chs. 4–7.

164. *See generally* ERIC SIEGEL, PREDICTIVE ANALYTICS: THE POWER TO PREDICT WHO WILL CLICK, BUY, LIE, OR DIE 103–220 (2013) (providing a general overview of predictive analytics and its uses).

165. *See generally* NATE SILVER, THE SIGNAL & THE NOISE: WHY SO MANY PREDICTIONS FAIL BUT SOME DON'T xiii–xvii (Penguin paperback ed. 2015).

166. *See* Dan Meisler, *Projects use Big Data to predict diseases, advance genomics analysis*, THE UNIVERSITY RECORD (Jan. 24, 2017), https://record.umich.edu/articles/projects-use-big-data-predict-diseases-advance-genomics-analysis; *see also* W. Nicholson Price II, *Regulating Black-Box Medicine*, 116 MICH. L. REV. 421, 425–31 (2017) (discussing the use of computer-generated medical algorithms to assist physicians in making medical diagnoses).

analytics has been used for gambling purposes to predict the outcome of games and tournaments, as well as by teams to predict (e.g., for purposes of determining how much to spend on a free-agent, or which rookie to draft) the likelihood that a player's career will continue its current trajectory or improve.[167] In the field of meteorology, predictive analytics has been used to improve weather forecasts.[168] And in the business world, predictive analytics has been successfully used for a variety of purposes. Most notably, it is used for marketing and advertising purposes to identify consumers in a targeted manner who might be most likely to purchase particular products.[169] However, there are a host of other business uses for predictive analytics,[170] ranging from consumer fraud detection, to evaluating consumer debt risks, to helping dating services find promising matches, to enabling autonomous cars to operate, to automatically customizing music "stations" for individual listeners,[171] and so on.

The success of predictive analytics over the past decade or so is largely due to advances in the field of artificial intelligence, which have enabled predictive analytics to make more accurate predictions than the traditional forecasting models that were used to facilitate predictions in earlier years. The traditional forecasting models required the researcher

167. *See* SILVER, *supra* note 165, at 74–107.

168. *Id*. at 108–41.

169. *See* SIEGEL, *supra* note 164, at 38–40. Siegel discusses a certain retail chain that wanted to target a marketing campaign to those among its existing customers who are pregnant (e.g., in order to send them ads for baby related products). *Id.* at 38. Under traditional methods of statistical analysis, marketers would first have to specify variables that the marketers believed to have predictive import, focusing for example, on women within a certain age range, who had purchased items such as pregnancy tests and diapers within the past several months. *Id.* at 39. But by using artificially intelligent predictive analytics, the retailer was able to identify previously unknown variables within a sample set of customers known to be pregnant (because they had signed up on a baby register), thereby improving the store's ability to predict which customers were pregnant, and allowing it to target its marketing campaign more efficiently. *Id.*

170. *Id*. at 54–59, 116–18; *see also* Spyros Markridakis, *Forecasting the Impact of Artificial Intelligence Part 3 of 4: The Potential Effects of AI on Businesses, Manufacturing, and Commerce*, FORESIGHT: THE INT'L. J. OF APPLIED FORECASTING, Spring 2018, at 18. Another important use of predictive analytics is business forecasting, whereby analytics are now used to predict such things as a company's future revenues or growth. *See generally* MICHAEL GILLILAND, LEN TASHMAN & UDO SGLAVO, BUSINESS FORECASTING: PRACTICAL PROBLEMS AND SOLUTIONS (2016) (containing a collection of leading articles on business forecasting).

171. For a graphical overview of these different business uses, *see* SIEGEL, *supra* note 164, at 142 (Tables 1–9).

to specify the variables that the researcher believed to be significant for purposes of prediction.[172] In law, for example, a lawyer may think that the court and particular judge involved, the location of the trial, the particular lawyers representing the parties, and, of course, the nature of the cause(s) of action involved are the most important predictive variables. The lawyer can then focus on those variables when comparing the facts of precedent cases to the facts of a prospective case. The more sophisticated tools that employ artificial intelligence advance the analysis further by using algorithms to identify their own predictive variables.[173] Thus, instead of relying just on a researcher's intuition as to what factors have predictive import, some artificially intelligent tools are capable of identifying patterns and automatically isolating predictive variables that the researcher may not have considered. The tools do this by automatically identifying patterns in training sets of data, and then creating predictive models based upon these patterns.[174] In addition, some of the newer tools differ from the earlier, more basic analytics tools in that they employ machine-learning techniques, which means that they are able to learn from their mistakes, and thereby continue to hone over time the accuracy of their predictions.[175] If a particular variable turns out to be a less promising predictor than originally hypothesized, a sophisticated predictive-analytics model will automatically adjust the weighting it gives that variable going forward to improve the accuracy of the model.[176]

Perhaps the most high-profile example in recent years of using artificial intelligence to drive more accurate predictions has been the development of applications based upon IBM's *Watson* platform.[177] To demonstrate the capability of Watson, IBM first used it to develop an artificially intelligent *Jeopardy* contestant, equipping the computer with memory capable of accessing millions of documents very quickly, and then training it with appropriate sample sets to predict the correct

172. *See id*. at 26–27.

173. *Id*. *See generally* Lyra Bennett Moss & Janet Chen, *Using Big Data for Legal and Law Enforcement Decisions: Testing the New Tools*, 37 U.N.S.W. L.J. 643 (2014) (providing an overview of the process by which artificial intelligence uses algorithms to identify predictive variables).

174. SIEGEL, *supra* note 164, at 111–15.

175. *Id*. at 110.

176. *Id*. at 122–23.

177. *See* Katz, s*upra* note 16, at 925–26.

answers to *Jeopardy* questions.[178] At first, Watson was unable to beat a group of Jeopardy champions. However, because it had machine-learning capability, Watson was able to improve its performance as time went on to the point where it was able to beat these champions regularly.[179] Subsequently, IBM has used the Watson platform to enable such tools as an artificially intelligent chess player[180] and an artificially intelligent chef,[181] both of which are able to compete well with masters of their respective crafts. Also, predictive analytics tools are now being used in fields like medicine as well, where, among other things, they can help predict disease patterns and aid doctors in making diagnoses.[182]

As discussed in the following sections, applications based on the Watson platform are also now being used, along with other predictive-analytics tools, to assist the practice of law.

B. Data Science in the Practice of Law

In the practice of law, data science has been assuming an increasingly important role over the past few years. This began in the area of e-discovery, where data science has enabled law firms and corporate legal departments to conduct discovery investigations in a significantly more cost efficient and timely manner, using techniques such as auto classification and predictive coding.[183] But data science techniques have also been used increasingly for other practice related purposes as well, such as: case management, billing, and budgeting; records management and other types of information governance; contracts review and management; selection of outside counsel; and, most pertinent to this article, outcome prediction.[184] In addition, legal

178. *See id.* at 925–28 (providing a summary of how the Watson platform works).

179. *See* SIEGEL, *supra* note 164, at 151–52, 178–84.

180. *See* SILVER, *supra* note 165, at 265–89 (describing the development of IBM's artificially intelligent chess player that is based on the *Watson* platform, and the machine's ultimate victory over the reigning world champion, Gary Kasparov, in 1997).

181. *See* Alexandra Kleeman, *Cooking with Chef Watson, IBM's Artificial Intelligence App*, NEW YORKER (Nov. 28, 2016), https://www.newyorker.com/magazine/2016/11/28/cooking-with-chef-watson-ibms-artificial-intelligence-app.

182. *See, e.g.*, Meisler, *supra* note 166.

183. Stevenson & Wagoner, *supra* note 22, at 1348 (discussing the enormous growth of e-discovery over the past decade).

184. Katz, *supra* note 16, at 928–49 (discussing in detail some of the legal practice use of predictive analytics); Warren A. Agin, *A Simple Guide to Machine Learning*,

research services have employed some important data-science advances to improve the responsiveness of their searches.[185]

The legal profession's demand for data analytics services appears to be growing quickly. For the past several years, an organization called The Coalition of Technology Resources for Lawyers (CTRL) has published an annual survey of the use of data analytics among corporate legal departments in the United States. In the 2015–2016 survey, 93% of practitioners reported that they thought data analytics will become more important and more widespread in the legal profession in the coming ten years, including 31% who predicted that data analytics would be "very important," considered "indispensable," and its use "widespread" within the next ten years.[186] One year later, the 2016–2017 survey revealed that 99% of practitioners now thought that data analytics will be very important, considered indispensable, and its use widespread within the next decade.[187] According to the survey, the principal purposes for which corporate legal departments use data analytics at the present time are for (1) e-discovery (including document culling, early case assessment, and fact-finding), followed by (2) case management (including management of outside counsel, comparing projected spending to actual spending, resource allocation, and budgeting), (3) review and analysis of contracts,[188] and (4) information governance (including facilitating defensible disposition, facilitating compliance with records policies and

BUSINESS LAW TODAY (Feb. 2017), https://www.americanbar.org/publications/blt/2017/02/07_agin.html.

185. *See, e.g.*, Robert Ambrogi, *Bloomberg Law Launches AI Research Tool to Find Key Points of Law*, LAWSITES BLOG (Sept. 26, 2017), https://www.lawsitesblog.com/2017/09/bloomberg-law-launches-ai-research-tool-find-key-points-law.html; *see also* Ashley & Brüninghaus, *supra note* 114, at 310–333 (discussing the use of artificial intelligence in current research platforms and the possibilities for further use).

186. *See* COALITION OF TECHNOLOGY RESOURCES FOR LAWYERS, DATA ANALYTICS IN THE LEGAL COMMUNITY: 2015–2016 TRENDS 3–6 (2015), http://ctrlinitiative.com/wp-content/uploads/2016/01/CTRL-Survey-Data-Analytics-in-the-Legal-Community-2015-2016.pdf.

187. *See* COALITION OF TECHNOLOGY RESOURCES FOR LAWYERS, DATA ANALYTICS IN THE LEGAL COMMUNITY: 2016–2017 TRENDS 2–5 (2017), http://ctrlinitiative.com/wp-content/uploads/2016/01/2017-CTRL-Report-R2.pdf [hereinafter CTRL, 2016–2017 TRENDS].

188. Both Microsoft and Cisco recently announced that they are instigating pilot projects to develop and test artificially intelligent software that will help law firms manage their contracts. *See* Rhys Dishpan, *Microsoft and Cisco Test the Waters with AI Contract Management Pilot Programs*, LEGALTECH NEWS (May 15, 2017, 2:38 PM), https://www.law.com/legaltechnews/almID/1202786204660/.

other requirements, and facilitating data migration).[189] Beyond these uses, some smaller ventures have been exploring other possible uses for predictive analytics in the practice of law. For example, predictive analytics can be used to assist lawyers in the jury selection process.[190] Additionally, in the area of criminal law, predictive analytics now offers researchers a powerful new tool to assess the potential for recidivism among defendants as a routine part of sentencing decisions.[191]

Data science, therefore, will undoubtedly play an increasingly important role in the practice of law in future years.[192] And while outcome prediction has not thus far been at the forefront of data science applications in the law, that appears to be changing. CTRL's 2018 Survey revealed a 43% increase from 2017 in law-firm use of data science for purposes of outcome analysis, and a 175% increase (from 16% to 44% percent of the surveyed firms) in anticipated spending for purposes of outcome analysis in the coming year.[193]

The following section looks at how lawyers are likely to incorporate these predictive analytics tools into their arsenal of traditional predictive tools to facilitate more accurate outcome predictions, and it discusses some challenges predictive analytics will have to overcome for it to be a true game changer.

C. The Prospects for Using Predictive Analytics to Improve Outcome Predictions

An increasing number of legal commentators have begun to look at predictive analytics as a potentially powerful new tool in the area of outcome prediction.[194] In fact, there is good reason to believe that

189. *See* CTRL, 2016–2017 TRENDS, *supra* note 187, at 1–5.

190. *See* Leslie A. Gordon, *Big Data Juries*, A.B.A. J., Sept. 2016, at 16. For an overview of other law-practice applications of predictive analytics, see Katz, *supra* note 16, at 929–36.

191. *See* Richard Berk, *Machine Learning Forecasts of Risk to Inform Sentencing Decision*, 27 FED. SENT'G R. 222 (2015).

192. Katz, *supra note* 16, at 963–64.

193. *See* COALITION OF TECHNOLOGY RESOURCES FOR LAWYERS, 2018 ANALYTICS REPORT 4–5 (2018), http://ctrlinitiative.com/wp-content/uploads/2018/02/2018-CTRL-IGI-Analytics-Report-Final.pdf.

194. *See* Katz, *supra* note 16, at 948–49; Stevenson & Wagoner, *supra* note 22, at 3 ("[B]y leveraging the quantitative strength of computers, lawyers can accurately forecast how events are likely to play out in the litigation."); Josh Blackman, *The Path of Big Data and the Law*, *in* BIG DATA AND THE LAW (West Academic Press 2014); Snyder, *supra* note 33, at 854–66.

predictive analytics may well drive some significant changes in the way lawyers assess potential case outcomes in their day-to-day practices. Thus, in the not-too-distant future, we can expect to see lawyers relying heavily on predictive analytics to complement the traditional tools of prediction, such as the element-focused analysis.

This Section looks at the current state of predictive analytics in the legal profession, and it assesses the potential going forward for predictive analytics to supplant, or more plausibly, complement, the traditional tools of outcome prediction discussed above in Part III. First, Section V.C.1 looks at the historical development of predictive analytics as a tool for making outcome predictions. Next, Section V.C.2 looks at the current state of predictive analytics as a tool for assessing outcome predictions in the practice of law. Finally, Section V.C.3 examines some key challenges predictive analytics will have to overcome going forward if it is to have a significant effect on the way lawyers make outcome predictions.

1. The Development of Predictive Analytics as a Tool for Outcome Prediction

Fundamentally, predictive analytics is an extension of the use of empirical information, which is one of the traditional tools of outcome prediction.[195] Like that traditional tool, it helps lawyers predict case outcomes by comparing information about past cases with a prospective case. Underlying both tools is an assumption (grounded in the concept of *stare decisis*) that similar cases are likely to be decided similarly.[196] In contrast, whereas the use of jury verdict reporters and settlement data rely on the lawyer's subjective assessment of similarity, predictive analytics employs computer algorithms to detect objective patterns in the language of court opinions and other court documents that can then be compared to the prospective case.[197]

For more than half a century now, researchers have been exploring the potential use of such computational analyses to predict the outcome of legal cases. Most of the early efforts, however, were made in cognate fields, such as political science and artificial intelligence.[198] And for the

195. *See supra* Section III.C.
196. *See* Smith, *supra* note 42, at 15–16, 55.
197. *See* Stevenson & Wagoner, *supra* note 22, at 16–18.
198. *See, e.g.*, Loevinger, *supra* note 10; Kort, *supra* note 10; Schubert, *supra* note 10; Fisher, *supra* note 10.

most part, these initiatives were not disseminated within the legal academy. However, in 1964, a young political science professor named Stuart Nagel published an article in the Texas Law Review entitled *Applying Correlation Analysis to Case Prediction*.[199] This article expanded on an article Nagel had written four years earlier, entitled *Using Simple Calculations to Predict Judicial Decisions*, which was published in *The American Behavioral Scientist*.[200] In these articles, Nagel used reapportionment cases to demonstrate how "correlation analysis" can be used to identify patterns in cases where the party attacking apportionment is successful. Nagel then described the process for conducting this analysis as follows:

> This process can be partially mechanized by converting the full text of the relevant cases into punched tape either by a typist or an optical scanner. Which side won in each case as well as the full text should be punched on the tape. The punch tape can then be processed by a program computer to read out each word (including its grammatical variations and synonyms) that has a +20 correlation or more (at a given level of probability) with victory for a given side (e.g., the apportionment attacker). If too few or too many predictive words are read out, the specified correlational probability levels can be raised or lowered accordingly. The resulting list of predictive words should generate insights as to what some of the relevant predictive variables are.[201]

Nagel thus set out over 50 years ago, in in a rudimentary form, the basic strategy for using predictive analytics to identify patterns in the language of case law that can be used to predict case outcomes.

Over the next several decades, data scientists, political scientists, and researchers in the area of artificial intelligence continued to work on refining techniques for using computational analyses to predict case outcomes (particularly Supreme Court decisions), with limited

199. Stuart Nagel, *Applying Correlation Analysis to Case Prediction*, 42 TEX. L. REV. 1006 (1964) [hereinafter Nagel, *Correlation Analysis*].

200. Stuart Nagel, *Using Simple Calculations to Predict Judicial Decisions*, AM. BEHAV. SCIENTIST, Dec. 1960, at 24. The most recent formulation of Nagel's theory can be found in his book, COMPUTER-AIDED JUDICIAL ANALYSIS: PREDICTING, PRESCRIBING, AND ADMINISTERING (1992).

201. Nagel, *Correlation Analysis*, *supra* note 199, at 1009.

success.[202] But the efforts never really generated much interest among legal academics until personal computers started becoming ubiquitous in the practice of law during the 1990s.

In 1998, Professor Frederick Schauer wrote an article entitled *Prediction and Particularity*[203] that laid an important theoretical foundation for predictive analytics. In that article, Schauer discussed the role of a legal doctrine in enabling outcome predictions, by contrasting the views of Oliver Wendell Holmes and Carl Llewellyn. Under Holmes's view, Schauer argued, a lawyer predicts case outcomes by evaluating how courts resolved precedents by reference to traditional legal concepts such as "contract," "consideration," "waiver," among others. In so doing, the lawyer determines precisely how courts apply these legal concepts and compares their applicability to a prospective case in order to predict how the prospective case will likely be resolved.[204] Llewellyn, on the other hand, put a greater emphasis on non-doctrinal factors in analyzing the likely outcome of cases. As Professor Schauer described Llewellyn's view:

> Llewellyn did not deny that there were regularities in the law. Nor did he deny those regularities might facilitate the process of predicting future legal outcomes. He did, however, deny that those regularities were regularly captured by the generalizations typically referred to as "legal doctrine," and thus claimed that legal doctrine did not reflect empirical regularities, and that legal regularities reflected by categorizations that did not resemble traditional legal doctrine.[205]

Thus, for example, in analyzing injunctions decided by the West Virginia Supreme Court of Appeals from 1920 to 1954, Holmes would look to traditional rules such as "a party who delays claiming its rights to the detrimental reliance of another party is precluded from obtaining an injunction" in order to predict the likely outcome of a prospective case.[206] Llewellyn, on the other hand, would rely on a non-doctrinal principle

202. *See generally* T.W. Rutger et al., *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decision-making*, 104 COLUM. L. REV. 1150 (2004).

203. Frederick Schauer, *Prediction and Particularity*, 78 B.U. L. REV. 773 (1998) [hereinafter Schauer, *Prediction and Particularity*].

204. *Id*. at 781.

205. *Id*. at 782.

206. *Id*. at 783.

such as "the coal company wins" to predict the outcome of future cases.[207]

Professor Schauer went on to note that, while legal scholars had largely ignored Llewellyn's call to focus on extralegal considerations in making outcome predictions, social scientists had been quite active in this regard.[208] He cited as an example a large body of political science scholarship that has analyzed Supreme Court decisions with respect to extra-legal variables that enable outcome predictions.[209] Schauer concluded that while there is slim evidence for the view that traditional doctrinal analysis enables accurate outcome predictions, "there is great empirical support for what [social scientists] call the 'attitudinal model,' the view that the best predictors of Supreme Court decisions are the policy attitudes or preferences of the justices, and that, often, the best predictors of those are the party affiliations of the presidents who appointed them."[210] Schauer's analysis is important, therefore, because it emphasizes the importance, for purposes of outcome prediction, of looking for meaningful patterns among precedents that go beyond the traditional doctrinal concepts the courts purport to rely on in those precedents, which is a task for which predictive analytics is well suited.[211] For even if such factors are deemed to be inappropriate for some purposes, such as legal explanations and arguments, their predictive value for purposes of outcome prediction should not be disregarded.

2. The Current Status of Predictive Analytics as a Tool for Outcome Prediction

Over the past several years, legal scholars have begun to take an increasing interest in the topic of prediction in the law, particularly the prospects for using data science to enable more accurate outcome predictions. Among the more prominent voices in the field at present

207. *See id*. at 783–84.

208. *Id*. at 784–85.

209. *Id*. at 784 n.31; *see also* Katz, *supra* note 16, at 936–39 (discussing social science research beginning in the 1980s that focuses on using non-doctrinal considerations to inform outcome prediction in the context of Supreme Court cases); Stevenson & Wagoner, *supra* note 22, at 1352 n.62 (summarizing research on data-driven attempts to predict Supreme Court decisions).

210. *See* Schauer, *Prediction and Particularity supra* note 203, at 784–85.

211. *See* Bennett & Chen, *supra* note 173, at 647–650 (describing how predictive analytics goes beyond traditional legal concepts in employing predictive variables).

time is Professor Daniel Katz, whose important paper, *Quantitative Legal Prediction – or – How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*,[212] provides a summary of the current uses of data analytics in the practice of law, including outcome prediction. With respect to outcome prediction, the piece examines the early efforts to employ machine learning to enable outcome predictions, particularly in the fields of patent law and securities fraud class actions. Katz stresses the preliminary nature of these efforts and argues that predictive analytics will soon be employed widely to assist lawyers in making outcome predictions.[213] He concludes that "the age of quantitative legal predictions is about a mixture of humans or machines working together to outperform either working in isolation."[214]

The study of predictive analytics to assess potential case outcomes has also led to the development of products that can be used to aid outcome prediction. The principal commercial online research services began offering tools for evaluating potential case outcomes several years ago. These tools essentially aggregated data from jury verdict reports and related publications, allowing users to filter results by category such as type of case, lawyer, judge, and location to find comparable cases. LexisNexis, for example, introduced a product called *Verdict & Settlement Analyzer*,[215] and Westlaw introduced a product called *Case Evaluator*.[216] Both provide information about verdicts (and some settlements) for particular causes of action in various jurisdictions, showing averages and ranges of recovery for a variety of different types of cases. They break the information down by jurisdiction and court, by party (plaintiff versus defendant), by subject matter, and by amount of verdict, and they provide case summaries, along with trial and appellate documents that were filed in the actions. These tools have brought a greater degree of precision to the process of comparing and analyzing

212. Daniel Martin Katz, *Quantitative Legal Prediction – Or – How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 EMORY L.J. 909 (2013).

213. *Id*. at 936–942.

214. *Id*. at 929.

215. *See LexisNexis Verdict & Settlement Analyzer*, LEXISNEXIS, https://www.lexisnexis.com/en-us/products/verdict-and-settlement-analyzer.page (last visited July 30, 2018).

216. *See Westlaw Case Evaluator*, THOMSON REUTERS LEGAL SOLUTIONS, https://tmsnrt.rs/2LQzDN8 (last visited July 30, 2018).

past cases that are similar (per various metrics that the user can specify) to the case in question. Instead of just perusing summaries of cases in jury verdict reporters and looking for similarities to the case in question, the lawyer can rely on data automatically compiled from such cases by the search engine.[217]

In addition, somewhat more sophisticated tools that incorporate artificial intelligence and machine-learning techniques have been introduced in recent years by smaller commercial ventures (mostly incubated by universities) to assist lawyers in evaluating cases. The pioneer in this area of predictive analytics has been a company called Lex Machina, which is now owned by LexisNexis.[218] Based in Silicon Valley, Lex Machina began as a public interest project at Stanford University, and was a spin-off from the law school and the Computer Science Department.[219] The company focused on two areas of law: patent litigation (its original focus) and securities litigation, though it now provides some analytic services with respect to other types of cases as well, such as antitrust, employment, and products liability cases.[220] The company uses predictive analytics tools to provide insights on opposing lawyers, law firms, parties, judges, venues, and other information, and it offers individualized early case assessment.[221] On the patent litigation side, it has compiled a huge database of information from the Electronic Document Information System of the United States International Trade Commission, as well as from the United States Patent and Trademark Office, including documents from trials before the Patent Trial and Appeal Board. It then supplements this information with trial-court documents from the federal government's PACER service.[222] On the

217. Some jury verdict reporters now offer these tools as well, as noted above. *See supra* note 85 and accompanying text.

218. *See* Casey Sullivan, *LexisNexis Acquires Lex Machina*, BIG LAW BUSINESS (Nov. 23, 2015), https://bit.ly/2KyJBSB.

219. *See About Lex Machina*, LEX MACHINA, https://bit.ly/2IEPSXT (last visited Aug. 1, 2018).

220. *See About LexisNexis*, LEXISNEXIS, https://bit.ly/2MGtgc2 (last visited Aug. 1, 2018).

221. *See Legal Analytics Platform*, LEX MACHINA, https://lexmachina.com/legal-analytics/ (last visited Aug. 1, 2018); *Legal Analytics Apps*, LEX MACHINA, https://lexmachina.com/legal-analytics-apps/ (last visited Aug. 1, 2018).

222. *See How It Works*, LEX MACHINA, https://bit.ly/2yYNPOJ (last visited Aug. 1, 2018).

securities side, the company analyzes data regarding damages from reports of SEC penalties, discouragements, and approved settlements.[223]

With respect to outcome prediction, Lex Machina's *Case Resolution Analytics* tool tracks different variables and case outcomes to enable more accurate predictions. Other tools provide information on trends in holdings among judges and courts, evaluations of opposing counsel parties, information on a party's litigation history, and damages analytics. A fairly new product is the company's *Motion Kickstarter* tool, which assists lawyers in drafting motions by identifying which arguments and motion styles are likely to be the most successful in a certain type of case, taking into account the particular court or judge.

While Lex Machina was the first significant venture to make predictive analytics available to lawyers, it has now been joined by several other significant ventures. These include Bloomberg Law, which introduced its *Litigation Analytics* tool in 2016,[224] Judicial Perspectives (which is owned by ALM),[225] Premonition (which is based on IBM's Watson platform and focuses on lawyer selection analytics),[226] and Ravel Law.[227] Of these, Ravel Law appears to be making the biggest splash. Ravel Law, like Lex Machina, is a 2012 spinoff from Stanford University's law, computer science, and "d.school" departments.[228] Also like Lex Machina, Ravel Law is now owned by LexisNexis and available as part of Lexis's subscription package.[229] The program is best known for its innovative legal research platform, which uses visualization tools to

223. Another venture that provides analytics services in the area of securities litigation is NERA Economic Consulting. *See* NERA, http://www.nera.com (last visited Aug. 1, 2018).

224. *See Litigation Analytics*, BLOOMBERG LAW, https://bit.ly/2KkXjsO (last visited Aug. 1, 2018); *see also* Robert Ambrogi, *Bloomberg Law's New Litigation Analytics Peeks Under the Robes of Judicial Data*, LAWSITES BLOG (Oct. 19, 2016), https://bit.ly/2MCTMTl (discussing how Bloomberg Law's *Litigation Analytics* tool "aims to help attorneys gain insights into questions such as how long federal judges typically take to resolve cases, how they rule on dispositive motions, and how often they are overturned on appeal.").

225. *See Legal Solutions to Build a Better Case*, ALM INTELLIGENCE, https://bit.ly/2yXmtsr (last visited Aug. 1, 2018).

226. *See Legal Analytics*, PREMONITION, https://bit.ly/2z0jMX0 (last visited Aug. 1, 2018).

227. *See Our Products*, RAVEL LAW, http://ravellaw.com/products/ (last visited Aug. 1, 2018).

228. *See Who We Are*, RAVEL LAW, http://ravellaw.com/who-we-are/ (last visited Aug. 1, 2018).

229. *See* Stephen Rynkiewicz, *LexisNexis Acquires Case Analytics Firm Ravel Law*, A.B.A. J. DAILY NEWS (June 8, 2017, 2:49 PM), https://bit.ly/2KwvWvc.

show users at a glance the complex relationships between the various precedents interpreting a particular rule.[230] But Ravel Law has also moved in the past two years into the business of predictive analytics. Its *Court Analytics* and *Judge Analytics* tools analyze for a particular judge or court, or by jurisdiction, case outcomes, language patterns, and citation history to provide insight on past rulings and to shed light on anticipated future case outcomes.[231] The tools are thus valuable both for outcome predictions, as well for crafting persuasive arguments. In addition, Ravel Law has introduced a tool called *Firm Analytics* that tracks, for various legal specialties, the success rates and volume of work of various large law firms in order to assist consumers of legal services in choosing the best firm for a particular job.[232] LexisNexis is currently in the process of rolling many of these tools into its Lexis Advance platform.[233]

Another venture that has drawn significant attention is ROSS Intelligence. Incubated at the University of Toronto, ROSS Intelligence is now located in Silicon Valley.[234] The company is primarily associated with legal research; it is based on IBM's Watson platform and relies upon artificial intelligence and machine learning tools. [235] ROSS Intelligence touts its product as the "world's first digital lawyer" because it allows lawyers using it to ask natural language questions, to which it provides answers by predicting the most applicable solution to the problem posed by the question.[236] Earlier this year, ROSS also

230. Ravel Law has also drawn significant attention for its *Case Law Access Project*, which has digitized and provided open access to all the case law in the Harvard Law School library. *See Case Law Access Project*, LIBRARY INNOVATION LAB, https://lil.law.harvard.edu/projects/caselaw-access-project/ (last visited Aug. 1, 2018).

231. *See Our Products*, RAVEL LAW, http://ravellaw.com/products/ (last visited Aug. 1, 2018).

232. *See* Daniel Lewis, *Introducing Firm Analytics*, RAVEL LAW (May 23, 2017), http://ravellaw.com/introducing-firm-analytics/.

233. *See* Robert Ambrogi, *Exclusive First Look: Ravel Law's Integration with Lexis Advance*, LAWSITES BLOG (Feb. 2, 2018), https://bit.ly/2z0m4FB.

234. *See AL Interview: Andrew Arruda, CEO and Co-Founder, ROSS Intelligence*, ARTIFICIAL LAWYER (Oct. 12, 2016), https://bit.ly/2ILJttQ.

235. *See id*. Westlaw has recently introduced a similar tool, called *Westlaw Answers*, as has LexisNexis, which recently introduced *Lexis Answers*. Both provide ordinary language answers to common legal inquiries. *See Westlaw Recent Enhancements*, THOMSON REUTERS LEGAL SOLUTIONS, https://tmsnrt.rs/2z2FMR8 (last visited Aug. 1, 2018); *You Ask. Lexis Answers—new machine-learning feature on Lexis Advance*, LEXISNEXIS, https://bit.ly/2v7co7C (last visited Aug. 1, 2018).

236. *See ROSS Intelligence: Overview*, LINKEDIN, https://www.linkedin.com/company/ross-inc/ (last visited Aug. 1, 2018); *see also* KEVIN

introduced a free product called *EVA* that analyzes drafts of legal briefs to ensure that the citations are still good law and to locate additional cases with similar language to those in the brief.[237] While ROSS currently does not provide a tool for making outcome predictions (given that it is not equipped to handle questions such as, "[w]hat is the likelihood that my client wins this case, given the following facts?"), it does not seem far-fetched to think that the product could eventually be enabled to make such predictions.[238]

In sum, while the available predictive-analytics technology is not yet at a state where it can enable accurate outcome predictions over a broad variety of case types, the landscape is evolving quickly, and it would seem reasonable, based upon the trajectory of growth in the field, to envision that these tools will be widely used by practitioners to supplement the traditional tools of outcome prediction in the not-too-distant future. For this vision to be realized, however, data scientists will need to overcome the obstacles discussed in Section V.C.3 below.

3. Potential Limitations on the Use of Predictive Analytics as a Tool for Outcome Prediction

While the future of predictive analytics in the legal profession looks bright, it is important to keep in mind some of the potential limitations on its ability to improve outcome predictions. While the technological advances hold significant potential, it is easy to get caught up in the enthusiasm and lose track of their limitations. As Nate Silver emphasized in his book on predictive analytics, *The Signal and the Noise—Why So Many Predictions Fail, But Some Don't*, "[I]f science and technology are the heroes of this book, there is risk in the age of Big Data about becoming too starry-eyed about what they might accomplish."[239] So while predictive analytics has some advantages over the traditional tools of outcome prediction, particularly the element-focused analysis discussed in Parts III and IV, it is premature to say that it will replace

D. ASHLEY, ARTIFICIAL INTELLIGENCE AND LEGAL ANALYTICS 14–18 (explaining generally how Watson works).

237. *See* Robert Ambrogi, *ROSS Unveils EVA, a Free AI Tool to Analyze Briefs, Check Cites and Find Similar Cases*, LAWSITES BLOG (Jan. 29, 2018), https://bit.ly/2EmP1L2.

238. *But see* ASHLEY, *supra* note 236, at 18–31 (discussing the challenges Watson faces in terms of its ability to engage in "legal reasoning" and solve legal problems).

239. SILVER, *supra* note 165, at 447.

these traditional tools in the near future; rather, predictive analytics can be expected to complement the traditional tools of outcome prediction.

On the plus side, predictive analytics is not subject to some of the problems with the traditional tools discussed in Part IV above. The difficulty in assessing probabilities that afflict the element-focused analysis, for example, is not an issue because predictive analytics does not rely on an analysis of independent variables,[240] relying instead on the detection of subtle correlations to enable predictions. Likewise, predictive analytics is better able to account for extra-legal considerations than the traditional tools because it can look for patterns among the holdings of individual judges, courts, and party types based on factors the courts may not have enunciated in the opinions (e.g., it can quantify the percentage of time the coal company actually does win in injunction cases).[241] And the data set on which the analysis is based could be significantly broader than the limited collection of precedents a lawyer uses in an element-focused analysis because predictive analytics could take into account trial documents and other information about a case beyond mere published opinions.[242] Nevertheless, as discussed further below, the availability of meaningful data is not as comprehensive as one would hope.

According to Nate Silver, a lack of meaningful data is one of the two principal factors that limits the success of predictive analytics generally.[243] The other is the difficulty in separating what he calls the "noise" from the "signal." As Silver put it:

> The goal of any predictive model is to capture *as much signal as possible and as little noise as possible*. Striking the right balance is not always so easy, and our ability to do so will be dictated by the strength of the theory and the quality and quantity of the data.[244]

240. *See supra* notes 132–142 and accompanying text.

241. *See* Ashley & Brüninghaus, *supra* note 114, at 317–318 (discussing research showing that predictive analytics can make accurate predictions based on rules that do not correspond to patterns of reasoning that are familiar to lawyers).

242. *See* Stevenson & Wagoner, *supra* note 22, at 1354–1368 (discussing sources of potentially useful data); *see also* Kevin W. Clement & Theodore Eisenberg, *Litigation Realities*, 88 CORNELL L. REV. 119, 125–126 (2002) ("On the one hand, judicial decisions represent only the very tip of the mass of grievances. . . . On the other hand, *published* decisions are a skewed sample of that tip . . . .") (emphasis added).

243. *See* SILVER, *supra* note 165, at 80.

244. *Id.* at 388.

This difficulty in teasing out the signal from the noise is the central metaphor that runs throughout Silver's book. What he seems to be referring to is the unique feature of predictive analytics that is both a strength and a weakness. Because it does not try to tease out causal factors for purposes of explanation, but merely looks to find predictive patterns, predictive analytics is able to identify a broader array of meaningful (for purposes of prediction but not necessarily explanation) correlations than traditional scientific methods.[245]

To use a simple example, predictive analytics may reveal that persons living in a certain geographic area tend to be afflicted with lung cancer at a higher rate than average. But this correlation does not show that living in the locale is itself the *cause* of lung cancer—it may just be that persons living in that locality smoke at a higher rate than average. But for purposes of prediction (e.g., for healthcare planning purposes), that distinction doesn't really matter.

This strength, however, can also be a weakness. Because it does not deal in the realm of causation, predictive analytics is subject to identifying accidental correlations that are not meaningful and do not inform predictions.[246] The key challenge for predictive analytics is thus to find ways to eliminate statistical anomalies (i.e., randomness) that do not enable accurate predictions, and in fact, impede them. As Silver put it:

> It would be nice if we could just plug data into a statistical model, crunch the numbers, and take for granted that it was a good representation of the real world. Under some conditions, especially in a data-rich fields like baseball, that assumption is fairly close to being

245. SIEGEL, *supra* note 164, at 90 ("When applying PA, we usually don't know about causation, and we often don't necessarily care. For many PA projects, the objective is more to predict than it is to understand the world and figure out what makes it tick."); *see also* Katz, *supra* note 16, at 952. Of course, the use of predictive analytics is not limited to finding these types of "black box" correlations. Predictive analytics can also be used to help lawyers craft effective arguments by identifying legal arguments, phrases, or cases that have proven to be particularly persuasive, either in general, or with respect to a particular judge. Ravel Law, for example, touts its *Court Analytic*s and *Judge Analytics* tools for this purpose. *See Our Products*, RAVEL LAW, http://ravellaw.com/products/ (last visited Aug. 1, 2018).

246. *See* SIEGEL, *supra* note 164, at 121 ("For any predictive model a pressing question persists: Has it learned something true that holds in general, or only discovered patterns that hold within this data set?").

> correct. In many other cases, a failure to think carefully about causality will lead us up blind alleys.[247]

Thus, predictive analytics is still a work in progress, and significant work remains to be done to improve its ability to distinguish meaningful (i.e., predictive) patterns from non-meaningful patterns.[248]

Furthermore, the use of predictive analytics to inform outcome predictions in the law also suffers from the other significant challenge that Silver referenced (i.e., a lack of quality data). The world of judicial decision-making is not the world of baseball, where "pretty much everything that has happened on a major-league playing field in the past 140 years has been dutifully and accurately recorded, and hundreds of players play in the big leagues every year."[249] Rather, as discussed above in Section IV.A, there are some significant limitations on the types of information available with respect to the actual bases for judge and jury decisions. This is perhaps particularly apparent with respect to settlement information, where the lack of widely available data concerning confidential settlements poses a significant challenge for lawyers using predictive analytics to inform outcome predictions. The problem is that most litigation matters are resolved through settlement, and most of the litigation matters that are resolved through settlement rely on confidentiality clauses to limit public access to the terms of the settlement.[250] But unless data analytics companies can tap into such information, their predictions will not be highly accurate, given the limited data sets that they are able to draw from with respect to case outcomes. To be sure, there are some types of litigation where settlement data is more widely available, such as securities class-action litigation, in which many settlements require court approval, and are therefore in the public domain. Thus, it is not surprising that companies like Lex Machina have so far focused on specialty areas, such as securities litigation and intellectual property litigation. However, this is the exception, not the rule; for most litigation matters, information about

247. SILVER, *supra* note 165, at 372.

248. For a detailed discussion of the various techniques AI researchers have employed in their effort to improve outcome predictions, see ASHLEY, *supra* note 236, at 107–126.

249. SILVER, *supra* note 165, at 80.

250. *See supra* note 82 and accompanying text.

settlement amounts remains largely outside the reach of lawyers not privy to the case.[251]

Of course, much of this data is out there; the challenge is to make it readily available to lawyers for purposes of outcome prediction. Currently, insurance companies and other corporations that are involved in frequent litigation compile settlement data for their own use, and this gives them a significant advantage in predicting how litigation matters are likely to be resolved. But most lawyers do not have access to this type of information, and they are thus left to draw primarily on the traditional outcome prediction tools discussed in Parts III and IV above. If, going forward, insurers and other companies with large quantities of settlement data would agree to pool such data and make it publicly available (perhaps for a fee), that would go a long way toward enabling predictive analytics to provide reasonable assessments of litigation prospects in prospective cases. However, there is no indication at present to think that such companies plan to do so.[252]

A further limitation on the quality of data that predictive analytics relies on is its generic nature, which makes it difficult to track individual factual distinctions between cases. Just evaluating the data from published opinions does not provide a lot of information about the facts that may have weighed on a court's decision in a precedent beyond what the court specifically identifies as legally relevant.[253] Without reference to the entire factual record in a case, predictive analytics will be limited in its ability to find meaningful factual similarities between past cases and a prospective case, thus limiting its predictive potential. To be sure, including pleadings and other trial documents that contain factual information in the data set can help significantly in this respect, which is presumably why a company such as Lex Machina includes trial-level

251. *See supra* note 82 and accompanying text.

252. An argument can be made that insurers and other companies with large amounts of compiled settlement data would be doing everyone a favor, including themselves, were they to make their settlement information more widely available. As discussed in the previous section, there is a tendency for lawyers to overestimate the likelihood of their being successful in a given action that can skew accurate outcome predictions. If these lawyers were privy to more actual data about the types of cases they were involved in, they might be able to temper their expectations, which would lead to earlier and more efficient settlements. *See* Stevenson & Wagoner, *supra* note 22, at 1374–1377 (discussing research that indicates parties are more likely to settle if each side has better information about its opponent's case).

253. For a helpful overview of the prospects for using predictive analytics to assess legal relevance, see Katz, *supra* note 16, at 954–57.

documents (such as those available on the PACER database) in its analysis.[254] But even with such trial documents, factual information is limited because some of the documents that would be the most data-rich, such as deposition transcripts, are generally not publicly available. This lack of factual information is particularly acute with respect to non-doctrinal considerations that may affect the outcome of a case. The likability and credibility of the individual parties, for example, are widely recognized as factors affecting the outcome of trials.[255] Yet this not the kind of information that generally gets compiled in court documents, particularly with respect to jury trials.

In sum, the use of data analytics to predict legal outcomes is not likely to be a cure-all for the problems associated with the traditional predictive tools. But it may well provide a useful supplemental tool in the not-too-distant future to augment the type of predictive analysis undertaken in the traditional legal memorandum. And as Professor Katz argues,[256] predictive analytics tools used in conjunction with the traditional predictive tools will likely outperform either type of tool used individually.

VI. CONCLUSION

Outcome prediction has always been a vital part of practicing law. Clients of all types rely on their attorneys to provide accurate assessments of the potential legal consequences the clients face when making important decisions. And yet, notwithstanding its enormous importance to the practice of law, outcome prediction in the law remains a very imprecise endeavor. The three traditional tools lawyers rely on when making predictions, the element-focused analysis, lawyerly experience, and empirical information, are all subject to significant limitations that hinder their effectiveness as predictive tools.

Fortunately, however, recent advances in data science are enabling new predictive tools that look to be potential game-changers. Already, these advances are bringing about significant changes in the way lawyers practice law, and they hold significant promise for outcome prediction as well. Thus, it seems quite likely that predictive analytics, while not a

254. For helpful discussion of PACER as a source of litigation data at the trial level, see Stevenson & Wagoner, *supra* note 22, at 1357–1364.

255. MAUET & MARCUS, *supra* note 17, at 405–406.

256. Katz, *supra* note 16, at 929 ("The equation is simple: Humans + Machines > Humans or Machines.").

panacea that can replace the traditional tools of outcome prediction in the foreseeable future, will increasingly emerge as an important supplemental tool that should help to make outcome predictions more accurate. And that is very good news for the clients who rely on the predictions.