Boston College Law Review

Volume 62 | Issue 7 Article 4

10-26-2021

The Reasonableness Machine

Brian Sheppard
Seton Hall University School of Law, brian.sheppard@shu.edu

THE REASONABLENESS MACHINE

BRIAN SHEPPARD

THE REASONABLENESS MACHINE

BRIAN SHEPPARD*

Abstract: Automation might someday allow for the inexpensive creation of highly contextualized and effective laws. If that ever comes to pass, however, it will not be on a blank slate. Proponents will face the question of how to computerize bedrock aspects of our existing law, some of which are legal standards—norms that use evaluative, even moral, criteria. Conventional wisdom says that standards are difficult to translate into computer code because they do not present clear operational mechanisms to follow. If that wisdom holds, one could reasonably doubt that legal automation will ever get off the ground. Conventional wisdom, however, fails to account for the interpretive freedom that standards provide. Their murkiness makes them a fertile ground for the growth of competing explanations of their legal meaning. Some of those readings might be more rule-like than others. Proponents of automation will likely be drawn to those rule-like interpretations, so long as they are compatible enough with existing law. This complex dynamic between computer-friendliness and legal interpretation makes it troublesome for legislators to identify the variable and fixed costs of automation. This Article aims to shed light on this relationship by focusing our attention on a quintessential legal standard at the center of our legal system—the Reasonably Prudent Person Test. Here, I explain how automation proponents might be tempted by fringe, formulaic interpretations of the test, such as Averageness, because they bring comparatively low innovation costs. With time, however, technological advancement will likely drive down innovation costs, and mainstream interpretations, like Conventionalism, could find favor again. Regardless of the interpretation that proponents favor, though, an unavoidable fixed cost looms: by replacing the jurors who apply the test with a machine, they will eliminate a long-valued avenue for participatory and deliberative democracy.

INTRODUCTION

As technologists seek to computerize aspects of our legal apparatus, the notion that legal functions can be automated has transformed from object of imagination to corporate mission statement. Legal technology investment ex-

* Associate Dean and Professor of Law, Seton Hall University School of Law. I am grateful for the feedback of the participants in the (Im)Perfect Enforcement Conference at Yale Law School and the World Congress of the International Association for the Philosophy of Law and Social Philosophy at the University of Lucerne. I am also grateful to Nicholas Almendares, Michael Coenen, Markus Kneer, Levin Güver, Andrew Moshirnia, and Kevin Tobia for their suggestions.

ploded in 2018 to at least one billion dollars in new investments—a seven hundred percent increase from the prior year[1]—and then broke that annual investment record by the end of the third quarter of 2019.[2] Several splashy investments in 2020 show that COVID-19 did little to change the trend.[3] Although investors have funded a variety of legal tech firms, a significant portion of this capital ended up in the hands of companies that aim to automate core aspects of lawyering, such as legal interpretation, legal argumentation, legal writing, and legal advice.[4] Legal scholars, too, have begun to contemplate the possibility of automated "robot" judges or lawyers.[5]

It is far from clear that we will ever replace core legal functions with machines, but a lack of certainty does not mean that we should refrain from studying the possibility until it is clearly upon us.[6] The law is the most important human-constructed means of regulating conduct; it claims authority over those who are subject to it and wields terrifying coercive power. There is something fundamentally risky in delegating features of our legal apparatus to machines, so it stands to reason that we should consider how such delegation might take shape, both in small and large scale, so long as automation is plausible.[7]

---

[1] *See* Valentin Pivovarov, *713% Growth: Legal Tech Set an Investment Record in 2018*, FORBES (Jan. 15, 2019), https://www.forbes.com/sites/valentinpivovarov/2019/01/15/legaltechinvestment2018/ [https://perma.cc/P784-Q4JL].

[2] Meg McEvoy, *Analysis: 2019 Legal Tech Investments Top $1B After Strong Q3*, BLOOMBERG L. (Oct. 11, 2019), https://news.bloomberglaw.com/bloomberg-law-analysis/analysis-2019-legal-tech-investments-top-1b-after-strong-q3/ [https://perma.cc/DCQ6-VJBV].

[3] *See* Bob Ambrogi, *20 for 2020: The Legal Tech Trends That Defined the Year*, LAWSITES (Dec. 30, 2020), https://www.lawsitesblog.com/2020/12/20-for-2020-the-legal-tech-trends-that-defined-the-year.html [https://perma.cc/W5BW-VLQG].

[4] *Id.* (showing contract automation and small claims court appeals among the top ten investments and discussing investments in eDiscovery and automated legal consultation).

[5] *See, e.g.*, Milan Markovic, *Rise of the Robot Lawyers?*, 61 ARIZ. L. REV. 325, 326 (2019); Eugene Volokh, *Chief Justice Robots*, 68 DUKE L.J. 1135, 1137 (2019).

[6] Frank Pasquale & Glyn Cashwell, *Four Futures of Legal Automation*, 63 UCLA L. REV. DISCOURSE 26, 28 (2015), https://heinonline.org/HOL/P?h=hein.journals/ucladis63&i=26 ("Will software substitute for lawyers, or increase their earning power? There will be evidence of each in coming decades: Routine work will continue to be automated, while new opportunities will also emerge. The critical question is which trend will be dominant, and what its effect will be.").

[7] Admittedly, if the important aspects of applying legal doctrine were to be automated, that process would surely take place during a time in which other important dimensions of lawyer or judge services were automated as well. For simplicity's sake, however, the discussion here will assume that the breach determination under the Reasonably Prudent Person Test (RPPT) is the only dimension of the process of establishing negligence in an American trial that will be handled by machines. With respect to other aspects of a negligence case, then, my assumption is that arguments will be raised by human lawyers, factual determinations will be made by human jurors, and legal determinations will be made by human judges. This artificiality will not undo the analysis, I hope; we can draw some confidence from the fact that there are already analogous situations in the United States legal system, after all. For example, California recently passed a law, only to have it later repealed on referendum, that sought to replace its cash bail system with one in which judicial decisions regarding pre-trial detention turn, in part, on flight risk assessments from a "a validated risk assessment tool," which is an algo-

There are automation skeptics, to be sure, but they too benefit from this research. Even if computerization of important legal functions never comes to pass, features of our *present-day* legal system are placed in a new light when we carefully examine how automation might proceed. Even if only as an intuition pump, carefully envisioning automated laws can reveal features of the law or legal system that are unavoidably human, governmental functions for which we overrate or underrate human performance, and underlying social values that need resuscitation. In short, there are benefits to taking automation seriously even if we don't know when or if it will happen.

For those who are nevertheless fixated on calculating the probability of legal automation, I offer a framework. The likelihood that lawyer or judge functions will be displaced by machines is contingent upon at least three forces, which work in combination: (1) the pace and trajectory of technological innovation; (2) the perceived benefits of using technology in that context; and (3) our willingness to bend the law to accept lower performance standards.[8] Frequently, scholars focus on the first two forces, which has given rise to an abundance of futurist research on technology's capacity to meet or exceed the performance standards as presently set by humans[9] and of progressive research on the ways in which technology might improve access to justice in the short-term.[10] The possibility that law's stakeholders will accept different, and more

---

rithm-powered computer program. Act of Aug. 28, 2018, ch. 244, § 4, CAL. PENAL CODE § 1320.7 (2018), *repealed by* Proposition 25 (Cal. 2020); Meagan Flynn, *California Abolishes Money Bail with a Landmark Law. But Some Reformers Think It Creates New Problems.*, WASH. POST (Aug. 29, 2018), https://www.washingtonpost.com/news/morning-mix/wp/2018/08/29/california-abolishes-money-bail-with-a-landmark-law-but-some-reformers-think-it-creates-new-problems [https://perma.cc/4BPB-LVSE]. So, although this analysis proceeds in the form of a thought experiment, it is not one that suffers from lethal defects of implausibility.

[8] *See generally* Brian Sheppard, *Incomplete Innovation and the Premature Disruption of Legal Services*, 2015 MICH. ST. L. REV. 1797 (focusing on the role the third force could play in displacing human-powered legal services).

[9] *See, e.g.*, Daniel Martin Katz, *The MIT School of Law? A Perspective on Legal Education in the 21st Century*, 2014 U. ILL. L. REV. 1431, 1431–33; Markovic, *supra* note 5, at 326; Frank Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, 87 GEO. WASH. L. REV. 1, 2–3 (2019); RICHARD SUSSKIND, TOMORROW'S LAWYERS: AN INTRODUCTION TO YOUR FUTURE (2d ed. 2017).

[10] *See, e.g.*, Raymond H. Brescia et al., *Embracing Disruption: How Technological Change in the Delivery of Legal Services Can Improve Access to Justice*, 78 ALB. L. REV. 553, 553–54 (2014); James E. Cabral et al., *Using Technology to Enhance Access to Justice*, 26 HARV. J.L. & TECH. 241, 243–44 (2012); Sherley E. Cruz, *Coding for Cultural Competency: Expanding Access to Justice with Technology*, 86 TENN. L. REV. 347, 348–49 (2019); Jessica Frank, *A2J Author, Legal Aid Organizations, and Courts: Bridging the Civil Justice Gap Using Document Assembly*, 39 W. NEW ENG. L. REV. 251, 251–52 (2017); Rochelle Klempner, *The Case for Court-Based Document Assembly Programs: A Review of the New York State Court System's "DIY" Forms*, 41 FORDHAM URB. L.J. 1189, 1192– 93 (2014); J.J. Prescott, *Improving Access to Justice in State Courts with Platform Technology*, 70 VAND. L. REV. 1993, 1994–2001 (2017); Kathleen Elliott Vinson & Samantha A. Moppett, *Digital*

technologically feasible, performance of legal functions has gotten less attention,[11] but it has the capacity to play a pivotal role in the process of automation. Increased willingness to change our law's substance and procedure to make automation easier decreases the wait time for adequate technological performance. This, in turn, lowers innovation costs[12] and makes it more likely that the technology will be perceived to provide a net benefit.[13]

Central features of existing legal systems are difficult to change, both because citizens likely perceive them to be of great benefit[14] and because wide-scale revisions present challenging logistical or coordination problems.[15] Indeed, this may well be why extensive automation of law never comes to pass.

Pro Bono: Leveraging Technology to Provide Access to Justice, 92 ST. JOHN'S L. REV. 551, 551–53 (2018).

[11] *See, e.g.*, Sheppard, *supra* note 8, at 1876–96 (considering ways in which ease of automation could change practice of law and legal interpretation). Scholars who have considered changing laws to hasten automation have typically focused upon the ways in which legal ethics rules might be modified to allow legal technology to gain ground in the legal services market. *See, e.g.*, Agnieszka McPeak, *Disruptive Technology and the Ethical Lawyer*, 50 U. TOL. L. REV. 457, 474–75 (2019) ("While some of these barriers are justified by these other purposes of the ABA Model Rules, some rules can be seen as anti-competitive and slow to evolve to the realities of modern legal needs. . . . While lawyers must honor and uphold the law, they can also take on leadership roles in law reform. Although law reform activities should not conflict with duties to clients, it is certainly within the lawyer's role to move the law forward to allow for new innovation—especially lawtech.").

[12] I use "cost" to refer not only to the price of goods or services, but also to development time. *Economic Cost*, CAMBRIDGE BUSINESS ENGLISH DICTIONARY (2011) ("[T]he cost in money, time, and other resources needed in order to do something or make something"). I also consider error or risk of error to be a cost, particularly in the context of reliability.

[13] A simple, albeit implausible, illustration might be helpful: suppose that we were willing to accept that all legal disputes be resolved by the spin of a wheel of fortune. We presently have the technology to create and install a sufficient number of automated wheels to handle the volume of disputes at low cost. Moreover, this cost would be much smaller than the cost associated with supporting human judges, law clerks, lawyers, expert witnesses, and the like. Under those conditions, the perceived benefit of implementing the technology would be quite high, and it would therefore be reasonable to predict that automation of dispute resolution would occur in short order. Thankfully, such a departure from the dictates and inherent values of our legal system is scarcely imaginable, but it shows the impact that willingness to bend the law could have on the value proposition of automation.

[14] *See, e.g.*, *The Public, the Political System and American Democracy*, PEW RSCH. CTR. 40, 46 (Apr. 26, 2018), https://www.pewresearch.org/politics/wp-content/uploads/sites/4/2018/04/4-26-2018-Democracy-release-1.pdf [https://perma.cc/L9R8-35YS] (finding that members of both political parties agree on several critical election issues, even in a polarized political climate). "Republicans and Democrats widely agree on the most important electoral components for the U.S. Nearly nine-in-ten across both parties say it is very important that elections are free from tampering . . . . Comparable majorities in both parties also say it's very important that no eligible voters are prevented from voting (85% of Republicans, 83% of Democrats)," and "[a] large majority of Americans say it's important for there to be a balance of power between the three branches of the federal government." *Id.* (emphasis omitted).

[15] *See, e.g.*, David S. Law, *A Theory of Judicial Power and Judicial Review*, 97 GEO. L.J. 723, 753–54 (2009) ("In order for the people to maintain control over the government, they must not only find a way to monitor the government, but also overcome large-scale coordination problems.").

It is far easier, however, to repeal or add individual laws and perhaps easier still to change the dominant interpretation of an existing legal test. If we assume for the sake of analysis that there is a desire to automate, the malleability of law can be exploited to hasten automation by changing legal tests to make them easier to translate into computer code.

A single article does not permit us to analyze the automation potential of more than a small number of legal functions. Some selectivity is required to illustrate the push and pull of the three forces. The best choice is a legal function that is commonplace (which increases the likelihood that automation would be seen as a favorable way to cut expenses); that is potentially difficult to computerize under existing technology (which carves out a role for the pace and trajectory of innovation); and that has given rise to several competing interpretations (which illuminates the complex relationship between law's malleability and lowering costs of innovation).

The test for negligent breach in American tort law, popularly known as the Reasonably Prudent Person Test (RPPT), meets all of these conditions.[16] It is ubiquitous, potentially tricky to translate into computer code, and has given rise to divergent interpretations, some of which are easier to computerize.

To be sure, focusing on the RPPT means that I will not be able to spend as much time on other aspects of negligence law that could be subject to automation. That is to say: if courts or other stakeholders automated the test for breach, it is likely that other aspects of negligence cases would be automated as well, such as fact-finding or damages calculation.[17] I will not be considering these non-breach dimensions in detail, however. Doing otherwise in an article of this size would risk spreading the analysis too thinly, and the importance of the RPPT cannot be overstated.

The RPPT is the beating heart of negligence liability, serving as the criterion for satisfaction of breach in the many cases involving ordinary duties.[18] Breach is an element of the prima facie case for negligence,[19] and it is typically

---

[16] *See, e.g.*, Cornella v. Churchill Cnty., 377 P.3d 97, 102 (Nev. 2016) (identifying the test for ordinary negligence as the Reasonably Prudent Person Test); *Negligence*, BLACK'S LAW DICTIONARY (11th ed. 2019).

[17] Likewise, I will not have space to discuss in detail important administrative considerations like the manner in which funding would occur, data would be accessed and secured, or builders would be selected.

[18] *See, e.g.*, Christopher Brett Jaeger et al., *Justice Is (Change) Blind: Applying Research on Visual Metacognition in Legal Settings*, 23 PSYCH. PUB. POL'Y & L. 259, 263 (2017) ("Defendants in tort cases are, in theory, liable for negligence only if they breach a duty of care owed to the plaintiff. And, speaking generally, a defendant breaches his or her duty of care only if he or she fails to act as an ordinary, reasonably prudent person would have acted in the circumstances.").

[19] *See* Karni Chagal-Feferkorn, *The Reasonable Algorithm*, 2018 U. ILL. J.L. TECH. & POL'Y 111, 118–21 (discussing negligence and assigning a "'reasonable algorithm' standard" to non-human decision-makers).

a jury's duty to resolve whether a breach has occurred. Ordinary duty cases make up a large share of all negligence cases, which give rise to hundreds of billions of dollars in costs annually.[20]

The RPPT is also a fruitful object of study for legal theorists. It is the paradigmatic legal standard[21]—it is what we think of when we think of a ubiquitous but paltry legal constraint. Although legal interpretation and the processing of computer code are not the same enterprise, the oft-discussed rules/standards distinction in legal theory is a good rule-of-thumb for identifying the challenges to automation of interpretative tasks in law.[22] A bright-line rule, such as the tax provision that imposes upon married, joint filing individuals a ten percent tax rate for taxable income up to $19,050,[23] is far easier to computerize than a standard, such as the statute that requires removal of a legal permanent resident for a crime of moral turpitude.[24] Given the predominance of rules in the U.S. Tax Code,[25] it is hardly surprising that tax preparation software has a track record of success; estimates suggest that more than thirty-four and a half million taxpayers used it in preparing their tax returns in 2019.[26]

Like many standards, the RPPT is meager and murky on its face: it simply asks the jury to decide whether a party or parties acted as a reasonable person would act.[27] The test's reliance upon short, evaluative content ("reasonably

---

[20] According to one study, albeit one from a tort reform advocate, "[l]iability related to automobile accident claims accounted for $160 billion" in 2016. *Costs and Compensation of the U.S. Tort System*, U.S. CHAMBER INST. FOR LEGAL REFORM 4 (Oct. 2018), https://instituteforlegalreform.com/wp-content/uploads/2020/09/Tort_costs_paper_FINAL_WEB.pdf [https://perma.cc/SU49-J442].

[21] *E.g.*, Jack L. Goldsmith, *Book Reviews and Notes*, 91 AM. J. INT'L L. 389, 393 (1997) (reviewing ANDREAS F. LOWENFELD, INTERNATIONAL LITIGATION AND THE QUEST FOR REASONABLENESS: ESSAYS IN PRIVATE INTERNATIONAL LAW (1996)) ("The reasonableness principle is a paradigmatic standard.").

[22] *See, e.g.*, Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 89 HARV. L. REV. 1685, 1687–1713 (1976) (tying the rules/standards distinction to rhetorical modes of individualism and altruism in private law judicial decision-making); Kathleen M. Sullivan, *The Supreme Court, 1991 Term–Foreword: The Justices of Rules and Standards*, 106 HARV. L. REV. 22, 56–122 (1992) (utilizing the rules/standards distinction for assessing Supreme Court opinions).

[23] 26 U.S.C. § 1(j)(2)(A).

[24] 8 U.S.C. § 1227(a)(2)(A)(i).

[25] *But cf.* Alice G. Abreu & Richard K. Greenstein, *The Rule of Law as a Law of Standards: Interpreting the Internal Revenue Code*, 64 DUKE L.J. ONLINE 53, 54–57 (2015), https://heinonline.org/HOL/P?h=hein.journals/duljo64&i=52 (arguing that rule of law values permit stakeholders to treat tax code provisions as standards).

[26] Paul Kiel & Justin Elliott, *TurboTax and Others Charged at Least 14 Million Americans for Tax Prep That Should Have Been Free, Audit Finds*, PROPUBLICA (Feb. 5, 2020), https://www.propublica.org/article/turbotax-and-others-charged-at-least-14-million-americans-for-tax-prep-that-should-have-been-free-audit-finds [https://perma.cc/Q6M2-7EZQ].

[27] *See* Kenneth W. Simons, *The Hand Formula in the Draft Restatement (Third) of Torts: Encompassing Fairness as Well as Efficiency Values*, 54 VAND. L. REV. 901, 930 (2001) ("First, the 'reasonable person' test is obscure. Although it does tell us that the subjective views of the actor are

prudent") appears to direct jurors to engage in a loosely guided normative evaluation of the facts before them. This could very well call for them to make complex, context-based assessments. Linguistically, the RPPT is the hot fudge sundae of indeterminacy: it is semantically vague because it is likely to produce a lot of borderline cases, it is general because a variety of conduct falls under its umbrella, and it is multidimensionally polysemic because it is unclear which features of decision-making that underlie reasonableness must be present for the test to be satisfied and in what amount.[28]

Therefore, it is not obvious that the RPPT can be satisfactorily[29] encoded into machine language. If interpretation of the RPPT can be automated, then perhaps anything in law can be.

But the notion that rules are easy to encode and standards are difficult is simply a rule of thumb. Rules can close off avenues for automation, and standards can provide opportunities. An existing legal rule might set forth conditions for satisfaction that are difficult to detect using machines due to their complexity or obscurity.[30] Replacing these machine-unfriendly rules is potentially costly because of the risk that people will view the change as a repeal or amendment to existing law.[31] Perversely, the standard-like quality of a legal norm could serve as an opportunity to automate under the right circumstances. The openness of a standard might give rise to a range of permissible legal interpretations, some of which are compatible to translation into machine code. If so, the perceived benefits of automation might lead to a new consensus around the computerizable interpretation but without the same risk that people will view adoption of that interpretation as a repeal or amendment. In other words, the openness of a standard is a double-edged sword; it can make translation of its meaning into computer code difficult, but it can also make it easier to offer attractive, translation-friendly conceptions of that meaning.

The RPPT has this character. Its openness appears, at first, to be incompatible with machine coding. But the frequency with which the RPPT has been applied has given rise to a number of alternative, legally plausible interpreta-

---

not decisive of fault, it does not explain which factors are relevant to fault." (emphasis omitted) (footnote omitted)).

[28] Brian G. Slocum, *Replacing the Flawed* Chevron *Standard*, 60 WM. & MARY L. REV. 195, 214–17 (2018) (describing linguistic indeterminacy of words such as "intelligent" and "reasonable").

[29] By saying "satisfactorily," I hope to signal that my analysis here does not set the bar at perfection. Perfect mimicry of human interpretation or perfect performance of ideal interpretation might be an interesting subject to ponder, but it is not a precondition for automation. Imperfect and even inferior machine work can displace human work if it has a lower price/performance ratio.

[30] For example, a rule that conditions fishing of a certain species based on the population of that species in particular waters might suffer from such difficulties.

[31] *See, e.g.*, Frank Cross et al., *A Positive Political Theory of Rules and Standards*, 2012 U. ILL. L. REV. 1, 39 (discussing the difficulty in changing rules); Gregory C. Shaffer & Mark A. Pollack, *Hard Versus Soft Law in International Security*, 52 B.C. L. REV. 1147, 1163 (2011) (same).

tions of the test. This intrinsic malleability could facilitate computerization by allowing the selection of the most computerizable, yet legally plausible, alternative.

Thus, the frequency, importance, and challenge of the RPPT make it the perfect focal point for an analysis of the automation of legal interpretation. This undertaking will require identification and consideration of competing interpretations of the standard, as well as assessment of how the project of automation might falter in practice. In this Article, I undertake this project.

Although my focus is on the future, I am not seeking to predict it. As legal technologist and scholar Daniel Martin Katz warns, "Despite all technical possibilities, prediction is a difficult enterprise, and as such, one should confront the question with humility."[32] I seek only to increase our understanding of the complex relationship between legal constraint, machine interpretation, and cost. It is true that a significant portion of my analysis will discuss how some interpretations of the RPPT are easier to computerize than others. At those points, I will periodically reference the extensive literature aiming to predict whether machines will reach important milestones such as consciousness or feeling. Those references are not endorsements, and I do not wish to make such heady prognostications here. Rather than hypothesizing when or if we will arrive at a destination, I take on the humbler task of describing which routes appear shorter or longer on a map that doesn't show the traffic.

In Part I, I discuss the development of interpretive technology, focusing on the development of Natural Language Processing (NLP) and identifying its strengths and weaknesses with respect to rules and standards under typical conditions of application.[33] In so doing, I identify the parameters by which we can judge the difficulty of computerization. In Part II, I introduce the RPPT, identifying the major competing conceptions of its meaning and then subjecting them to the parameters of computerization.[34] After taking stock of those results, I step outside of the legal meaning of the RPPT in Part III to reveal potentially vital structural aspects of our legal system *as a whole* that could be lost if computerization of RPPT is adopted.[35]

I. RULES, STANDARDS, AND THEIR AUTOMATION

It is fair to wonder whether legislators will have any interest in automating legal interpretation and, if so, whether there are any factors they will con-

[32] 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, 958 (2013).

[33] *See* discussion *infra* Part I.

[34] *See* discussion *infra* Part II.

[35] *See* discussion *infra* Part III.

sider in that endeavor. In this political environment, it is hard to imagine American legislators doing anything of this scale and ambition. If our forecasting window is large enough, however, it becomes fair to predict that the political winds will shift enough that grand legislative initiatives will have a chance of passage. If that much is true, then the economic appeal of an initiative will be a good predictor of the likelihood of a particular initiative's passage. The rules/standards distinction has proven to be a useful prism through which to evaluate resource management in legislation.[36] I will start there.

A. The Cost Centers of Rules and Standards

Rules set forth specific, concrete directives, and it is commonplace that a rule will condition compliance on numerical minimums or maximums.[37] Standards, by contrast, set forth evaluative, discretionary directives.[38] Rules are assumed to be better at constraining those subject to them and therefore are more likely to bring about desired conduct.[39] There is some empirical support for this proposition.[40] Thus, rules can have a big payoff for those with the authority to create them.[41]

A significant problem for legislators is that the constraining force of rules brings the risk that an unwise or poorly written rule could lead to undesired outcomes by channeling conduct in a fruitless or counterproductive direction.[42] This is particularly problematic if correcting the rule would be difficult, which is often the case.[43] Thus, legislative repeal or amendment is costly from both a resource and a political standpoint because it renews the process of engaging in potentially burdensome or expensive rule creation[44] and because it constitutes an admission of error, respectively.[45]

---

[36] *See, e.g.*, Prasad Krishnamurthy, *Rules, Standards, and Complexity in Capital Regulation*, 43 J. LEGAL STUD. S273, S275 (2014) (noting the utility of distinction).

[37] *See* Cass R. Sunstein, *Problems with Rules*, 83 CALIF. L. REV. 953, 961–62 (1995).

[38] *See id.* at 964–65.

[39] *See* Brian Sheppard, *Judging Under Pressure: A Behavioral Examination of the Relationship Between Legal Decisionmaking and Time*, 39 FLA. ST. U. L. REV. 931, 990 (2012) [hereinafter Sheppard, *Judging Under Pressure*].

[40] *Id.*

[41] *See* Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557, 621–22 (1992).

[42] *See* Robert G. Bone, *Who Decides? A Critical Look at Procedural Discretion*, 28 CARDOZO L. REV. 1961, 2002 (2007).

[43] *See id.*

[44] *See id.*

[45] *See* Daniel W. Drezner, *Why Can't World Leaders Ever Admit They Were Wrong?*, WASH. POST (June 9, 2015), https://www.washingtonpost.com/posteverything/wp/2015/06/09/why-cant-world-leaders-ever-admit-they-were-wrong/ [https://perma.cc/3CV7-ANBA] (describing how political incentives make it costly for political leaders to admit error, stating "[doing so] generates zero political upside and risks alienating one's base").

To minimize the risk of undesired outcomes, legislators must take great care in crafting rules, which often will include understanding, in detail, the state of affairs that they want and the constraints upon conduct that are necessary to bring it about. [46] Ordinarily, the domain in which legislators operate is dynamic, multifaceted, or otherwise complex, which makes it harder for them to know how to fashion a rule that appropriately responds to the domain's complexity.[47] Providing a directive that gives subjects the path to follow in every important context is hard and, therefore, costly.[48] In other words, rules are powerful instruments of change, for better or worse. Therefore, discerning the optimal course of conduct and thereafter creating a rule to bring it about are costly endeavors for the rule-maker who is keen to avoid serious error.[49]

Learning the ways in which rule drafting is expensive will help us identify the cost centers of legislation and, thereafter, to make better predictions regarding the desirability of automation. My account is largely minimalist, focusing on the work of creating norms to effectuate a state of affairs.[50]

The lodestar of legislative design should be the state of affairs that the legislator seeks to bring about.[51] With that aim, the primary task is to devise a

---

[46] *See* Gideon Parchomovsky & Alex Stein, *Catalogs*, 115 COLUM. L. REV. 165, 176 (2015) ("Rules require precision. In enacting rules, the legislature must accurately identify the circumstances under which the chosen rule will achieve the desired results. . . . To succeed in this task, the legislature must gather, analyze, and categorize an enormous amount of information—a process requiring the expenditure of substantial resources.").

[47] John C. Roberts, *Gridlock and Senate Rules*, 88 NOTRE DAME L. REV. 2189, 2190–91 (2013) ("[L]egislating is difficult under the best of circumstances . . . .").

[48] *See id.*

[49] Louis Kaplow, *Information and the Aim of Adjudication: Truth or Consequences?*, 67 STAN. L. REV. 1303, 1307–08, 1313 (2015) ("One involves the distinction between rules and standards . . . . [A distinction that] govern[s] the intensity of effort (and thus cost) of supplying legal content both ex ante and ex post. Moreover, they influence the law's consequences for individuals' behavior in the interim because such behavior depends critically on the extent to which individuals choose to become informed about the law before they act. . . . Rules are more expensive up front because more effort is expended then, and standards are more costly in adjudication because substantive legal content must be determined at that time.").

[50] I do not deny that there are other aspects of legislation that are important and potentially costly. Other academics have provided fulsome accounts, particularly in the field of legisprudence – which sets forth rational principles of legislation, such as justifying limits to freedom posed by the legislation under consideration. *See* Luc J. Wintgens, *Legisprudence as a New Theory of Legislation*, 19 RATIO JURIS 1, 10–24 (2006).

[51] It is certainly possible that a legislator has not identified a desired state of affairs or is at a loss to describe its features. They might seek only to appear as though they are engaged in an important legislative project while actually seeking to maintain the status quo; they might score political points without changing the current state of affairs. This is not so implausible. *Cf.* Tony Wright, *What Is It About Politicians?*, 84 POL. Q. 448, 448–53 (2013) (discussing the accuracy of the popular belief that politicians primarily engage in games rather than drafting good legislation for their constituents). In such a situation, they could do any number of things, but it is not obvious that a rule or standard would be the better approach. For instance, a bad faith legislator might create rules of exceedingly narrow scope so as to render them inert or create vague standards that are unlikely to bring about

function that draws a connection between certain conduct and the desired state of affairs. This is like asking, "What kind of behavior would bring about the state of affairs we want?" Once identified, the work of drafting the legal norm is largely a matter of devising a test that makes legal compliance contingent upon the desired conduct and an appropriate penalty for non-compliance. It sounds simple, but it can take hard work.[52] First, it might be challenging to identify the various forms of conduct that should be altered by the norm. Even with knowledge of the scope of conduct to be regulated, it might be difficult to determine a sufficiently effective way to alter that conduct. And even then, it might be difficult to articulate the conduct alteration function in the form of a norm that can be followed reliably by people who are subject to it and by officials who make determinations of compliance.

Suppose a legislator wants to minimize sidewalk injuries. The desired state of affairs covers a wide range of conduct; sidewalks can be busy places, hosting myriad activities and, therefore, myriad mishaps. Even if the legislator simplified the problem by limiting the state of affairs to, say, no slipping and falling from unshoveled snow, it might be difficult to derive a satisfactory norm to guide conduct. The legislator might begin with the norm, "All operators of land that abuts public sidewalks must have those sidewalks clear of snow after a snowfall or face a four hundred dollar fine."

Although this law provides somewhat specific guidance, it is arguably both too onerous and too lenient. On the one hand, it would seem to put too much of a burden on operators who experience a snowfall that ends at two o'clock in the morning. On the other hand, it might excuse people who live on highly traveled streets and let pedestrians eliminate the snow with their feet, thereby risking the very people the law was supposed to protect. It also suffers from semantic vagueness. What does it mean to be "clear of snow"? Does that include the thin layers of ice that remain after shoveling? The dusted snow that blows upon the sidewalk from post-storm winds? Further, it might not be obvious when snowfall has ended. What about intermittent waves of snow spaced out by a few hours?

A legislator seeking to provide adequate, up-front guidance will probably have to predict issues of that sort and revise the norm in a way that resolves them.[53] Doing so will likely require the accumulation of considerable background knowledge. The legislator should be aware of the variety and nature of conduct that affects the risk of snow-related falls. In this example, legislators should have some understanding of the relevant resources of the snow clearers,

---

change to existing law. That is not the focal point here; rather, I assume that the legislator is acting in earnest to bring about a known, desired state of affairs.

[52] Kaplow, *supra* note 41, at 568–69, 579–80.

[53] *See id.*

the character of winter storm systems, and the ways in which snow interacts with sidewalks both with and without removal efforts.

With that knowledge, the legislator should then consider the ways in which setting the norm's conditions will alter the behavior of those subject to the norm. Doing this part well requires knowledge of the direction in which behavior ought to be channeled as well as the severity of sanctions needed to motivate compliant conduct.[54] Here, the legislator should probably consider whether operators will need guidance as to the tools or chemicals that should be used, as well as whether the fine for non-compliance outstrips the costs of purchasing those resources in light of the frequency of snow.

This endeavor also forces the legislator to consider the reliability[55] of the norm under consideration. The good legislator will develop the capacity to draft a norm that will be comprehensible enough that those subject to it will change their conduct in a way that makes the desired state of affairs more likely to occur.[56] To succeed, the drafter might need to consider the reading levels of landowners or snow clearers and whether the test for compliance is precise enough. A judge or juror must also be able to understand the norm and have the capacity to apply it correctly and neutrally. This brings the additional concern of whether the adjudicator can adequately determine what the operative facts are.

Finally, the legislator should consider whether the norm would conflict with other existing laws.[57] If laws push subjects at cross-purposes, the norm might threaten the capacity of other laws to realize their ends. In this example, there might be noise ordinances that could place snow-removers in a catch-22, leading them not to clear snow at certain hours of the day.

Thus, a legislator's intent to provide clear guidance faces at least three areas that demand effort or otherwise impose costs: (1) knowledge regarding existing conduct and its surrounding conditions ("background knowledge");

---

[54] *See* Amanda M. Rose, *The Multienforcer Approach to Securities Fraud Deterrence: A Critical Analysis*, 158 U. PA. L. REV. 2173, 2188 (2010) ("Sanction setting itself requires a difficult balancing of under- and overdeterrence costs.").

[55] By reliability, I mean the likelihood that the norm, itself, leads to error. By error, I mean the norm brings about an unsatisfactory state of affairs from the perspective of those who legislated the norm. An error of this sort might occur because the norm sets an ineffective or inaccurate test for compliance.

[56] *See* LON L. FULLER, THE MORALITY OF LAW 63–65 (2d ed. 1969) (explaining laws must be comprehensible so that citizens should be able to identify what the laws prohibit, permit, or require).

[57] *See* Karen M. Gebbia, *The Keepers of the Code: Evolution of the Bankruptcy Community*, 91 AM. BANKR. L.J. 183, 198–99 (2017) ("Several factors contribute to clarity and coherence concerns in aging statutory law. These include a patchwork of conflicting judicial decisions, interim amendments that might be poorly drafted or might not integrate cleanly with existing law, and increasingly detailed and ossified statutory provisions that may reduce the law's flexibility to adjust to new and unanticipated circumstances.").

(2) creation of an instrument for conduct alteration that reliably leads to the desired state of affairs ("reliability"); and (3) understanding of potential conflict with existing law ("compatibility").

This is where the legislative value of standards becomes apparent.

Imagine that our snow-hating legislature decides that it would be too much work to fix their draft rule and so they scrap it for the standard, "All operators of land that abuts public sidewalks must keep their sidewalks in reasonable condition or face a four hundred dollar fine."

Their lack of background knowledge ceases to be a weighty cost to the legislator because the standard is designed to allow factual scenarios to play out before clear guidance is given. In this regard, it delegates to adjudicators the work of understanding the characteristics and variety of the conduct that is to be regulated by the norm.[58] Typically, this work falls to law-applying organs like judges who get those cases before them.[59]

A standard can also lower the cost to craft a reliable function for determining compliance.[60] Devising a standard like our, "Keep . . . sidewalks in reasonable condition" takes only a modicum of effort. As with scope, the hard work of reliably connecting conditions to their desired states of affairs is pushed downstream and broken up. The law-applying organs will likely serve as piecemeal legislators, announcing whether the facts before them comply with the standard and potentially creating a small rule in the process.[61] This is a less daunting assignment than the one that could have faced the initial legislator; judges or jurors are charged with determining compliance only for the case before them (and, in systems of precedent, those future cases that bear a sufficiently strong resemblance to it). Yes, these downstream legislators must learn more about the context in which they apply the rule, but at least they do not need to face every plausible context in which they might apply it. Ideally, this process of identifying the particular way to apply the principle in the standard (effectively converting it to a rule) in a particular case can be repeated, gradually providing an evolved and reliable norm—one that likely falls somewhere between a rule and a standard.[62] Even when judges have the impulse and opportunity to announce a more rule-like interpretation of a standard, they might do so with access to information regarding how human beings have interpreted and, thereaf-

---

[58] *See* Dale A. Nance, *Rules, Standards, and the Internal Point of View*, 75 FORDHAM L. REV. 1287, 1298 (2006) ("The use of standards can be criticized as involving inherent delegation of lawmaking authority to the decision maker employing the standard—that is, mostly judges, particularly trial judges.").

[59] *Id.*

[60] Kaplow, *supra* note 49, at 1307–09.

[61] *See id.*

[62] Frederick Schauer, *The Tyranny of Choice and the Rulification of Standards*, 14 J. CONTEMP. LEGAL ISSUES 803, 805–06 (2005).

ter, behaved under the existing legal standard (and its revisions). In theory, this will improve the reliability of their revision.

Finally, a standard lowers the costs of determining legal compatibility. Because the rule will gradually develop under the steady watch of the courts and is subject to the arguments of adversarial litigants, it is less likely that a standard will pose problems of compatibility. Although it is true that a legal standard is more likely to run afoul of the vagueness doctrine, that seldomly successful constitutional principle is unlikely to be a weighty concern, particularly in areas that do not involve criminal punishment or speech regulation.[63] Even those concerns are likely to dissipate over time as the standard becomes rulified.[64]

There are at least two important downsides, however. First, the legislator risks that the standard will be an unreliable conduct modifier, at least in its early stages. Cases might trickle into the courts, leaving ordinary folks largely rudderless in their attempt to interpret the standard when deciding how to act. Secondly, the legislator has relinquished a great deal of control over the eventual content of the norm to people who could have different ideologies or motivations. Consequently, they might end up dismayed with the manner in which the evolving norm regulates conduct or, worse, with the state of affairs that the evolving rule brings about—and it might take a long time to figure this out.

Although this brief hypothetical exercise might seem too abstract to be of much use, it has yielded two practical principles. First, rules and standards provide different but bilateral costs and benefits. Second, legislative costs can be divided into three meaningful categories: background knowledge, reliability, and compatibility.

B. Natural Language Processing of Rules and Standards

It is not controversial to claim that applying facts to the law is, at least in part, an interpretive exercise.[65] A lesser-known truth is that machine processing

---

[63] *See* Guyora Binder & Brenner Fissell, *A Political Interpretation of Vagueness Doctrine*, 2019 U. ILL. L. REV. 1527, 1538 (discussing frequency in criminal and speech contexts); Michael S. Kelley, *"Something Beyond": The Unconstitutional Vagueness of RICO's Pattern Requirement*, 40 CATH. U. L. REV. 331, 371 (1991) (explaining why invalidation under the vagueness doctrine is so rare).

[64] *See, e.g.*, Howell v. State Bar of Tex., 843 F.2d 205, 206 (5th Cir. 1988) (rejecting vagueness challenge); Villeneuve v. Connecticut, No. 10CV296, 2010 WL 4976001, at *5 (D. Conn. Dec. 2, 2010) (rejecting vagueness challenge because, "Although the plain text of the rules may lack precision and detail, it is clear from the rules, Official Comment, and case law interpreting substantively identical rules that the rules are limited to conduct that is relevant to and based on the practice of law, and therefore lawyers are sufficiently on notice of what the rules proscribe").

[65] *See generally, e.g.*, Owen M. Fiss, *Objectivity and Interpretation*, 34 STAN. L. REV. 739 (1982) (characterizing adjudication as an interpretive act). In saying this, I do not mean that it is only interpretive or that it is therefore subject to the same interpretive techniques that are used for literature. *Cf.*

is interpretive too. Computer code is a language consisting of rules that the machine is capable of interpreting and following.[66] Machines do not do this consciously (at least that is what we generally assume), but that does not prevent their performance of this function from providing considerable value to humans. The challenge of automating legal interpretation is essentially the challenge of mapping computer code to legal language. Legal language is natural language; it has evolved from human interaction.[67] Computer code is not natural language; scientists designed it according to a rigid, purpose-driven syntax.[68] Therefore, machine interpretation of natural language has the additional hurdle of translation.

Some scholars have suggested that law is a less challenging domain for translation because it is already code-like.[69] It is true that law uses norms to designate certain conduct as legally compliant or non-compliant, and that work, on a grand scale, could ideally create an organized network of directives.[70] Compared to ordinary talk, legal texts are orderly and contain a higher proportion of clear meanings and resolution protocols.[71]

Looking closer at a legal system, however, reveals that laws exist on a continuum with respect to their achievement of this ideal. Bright-line rules come closest,[72] so we can hypothesize that they will be the easiest to translate

---

Robin L. West, *Adjudication Is Not Interpretation: Some Reservations About the Law-as-Literature Movement*, 54 TENN. L. REV. 203, 205 (1987) (claiming that adjudication is an act of creation of law backed by force rather than an act of interpretation).

[66] RAMESH BANGIA, IT TOOLS AND APPLICATIONS 63 (2d ed. 2005).

[67] PRIMAVERA DE FILIPPI & AARON WRIGHT, BLOCKCHAIN AND THE LAW: THE RULE OF CODE 199 (2018) ("Not all laws can be easily translated into code. Legal rules are written in natural language, which is, by its very nature, inherently flexible and ambiguous.").

[68] ROLAND HAUSSER, FOUNDATIONS OF COMPUTATIONAL LINGUISTICS: HUMAN-COMPUTER COMMUNICATION IN NATURAL LANGUAGE 4 (3d ed. 2014).

[69] *See, e.g.*, Harry Surden, *The Variable Determinacy Thesis*, 12 COLUM. SCI. & TECH. L. REV. 1, 24 (2011) ("Given that most laws can be formulated into if-then statements, and that much of legal decision-making is characterized, at a high level, in terms of deductive structure, we can begin to see the natural appeal between computerized automated reasoning systems—which take complicated 'if-then' rules and analyze them applying deductive logic—and the law.").

[70] *See* Henry Prakken & Giovanni Sartor, *Law and Logic: A Review from an Argumentation Perspective*, 227 ARTIFICIAL INTELLIGENCE 214, 217–22 (2015) ("So far everything [regarding the use of logic-based artificial intelligence (AI)] has been compatible with viewing the law as an axiomatic system, where rules and facts are represented in a logical language and deductive logic is used to derive legal consequences from the representations. At first sight, it might be thought that the axiomatic approach to the law is committed to deductive logic. However, this approach has been broadened to include nonmonotonic techniques to deal with two very common structural features of legal regulations, the separation of general rules and exceptions, and the use of hierarchies over legislative sources to resolve conflicts between different regulations within a normative system.").

[71] *See id.*

[72] *See, e.g.*, Eric Engle, *Legal Interpretation by Computer: A Survey of Interpretive Rules*, 5 AKRON INTELL. PROP. J. 71, 92 (2011) ("Formalist rules are easiest to model computationally, because the results are most predictable; the lament of 'mechanical jurisprudence.'").

into computer code. A series of if-then commands might even do the trick.[73] Take a speed limit, for example. If the number entered for traveled speed is higher than the number entered for the speed limit, then the punishment function is turned on.[74] Of course, not all legal tests, even those that are rule-like, correspond to a clear numerical trigger. For example, the statutory rule that copyright protection be granted to works "fixed in any tangible medium of expression" does not have that feature.[75] And for a machine to be able to handle a variety of legal tests over a long period of time, the machine would need to be programmed with a robust capacity to interpret natural language in a variety of cases, including those that are less rule-like. Enter artificial intelligence.

The research field in which AI is used to interpret natural language is called Natural Language Processing (NLP). NLP scientists, drawing from linguistics and philosophy, find it useful to break natural language into three components: syntactics, semantics, and pragmatics.[76] Linguists and computer scientists often define syntax as the formal relations of signs to one another or, more broadly, to be the formal aspects of a language.[77] They understand semantics to be the relations between expressions or, more specifically, to be that to which a sign refers. Finally, they understand pragmatics to be the relations among expressions, their meanings, and the use that speakers make of these expressions in contexts of utterance.[78] These components are in ascending order of complexity, and this complexity correlates with resistance to computer codification.[79] Unsurprisingly, it also correlates with the amount of time we will have to wait for machine performance to meet human performance thresholds. Experts estimate that significant reduction of the gap between human and machine progress on semantics will occur over the next three decades, with significant progress occurring on pragmatics in the three decades after that.[80]

---

[73] Surden, *supra* note 69, at 24.

[74] Interestingly, it is not obvious that even the translation of speed limit enforcement to computer code would not require the use of discretion when it comes to letter-of-the-law or spirit-of-the-law choices. An experimental study in which fifty-two programmers were asked to take on this task resulted in the finding that "[u]nfortunately, laws were not created with automated enforcement in mind, and even seemingly simple laws have subtle features that require programmers to make assumptions about how to encode them." Lisa Shay et al., Abstract: Do Robots Dream of Electric Laws? An Experiment in the Law as Algorithm, Presentation at the We Robot Conference 2013 (Apr. 8–9, 2013), http://www.gregconti.com/publications/201303_AlgoLaw.pdf [https://perma.cc/C5CW-UETC].

[75] 17 U.S.C. § 102.

[76] *See, e.g.*, MAURIZIO GABBRIELLI & SIMONE MARTINI, PROGRAMMING LANGUAGES: PRINCIPLES AND PARADIGMS 27–28 (2010); Erik Cambria & Bebo White, *Jumping NLP Curves: A Review of Natural Language Processing Research*, 9 IEEE COMPUTATIONAL INTEL. MAG. 48, 48–49 (2014).

[77] Sheppard, *Incomplete Innovation*, *supra* note 8, at 1851.

[78] *Id.*

[79] Cambria & White, *supra* note 76, at 48.

[80] *Id.* at 51 fig.1.

Today, the AI powering NLP is primarily syntactical: it identifies correlations between natural language queries and natural language answers by relying on the sequence of characters or audio waveforms for each.[81] The most well-known systems that use this approach are Amazon's Alexa or Apple's Siri, which rely on a treasure trove of links between internet search field inputs and ranked search results that have been approved through the conduct of millions upon millions of users.[82] When these programs identify a strong enough correlation between an input and the desired user response—in the context of search, the desired response is often a user accessing a particular search result or a portion thereof—then the program can use that information to produce a sound that resembles a spoken answer to a question.[83] That answer is effectively a vocalization of the top search result.[84] Artificial intelligence of this sort might not be smart, but it is useful. It is clear that syntactical processing and rudimentary semantics can produce impressive results when sufficient high-quality data is analyzed.

Semantically powered NLP systems are not as ubiquitous as syntactically powered systems, but they are becoming more commonplace. Presently, there are numerous products and research initiatives that use AI to construct semantic webs of concepts known as ontologies.[85] These webs aim to allow machines to distinguish between identical sequences of characters or sounds based on linkages to categories of meaning. A semantically powered machine might be able to identify the themes in a brand-new poem even if those themes are abstracted and challenging to identify on a straightforward reading.[86]

As for pragmatics, the industry is at a relatively rudimentary state of development, lagging far behind human ability.[87] Humans are equipped with a massive store of structured background knowledge that enables us to identify how contextual factors can change the meaning of a statement. Computers are not. If I walk into a McDonald's and ask a friend, "What should I do?," the friend can readily understand that I mean, "What should I order?" And if that

---

[81] *Id.* at 51.

[82] Josh Hendrickson, *Alexa, Siri, and Google Don't Understand a Word You Say*, HOW-TO GEEK (Feb. 19, 2019), https://www.howtogeek.com/405011/voice-assistants-dont-really-understand-you/ [https://perma.cc/MG89-P7W4].

[83] *Id.*

[84] *See id.*

[85] *See* Maryam Ramezani et al., *Using Machine Learning to Support Continuous Ontology Development*, *in* KNOWLEDGE ENGINEERING AND MANAGEMENT BY THE MASSES 381, 381–90 (Philipp Cimiano & H. Sofia Pinto eds., 2010).

[86] *See* Geetanjali Rakshit et al., *Automated Analysis of Bangla Poetry for Classification and Poet Identification*, *in* PROCEEDINGS OF THE 12TH INTERNATIONAL CONFERENCE ON NATURAL LANGUAGE PROCESSING 247, 247 (Dipti Misra Sharma et al. eds., 2015).

[87] JASON WHITTAKER, TECH GIANTS, ARTIFICIAL INTELLIGENCE, AND THE FUTURE OF JOURNALISM 119 (2019).

friend also knows that I am trying to lose weight, they might understand my question as, "What's the best combination of flavor and healthiness on the menu?" For humans, this is no great feat, but for a machine it is.[88]

It turns out that the rules/standard distinction is useful for approximating the challenge that an individual norm would pose for NLP in its current state. Rules can limit the role that pragmatics and, to a lesser extent, semantics, play in interpretation.

Rules streamline decision-making and cabin discretion by limiting the number of considerations necessary to reach the correct[89] interpretation; they also provide relatively clear guidance regarding the analysis of those considerations. Thus, they are, comparatively speaking, self-contained. Interpretation of legal norms does not demand that we rely solely upon syntactics when interpreting legal rules, but individual canons of construction can come close. For instance, the popular "plain meaning rule" warns against unnecessary resort to pragmatics-laden factors such as public policy or morality and seeks to limit the world of permissible semantic meanings to those that are ordinary.[90] This is not to suggest that semantics or pragmatics play no role when it comes to norms at the rule end of the spectrum. Rules make it less likely that a sophisticated capacity to understand semantics and pragmatics will be necessary to apply the law correctly to a given set of facts.[91]

---

[88] To be sure, existing technology might be able to detect a correlation between the combination of the sound waveform "what should I do," the GPS coordinates at which the sound was detected, the character string "diets and fast food" that had been entered into a search engine on that phone hours earlier, and the ordering of a McDonald's Santa Fe salad, but only if the background data are there and in sufficient number and quality to give rise to statistical significance.

[89] By "correct," I mean only an interpretation that squares with the consensus intention of stakeholders, like legislators and judges. I do not mean to suggest that there is a context-independent, linguistically correct or default model of interpretation, which is still a topic of debate among philosophical linguists. Katarzyna M. Jaszczolt, Defaults in Semantics and Pragmatics, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2018), https://plato.stanford.edu/entries/defaults-semantics-pragmatics/ [https://perma.cc/2ZQ4-CQUW].

[90] William Baude & Ryan D. Doerfler, *The (Not So) Plain Meaning Rule*, 84 U. CHI. L. REV. 539, 540 (2017) ("Many tenets of statutory interpretation take a peculiar form. They allow consideration of outside information—legislative history, practical consequences, the statute's title, etc.—but only if the statute's text is unclear or ambiguous. These tenets are often expressed as a variation of the 'plain meaning rule.'" (emphasis omitted) (footnote omitted)); Melvin J. Sykes, *A Modest Proposal for a Change in Maryland's Statutes Quo*, 43 MD. L. REV. 647, 655 (1984) (describing the plain meaning rule as "a popular canon").

[91] In this regard, it is possible to draw a parallel between my views and the views of Lee and Mouritsen with respect to pragmatics. Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 YALE L.J. 788, 816 (2018) ("Literalist semantic meaning alone is not an indication of ordinary communicative content. Real human beings do not derive meaning from dictionary definitions and rules of grammar alone. Everyone takes nonsemantic context—pragmatics—into account in deriving meaning from language."). My characterization of the syntactics/semantics/pragmatics triad is slightly different than theirs, however. They group syntactics and semantics under formal aspects of utterances as opposed to pragmatic aspects and combine syntactics with semantics when they refer to

I should note that some scholars have pushed to harness technology to expand the realm of cases that could be decided through plain meaning to capture legal norms that might otherwise suffer from ambiguity. Lee and Mouritsen argue that computer-powered corpus linguistics—using computers to analyze large bodies of natural language data and identify patterns of usage—will allow courts to resolve interpretations of statutory meaning in some cases when the statute contains language that is not used in a specialized sense.[92] Regardless of whether corpus linguistics is going to be useful, rules are much less likely to make resort to contextual features necessary because they tend to contain terms that are more precise and are therefore more likely to exhibit uniformity of usage.

Standards, by contrast, place a premium on semantic, and especially-pragmatic, understanding. They typically enshrine moral and evaluative principles, like reasonableness, which are notorious for being objects of disagreement when applied to specific cases.[93] The determination of whether a punishment is illegally "cruel"[94] might call for complex semantic work, such as consideration of how the concept of cruelness links to the concepts that underlie other illegal conduct, such as torture,[95] or that underlie legal values, such as dignity.[96] Or it might require pragmatic considerations such as whether a punishment is compatible with state commitments to justification, such as proportionality[97] or deterrence.[98] Thus, a standard makes it hard for a machine to identify

---

"semantic meaning." *Id.* at 815, 821. I find it useful to make a harder distinction between semantics and syntactics; I link semantics to analysis of meaning and, thus, connect it to the effort to teach machines to process natural language with regard to word meaning rather than the effort to have machines use syntactical protocols for parsing sequences of characters. And although it is true that both syntax and semantics are contexts, the former is a relatively arid one—the rules of syntactical processing, such as grammar rules, are less sensitive to extra-textual factors that come from the environment of the utterer than semantics and pragmatics.

[92] *Id.* at 795, 872.

[93] *See* Matthew D. Adler & Eric A. Posner, *Rethinking Cost-Benefit Analysis*, 109 YALE L.J. 165, 218 (1999) (discussing the moral criteria of standards).

[94] *See, e.g.*, U.S. CONST. amend. VIII. ("Excessive bail shall not be required, nor excessive fines imposed, nor cruel and unusual punishments inflicted.").

[95] *Cf.* John F. Stinneford, *Rethinking Proportionality Under the Cruel and Unusual Punishments Clause*, 97 VA. L. REV. 899, 947 (2011) ("Although one function of such a prohibition was to prevent Congress from approving the use of torture, proponents of the Cruel and Unusual Punishments Clause also wanted to prevent the imposition of retroactive punishments and (most relevant for our purposes) excessive punishments.").

[96] *Cf.* Trop v. Dulles, 356 U.S. 86, 100 (1958) ("The basic concept underlying the Eighth Amendment is nothing less than the dignity of man.").

[97] *Cf.* Ewing v. California, 538 U.S. 11, 20 (2003) (discussing the narrow proportionality principle in noncapital cases).

[98] *Cf.* Roper v. Simmons, 543 U.S. 551, 571 (2005) (discussing the role of deterrence in a cruel and unusual punishment analysis).

and execute the correct mode of resolution because that process so often requires recognition of legal semantics or pragmatics.

To demonstrate this principle, I performed an at-home test with Siri (which relies primarily upon syntactical NLP):

> Me: Siri, set up a wake-up alarm tomorrow at 8:00 AM.
> Siri: Your alarm is set for 8:00 AM.
> Me: Siri, set an alarm at a reasonable wake-up time tomorrow.
> Siri: Set the 'at a reasonable wake up time' alarm for when?
> Me: At a reasonable time.
> Siri: Set the 'at a reasonable time alarm' for when?
> Me: [sigh]

Does this mean that machines will have to master pragmatics before they can reach the correct interpretive outcome under legal standards?[99] Almost certainly not. For example, empirical studies in a non-legal context have shown that syntactical approaches can sometimes reach the same interpretations of difficult texts as semantic or pragmatic approaches.[100] In a legal context, syntactically powered natural language searching in Westlaw or Lexis helped propel them to be industry leaders.[101] This comparatively anemic approach, however,

---

[99] It is possible, of course, that the law demands going through the same process to reach an interpretive outcome.

[100] Google's BERT system (a deep mind neural network) is able to provide state-of-the art results on language understanding and reasoning tests by using predictive models that parse sentences backwards and forwards at the same time. John Pavlus, *Machines Beat Humans on a Reading Test. But Do They Understand?*, QUANTA MAG. (Oct. 17, 2019), https://www.quantamagazine.org/machines-beat-humans-on-a-reading-test-but-do-they-understand-20191017/ [https://perma.cc/4UHX-CB7R]. BERT can even best human performance on advanced reading tests, though it does not rely on networked conceptual knowledge. Rather, it constructs trees on whatever patterns it happens to detect in the materials upon which it is trained. *Id.* Interestingly, a recent syntactical utilization of BERT improved BERT's performance, potentially even for semantic tasks. Adhiguna Kuncoro et al., *Syntactic Structure Distillation Pretraining for Bidirectional Encoders*, 8 TRANSACTIONS ASS'N FOR COMPUTATIONAL LINGUISTICS 776, 783–86 (2020), https://direct.mit.edu/tacl/article/doi/10.1162/tacl_a_00345/96469/Syntactic-Structure-Distillation-Pretraining-for [https://perma.cc/5FB4-2WBJ]; Reina Qi Wan, *DeepMind Says Syntactic Biases 'Helped BERT Do Better,'* SYNCED (May 29, 2020), https://syncedreview.com/2020/05/29/deepmind-says-syntactic-biases-helped-bert-do-better/ [https://perma.cc/GTQ4-34AW]. This result, however, should not be taken to mean that approaches that additionally utilize semantic and pragmatic methods will not eventually provide better results.

[101] Semantic processing is making headway in that space, though experts claim that Westlaw and Lexis are primarily syntactical. *See* Erin Yijie Zhang, *Legal Applications of Neural Word Embeddings*, TOWARDS DATA SCI. (Aug. 6, 2020), https://towardsdatascience.com/legal-applications-of-neural-word-embeddings-556b7515012f [https://perma.cc/5ERG-8EJ5] (describing how word embeddings improve semantic understanding in legal information retrieval and comparing that to Westlaw and Lexis, which primarily "look[] for literal matches or variants of the query keywords, usually by using string-based algorithms to measure the similarity between two text strings.").

increases the risk of error or non-answer, particularly if background data is sparse or does not indicate a consensus.

Eschewing additional and often messy considerations, as rules do, helps NLP performance in at least two ways. First, the limitation of discretion makes it more likely that the inputs of law and facts and the output of decision form statistically significant correlations—rules limit variation by channeling facts and connecting them more directly to outcomes.[102] This means that a statistical relation is more likely to emerge through a simple model of legal analysis, even one that uses syntactical parsing of the operative language.[103] Second, the simplicity of rule content is less likely to force interpreters into a complex web of conceptual meaning, either because they rely on bright-line triggers or because they use content with clearer terminology that provides satisfactory meaning without abstraction. Therefore, a simpler semantic ontology might do the trick, making hand-crafted semantic approaches more feasible.

We have covered the basics of the relationship between the rules/standards distinction and the costs of automation, but this Article has thus far contemplated only pure rules and standards. In the American legal system, pure rules and standards are exceptional and impermanent.[104] Although rules can become more standard-like over time, the centerpiece of this Article is a contemporary standard, so my focus will be the inverse process—the rulification of standards. In the following section, I discuss how the repeated application of standards can give rise to competing interpretations of that standard, some of which are rule-like.

C. NLP Between Rules and Standards

Even though norms that are categorizable as legal standards might pose a challenge for syntactical or rudimentary semantic NLP techniques, it would be unwise to assume that automation of interpretation of those norms is not possible using technology that currently exists or will exist in the coming decades.

For one, some norms containing predominantly standard-like norm content have significant rule content as well, making them fall between pure rules and pure standards. Lots of norms fall into this gap: a law might be an impure standard such that the language within the norm provides clear direction in a relatively small subset of the cases to which it is applied but is typically murky;

---

[102] *See generally* Sheppard, *Judging Under Pressure*, *supra* note 39 (analyzing the relative constraining power of rules and standards through behavioral experiments).

[103] *See id.*

[104] *See, e.g.*, Richard L. Heppner, Jr., *Conceptualizing Appealability: Resisting the Supreme Court's Categorical Imperative*, 55 TULSA L. REV. 395, 403 (2020) ("As a result, the various appealability doctrines are spread along the rules-standards continuum. Few doctrines are pure rules or pure standards.").

or it might be an impure rule that provides bright-line guidance within a limited range of situations but contains material, standard-like exceptions.[105] The norms might have this in-between character from their initial legislation, but they might alternatively develop it over time; each application of the standard in a case can create a sub-rule that applies to all factually similar cases, particularly in a system of precedent.[106] To be clear, the norms with these features are readily understood as being neither pure rules nor pure standards.

But what about situations in which there is no consensus that a standard has lost its standard-ness? The vagueness of a legal standard might lead to competing conceptions of the proper test to be used in the norm, some of which might be standard-like and some of which might be rule-like. One way that this occurs is when a court announces a new norm for applying a legal standard in hope that their norm will constrain future applications within a desired scope.[107] A norm of this sort might catch on with other judges to such an extent that it becomes the dominant interpretation for all applications of the standard, it might be utterly disregarded, or it might fall somewhere in between.[108] This process is repeated whenever other judges announce their own supplemental norms, and those too will face one of these three fates. At various points in the evolution of a legal standard's application in the courts, then, there may exist for some scope of cases competing interpretations of that standard. Some of those interpretations might be rule-like. Thus, the openness of standards can create a void that will be filled by norms that are less standard-like. These norms might come from judges or other influential interpreters, such as well-respected scholars or special-interest groups. Indeed, this sort of

---

[105] *See* Michael Coenen, *Rules Against Rulification*, 124 YALE L.J. 644, 648 (2014); Russell D. Covey, *Rules, Standards, Sentencing, and the Nature of Law*, 104 CALIF. L. REV. 447, 490 (2016) (discussing sentencing guidelines).

[106] Coenen, *supra* note 105, at 648; Covey, *supra* note 105, at 490.

[107] *See* Schauer, *supra* note 62, at 809–10 ("So when a court, say, makes a standard more precise and limits the choices available to that court in future cases, it also limits the choices available to future courts in future cases. If the current court anticipates that the future court will make more mistaken (from the perspective of the current court) decisions under a standard than the current court anticipates that it will itself be constrained by its own rules to make mistaken decisions, then it will have good reason, assuming its interest is in maximizing its own policy preferences over time, in imposing a rule-based constraint on itself. We could formalize this (more accurately, someone else could formalize this), discounting for the fact that neither precedent nor path dependence are insurmountable." (footnote omitted)).

[108] I take this to be an uncontroversial point, but if further explanation is needed about how a supplemental norm could be articulated in a system of precedent and not become dominant upon announcement, I will offer a few ways. For example, the supplement might be offered somewhat timidly, starting as dicta before eventually transforming into the dominant interpretation. Or it could come in a decision from a lower or intermediate court, which does not bind the courts higher in the food chain. Or, even judges on the same court might not engage in strict observance of horizontal stare decisis, so the supplement must vie against those offered by judges' colleagues.

work has historically been the bread and butter of prestigious commentators like the American Law Institute.[109] In other words, norms that begin as legal standards can become marketplaces for interpretation; they can give rise to alternative and competing approaches to interpretation, some of which might have a rule-like character.[110] This might be the case with the RPPT: although there is a consensus that the concept of reasonableness at the heart of the breach inquiry under negligence doctrine is fairly characterized as a paradigmatic[111] or even pure standard,[112] there is not a consensus as to the particular conception of reasonableness that should be used for all negligence cases.[113]

This competing conception scenario is an alluring one for proponents of automation. The prospect of a rule-like conception becoming dominant means that what initially appeared to be the most daunting translation of law into computer code has transformed into something much more feasible.

In summary, legal standards pose computational challenges because they are more likely to require resorting to pragmatics or complex semantics to arrive at legally permissible interpretations when they are applied to a case. Pragmatics and semantics are problematic for automation because existing NLP technology is not yet able to meet human performance with respect to those aspects of textual interpretation. Despite these challenges, the openness of standards gives rise to the possibility of rulification—through application of or through the introduction of legally permissible, rule-like interpretations of the legal standard—which can assist the process of automation.

---

[109] *See How the Institute Works*, AM. L. INST., https://www.ali.org/about-ali/how-institute-works/ [https://perma.cc/WLA4-7MNU] ("The Institute's mission is 'to promote the clarification and simplification of the law and its better adaptation to social needs . . . .'").

[110] *See, e.g.*, Deborah L. Rhode, *Moral Character as a Professional Credential*, 94 YALE L.J. 491, 551(1985) ("Moreover, in most jurisdictions, conduct is considered disabling only if it involves 'moral turpitude,' a standard open to competing interpretations.").

[111] Benjamin C. Zipursky, *Reasonableness In and Out of Negligence Law*, 163 U. PA. L. REV. 2131, 2133 (2015) ("The word 'reasonable' is a paradigmatic example of a standard in the law, and its meaning is, if nothing else, vague." (emphasis omitted)).

[112] Russell B. Korobkin, *Behavioral Analysis and Legal Form: Rules vs. Standards Revisited*, 79 OR. L. REV. 23, 27–28 (2000) ("The basic negligence requirement of acting as would a 'reasonable person' is close to a pure standard. The legal pronouncement specifies no facts that would automatically trigger a finding of negligence, and no facts that would trigger a finding of non-negligence. Furthermore, the standard does not even identify facts that would be evidentiary of either outcome."); *cf.* Larry Alexander, *Incomplete Theorizing: A Review Essay of Cass R. Sunstein's* Legal Reasoning and Political Conflict, 72 NOTRE DAME L. REV. 531, 542 (1997) (book review) ("Perhaps 'act as would a reasonable person' is the closest we get to a pure standard.").

[113] *See* John Inazu, *Beyond Unreasonable*, 99 NEB. L. REV. 375, 376 (2020) ("The concept of 'reasonableness' permeates the law: the 'reasonable person' determines the outcome of torts and contract disputes . . . . But as any first-year law student can attest, the line between reasonable and unreasonable is not always clear. Nor is that the only ambiguity." (footnote omitted)).

D. Automation and the Value Proposition of Standards

It remains to explain how the character and dynamics of legal norm content bears upon the utility of automation of legal interpretation in terms of background knowledge, reliability, and compatibility. Arguably, the most important contribution to this scholarship thus far comes from Anthony Casey and Anthony Niblett and, in particular, their claim that machines will lead to the end of the rules/standards tradeoff. They believe that machines will lower the cost of rule creation by providing an inexpensive means to solve the problem of path discernment—identifying the best path for conduct change—thereby enabling the legislation of context-driven rules of narrow scope called "microdirectives."[114]

The machine-produced directives Casey and Niblett imagine are clear and straightforward and, therefore, offer the constraining power of rules.[115] What makes them special, however, is that the machine uses its processing power to take contextual factors into account; AI will create a host of microdirectives tailored to the very particular circumstances faced by those subject to the law.[116] And they do it quickly. On their account, machines will bear the burden of processing contextual data to discern the best path to the desired social outcome.[117] Because machines will someday have the capacity to process this information at a scale and speed impossible under the existing common law system, Casey and Niblett suppose, the automated system better serves the legislators' ends than traditional legislative methods.[118] That is, legislators will no longer face the difficult choice between rules and standards because machines will someday enable the legislators to articulate broad objectives cheaply—like a standard—with downstream constraining power—like a rule—through the automated articulation of context-sensitive directives.[119] They hope that a broad directive like "minimize sidewalk injuries" will be disambiguated into narrowly applying directives to those who find themselves in those situations.

---

[114] Anthony J. Casey & Anthony Niblett, *The Death of Rules and Standards*, 92 IND. L.J. 1401, 1403 (2017).

[115] *Id.* at 1405 ("We describe the rise of microdirectives as the death of rules and standards. One might alternatively frame the coming change simply as the death of standards. After all, microdirectives are ex ante rules that govern behavior.").

[116] *Id.* at 1410–12.

[117] *Id.* at 1446 ("As standards disappear and judges have progressively less influence, legislative intent will be entrenched and concretized in the catalog of microdirectives.").

[118] *Id.* at 1412–14 ("If the predictive technology is very powerful and the machine is able to provide precise and accurate information, then this points in favor of using the machine to create a rule, rather than relying on a judge to adjudicate a standard.").

[119] *See id.*

Moreover, these directives will respond to changes in behavior so as to further the overall objective.[120]

To their credit, Casey and Niblett do not assert that an existing legal system could easily be converted to their microdirective approach. They imagine that conversion will happen incrementally over time, with administrative agencies playing a role. In particular, they foresee an era in which "[l]egislatures may continue to enact standards, but they will leave the machine-aided implementation to regulators[] . . . [who] will then translate that broad objective into specific sets of rules generated by machines."[121] In turn, they predict, legislatures will be able to focus primarily on "stat[ing] their objectives [regarding social policy] at increasingly higher levels of abstract[ion] . . . ."[122]

This is arguably a fanciful account of our legal future. But, for my purposes here, the value of their work is that it shows how automation might change the value proposition for legislators with respect to the factors of background knowledge, reliability, and compatibility even if a microdirective-based legal system never comes into being.

The realization of Casey and Niblett's vision would greatly reduce background knowledge costs, which are ordinarily higher with legal rules than with legal standards.[123] Artificial intelligence techniques could allow legislators to rely on machines to identify patterns of conduct and the environmental conditions in which they occur. Whether from carefully created models or brute force processing power, machines can process data and affix statistically determined weight to the behavior correlates of desired outcomes; the success rate of prediction increases when the process is capable of automated improvement from the acquisition of data regarding prior performance.[124] With these correlates known, the costs of devising a reliable function for maximizing their occurrence decrease as well: it is largely a matter of imposing conditions that mandate positive correlates or forbid negative correlates. Of course, statistical AI techniques are at the mercy of the availability of data, which

---

[120] *Id.* at 1433 ("There is a feedback effect. The devices that form part of the Internet of Things also collect data on how individuals and corporations behave. Lawmakers can generate even better predictions of human behavior by harnessing such data. In doing so, the Internet of Things will further reinforce the feasibility of the predictive technology." (emphasis omitted) (footnote omitted)).

[121] *Id.* at 1437.

[122] *Id.*

[123] *Id.* at 1410 ("Lawmakers will no longer have to think up rules to enact laws. Judges will no longer have to examine citizens' decisions on a case-by-case basis in order to apply laws. And the laws will be highly calibrated to policy objectives with no chance of judges introducing bias or incompetence.").

[124] *See* Lisa L. Harlow & Frederick L. Oswald, *Big Data in Psychology: Introduction to the Special Issue*, 21 PSYCHOL. METHODS 447, 447–57 (2016) (describing several articles that use massive data and AI techniques to understand or predict human behavior and personality).

might be insufficient to yield useful correlations.[125] As time goes on, however, it is quite likely that the quantity of collected data will grow[126] and the speed of data processing and access will increase,[127] which will drive performance costs down.[128]

Secondly, the technology that powers Casey and Niblett's vision can increase the reliability of interpretation by reducing the number of human interpreters that are necessary to convert the standard to a particularized rule. In this regard, automation might reduce errors of human cognition or personal bias.[129] This is not to suggest that automation will always increase reliability.

---

[125] *See, e.g.*, Dan McQuillan, *People's Councils for Ethical Machine Learning*, 4 SOC. MEDIA + SOC'Y 1, 2 (2018) ("Machine learning brings with it another characteristic which can cause or obscure harm, which is the opaque nature of its decision-making. As has been pointed out, machine learning depends on big data because the training set needs to be large enough to generate a useful predictive model.").

[126] In just the two-year span between 2017 and 2019, the quantity of worldwide stored data doubled from ten thousand to twenty thousand exabytes. POTOMAC INST. FOR POL'Y STUD., THE FUTURE OF DNA DATA STORAGE 6–7 (2018), https://potomacinstitute.org/images/studies/Future_of_DNA_Data_Storage.pdf [https://perma.cc/2AMM-DUJK] (discussing how the burgeoning field of DNA data storage "will usher in a new paradigm for computing with little to no limitations on the volume of data that we can produce, store, and access").

[127] Matthew Stewart, *The Future of Computation for Machine Learning and Data Science*, TOWARDS DATA SCI. (Oct. 24, 2019), https://towardsdatascience.com/the-future-of-computation-for-machine-learning-and-data-science-fad7062bc27d [https://perma.cc/LJK9-87FG] ("In summary, the future of computation looks like it will involve speeding up computations to handle the relentless and exponential increase in data production.").

[128] To be clear, I am not saying that increases will be uniform or exponential. Indeed, Moore's Law of Exponential Increase regarding the number of transistors on a chip might have failed in recent years as a matter of production, and it appears to have already failed to correspond to exponential increases in processing performance. *Id.* ("However, speeding up individual processors is difficult for various reasons, and Moore's law cannot last forever—it is becoming increasingly constrained by the limits of heat transfer and quantum mechanics."). Rather, my point is modest: over time, these measures will tend to go up, and performance costs for the same data task will tend to go down.

[129] *See* Chris Chambers Goodman, *AI/Esq.: Impacts of Artificial Intelligence in Lawyer-Client Relationships*, 72 OKLA. L. REV. 149, 150 (2019) ("[AI] can identify and minimize bias in client intake and initial consultations, it can assess the uniformity of criminal charging decisions made by prosecutors, and it can help to diversify law firm ranks, judicial ranks, and even juror pools. AI can also reduce the impacts of implicit bias by providing a mechanism for enhancing empathy, and by expanding the scope of information that lawyers rely upon . . . ."); Kimberly A. Houser, *Can AI Solve the Diversity Problem in the Tech Industry? Mitigating Noise and Bias in Employment Decision-Making*, 22 STAN. TECH. L. REV. 290, 294 (2019) ("The responsible use of artificial intelligence, however, can mitigate unconscious bias by reducing the impact of human decision-makers on the process, and create better employment decisions which are based on skills, traits and behaviors rather than factors (such as sex, race, or pedigree) that do not correlate with merit or success." (footnote omitted)); *cf.* Avital Mentovich et al., *Are Litigation Outcome Disparities Inevitable? Courts, Technology, and the Future of Impartiality*, 71 ALA. L. REV. 893, 898 (2020) ("We find that the shift from traditional in-person judicial proceedings to online (or at least differently organized) proceedings reduces measured age- and race-based disparities in litigation outcomes."). *But see* Rebecca Crootof, *"Cyborg Justice" and the Risk of Technological-Legal Lock-In*, 119 COLUM. L. REV. F. 233, 240

Sometimes the inverse occurs. For example, there is a growing literature on how machine learning techniques that rely and learn from unstructured and unmonitored human feedback risk enshrining reprehensible human biases and masking them behind opaque algorithms.[130]

Finally, an automated system might be able to perform in a way that minimizes internal inconsistency or other problems of legal compatibility. For example, programmable defeasible logic could be used to resolve conflicts between rules in the system by setting up a hierarchy of metarules;[131] conceivably, if human legislators select a trumping policy objective, the metarule could use that objective as a means to choose one rule over another based on a statistical assessment of the degree to which each rule furthers the goal.

In conclusion, Casey & Niblett's vision has been useful to illustrate that automation could very well be an economically attractive alternative for legislators because it has the capacity to reduce costs associated with background knowledge, reliability, and compatibility. Understanding, as Casey and Niblett do, that an automated legal future would come gradually, it serves to consider whether automation technology will be of use under existing law, particularly those laws that are already stated in abstract terms. Even on their account, then, there might come a time where there is piecemeal automation of the legal system. Such a project would likely face a crucial and potentially pesky challenge—the RPPT.

II. The Competing Conceptions of the Reasonably Prudent Person Test and the Variable Costs of Their Automation

On the rule/standard continuum, the RPPT is a standard.[132] It invokes the concepts of "reasonable" and "prudent," terms that we associate with evaluative judgment. Were computer scientists limited to the face of the test and not allowed to consider the supplemental interpretations that have emerged since its introduction, then there would be no way for them to handcraft a simple rule-based programming structure that would be satisfactory.

(2019) ("Further, despite its veneer of objectivity, AI will not solve the problem of human bias; it incorporates human bias and adds other kinds.").

[130] *See generally, e.g.*, Safiya Umoja Noble, Algorithms of Oppression: How Search Engines Reinforce Racism (2018); Cathy O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy (2016); Will Knight, *Biased Algorithms Are Everywhere, and No One Seems to Care*, MIT Tech. Rev. (July 12, 2017), https://www.technologyreview.com/s/608248/biased-algorithms-are-everywhere-and-no-one-seems-to-care/ [https://perma.cc/G66X-K2T9].

[131] *See* Noor Sami Al-Anbaki et al., *A Defeasible Logic-based Framework for Contextualizing Deployed Applications*, 10 Int'l J. Advanced Comput. Sci. & Applications 176, 177 (2019).

[132] *See* discussion *supra* Part I.

Helpfully, the RPPT is an *old* standard,[133] and therefore it has given rise to competing conceptions of the test within it. Importantly, some of these conceptions are amenable to simple computer programming solutions. Some might be solved by today's state-of-the-art AI while others might require something a bit beyond today's AI; and others still call for a leap in technological innovation. For the conceptions that might be formalized into computer code using something resembling existing techniques, we should first consider their cost with respect to two features of legislation: the background knowledge (or, in computational terms, background "data") and the reliability of the function that emerges.[134] Doing so will reveal that the various interpretations of the RPPT under consideration do not hit the cost centers of legislation uniformly.

Some approaches bring costly background knowledge challenges, such as the need for massive amounts of training data. For other approaches, the dominant cost is reliability. For example, the computer's approach poses a significant risk that it will not generate a legal test that brings about desired outcomes because it is not adequately sensitive to context.

Even if background data and reliability problems are surmountable, a further complication is that these alternative approaches to RPPT have varying degrees of legal acceptability. Some approaches are more popular or more coherent with existing laws than others and, therefore, present fewer compatibility costs.

In a paper of this size, it would be impossible to perform a fifty-state analysis of the evolution of the RPPT over the course of the last century. Thankfully, it is unnecessary. Two excellent articles have identified the leading alternatives: Patrick Kelly and Laurel Wendt's *What Judges Tell Juries About Negligence: A Review of Pattern Jury Instructions*[135] and Kevin Tobia's *How People Judge What Is Reasonable.*[136] The former is a comprehensive, fifty-state review of model jury instructions on the RPPT.[137] The latter is an empirical study of how laypeople judge what is reasonable.[138] Therein, Tobia identi-

---

[133] Stephen A. Zorn, *Innocent Spouses, Reasonable Women and Divorce: The Gap Between Reality and the Internal Revenue Code*, 3 MICH. J. GENDER & L. 421, 459 (1996) ("The 'prudent and reasonable man' first appears in the common law reports in 1856, in *Blyth v. Birmingham Waterworks Co*. He had been preceded by the 'man of ordinary prudence,' first mentioned in *Vaughan v. Menlove* in 1837. His twentieth-century English embodiment, the man on the Clapham omnibus, arrived in 1933, in *Hall v. Brooklands Auto Racing Club*, together with his suburban American counterpart, 'the man who takes the magazines at home and in the evening pushes the lawn mower in his shirt sleeves.'" (footnotes omitted)).

[134] *See* discussion *infra* Part II.A.

[135] Patrick J. Kelley & Laurel A. Wendt, *What Judges Tell Juries About Negligence: A Review of Pattern Jury Instructions*, 77 CHI.-KENT L. REV. 587 (2002).

[136] Kevin P. Tobia, *How People Judge What Is Reasonable*, 70 ALA. L. REV. 293 (2018).

[137] Kelley & Wendt, *supra* note 135.

[138] Tobia, *supra* note 136.

fies predominant reasonableness theories, which he incorporates into his experimental designs.[139] Using these sources, I set forth competing conceptions of the RPPT. I then evaluate each conception's legal compatibility using Kelley and Wendt's work, as well as other sources. For the most part, the conceptions are arranged in order of increasing complexity.

A. Average Conduct

Tobia explains that one important conception for reasonableness is the notion of averageness, which he characterizes as a statistical approach.[140] A small number of scholars have claimed that the RPPT asks interpreters to identify average[141] conduct for the activity in question and, once determined, to make a simple assessment of whether the conduct in the instant case fell above or below that line.[142] This Average Conduct Conception is most famously attributed (fairly or unfairly)[143] to Oliver Wendell Holmes, who wrote in *The Common Law*, "[W]hen men live in society, a certain average of conduct, a sacrifice of individual peculiarities going beyond a certain point, is necessary to the general welfare."[144]

---

[139] *Id.*

[140] *Id.*

[141] While not adopting the view himself, Christopher Brett Jaeger calls the notion that "reasonable behavior is average behavior" the "most basic empirical definition" of the RPPT. Christopher Brett Jaeger, *The Empirical Reasonable Person*, 72 ALA. L. REV. 887, 898 (2021). Averages or means might fail to capture the typical conduct of actors as well as medians would, such as when a distribution is skewed. *See* MELISSA HARDY & ALAN BRYMAN, HANDBOOK OF DATA ANALYSIS 32–34 (2004). The "median conduct" conception of the RPPT, however, has not achieved the prominence of the "average concept conception." For those who believe a median approach would be preferable in all or a subset of cases, the computational plusses and minuses described in this section would likely apply to medians in more or less the same way.

[142] *Cf.* Stephen G. Gilles, *The Invisible Hand Formula*, 80 VA. L. REV. 1015, 1041–42 (1994) ("But if juries do well when they imagine how the average person would have behaved, it is hard to see why they would not do even better if instructed to imagine how the average person would have behaved if the risks involved had been to his or her own person or property. The simple single-owner heuristic would eliminate the potential gap between average behavior under a negligence rule and optimal behavior, and would do so without confusing jargon.").

[143] Later in the same text, Holmes changes the articulation, indicating that the standard is what an average man would do. OLIVER WENDELL HOLMES, JR., THE COMMON LAW 51 (1881) ("The reconciliation of the doctrine that liability is founded on blameworthiness with the existence of liability where the party is not to blame, will be worked out more fully in the next Lecture. It is found in the conception of the average man, the man of ordinary intelligence and reasonable prudence. Liability is said to arise out of such conduct as would be blameworthy in him."). This is a different test; the conduct of the average person might be the same as average conduct, but it might not. An average person might only rarely operate complex machinery of a certain type, and they would presumably be far less skilled than the average user of that machine. If above average people are the ordinary users, then the average conduct might be considerably better than the conduct of the average person. In this section, however, I will be focusing on average conduct.

[144] *Id.* at 108.

1. Background Data and Reliability

The Average Conduct Conception has two virtues from a computational standpoint.

First, it paves the way to turn the RPPT into a pure rule. On this conception, the RPPT could be restated thusly: the test for breach is whether a person's conduct fell below average conduct for the circumstances under consideration. If there are robust statistical measures in those conditions, then the RPPT becomes a bright-line test of whether the measured conduct in this case violates a numerical minimum or maximum. From that point, the test's operation would require simple logical deduction, and even primitive computer programming would be sufficient. Imagine something like: "If defendant's conduct X is greater than average Y, then liability attaches." Or in cases where the conduct that played a causal role in the injury is binary (such as whether a person did or did not take a discrete precaution): "If a majority of people perform X and defendant did not perform X, then liability attaches." Assuming that sufficient data is available to determine reliable averages (or majority) conduct in comparable settings, there would be little to no risk of processing error, thereby eliminating an important aspect of reliability costs.

Second, there are significant contexts in which there may already be sufficient background data to arrive at robust statistical measures of average conduct. Take auto accidents, which Nora Freeman Engstrom called "the 800-pound gorilla of the tort liability system" because they account "for more than half of all trials, nearly two-thirds of all injury claims, and three-quarters of all damage payouts."[145] There are already enormous stores of data regarding traffic behavior and driving (known as "telematics"), much of it possessed by companies that own and license mapping software, such as Google, Facebook, and Foursquare,[146] or by automobile companies.[147] Indeed, for the last several years, automobiles have been outfitted with hundreds of sensors, which are linked to a persistent internet connection.[148] Adding to that, autonomous auto-

[145] Nora Freeman Engstrom, *When Cars Crash: The Automobile's Tort Law Legacy*, 53 WAKE FOREST L. REV. 293, 295 (2018).

[146] *See, e.g.*, Paris Martineau, *You May Have Forgotten Foursquare, but It Didn't Forget You*, WIRED (Mar. 8, 2019), https://www.wired.com/story/you-may-have-forgotten-foursquare-it-didnt-forget-you/ [https://perma.cc/Z42F-J4V3].

[147] Geoffrey A. Fowler, *What Does Your Car Know About You? We Hacked a Chevy to Find Out*, WASH. POST (Dec. 17, 2019), https://www.washingtonpost.com/technology/2019/12/17/what-does-your-car-know-about-you-we-hacked-chevy-find-out/ [https://perma.cc/Q2KV-A5DP] ("On a recent drive, a 2017 Chevrolet collected my precise location. It stored my phone's ID and the people I called. It judged my acceleration and braking style, beaming back reports to its maker General Motors over an always-on Internet connection. Cars have become the most sophisticated computers many of us own, filled with hundreds of sensors.").

[148] *Id.*

mobiles collect information, not only about themselves, but other drivers as well, at the rate of over three terabytes per hour for each vehicle.[149]

According to the leading market research firm, International Data Corporation (IDC), "[t]his data creation will continue to increase as vehicle-to-vehicle communication becomes commonplace, and as machine learning and AI continuously update pattern recognition integrated into vehicles' intelligent driving algorithms."[150]

If accessed, existing driving data could theoretically enable us to arrive at reliable statistics regarding average conduct on myriad driving behaviors,[151] such as speed, braking, turning, tire pressure, use of directional or headlights, and even movement. Environmental conditions, such as precipitation, wind, visibility, and temperature, are similarly tracked and stored by precise location.[152] Even if sufficient tracking occurs, however, there exist bottlenecks of network speed and data processing that must be surmounted.[153] As more people are tracked and as databases expand, it will likely be harder to upload data with sufficient speed and, thereafter, to access that information quickly. It is possible, though, that processing speed can keep pace with an increase in data scale.[154]

---

[149] DAVID REINSEL ET AL., INT'L DATA CORP., DATA AGE 2025: THE DIGITIZATION OF THE WORLD FROM EDGE TO CORE, https://downloads.snapaddy.com/external/seagate-data-age-idc-report-final.pdf [https://perma.cc/ED3B-3YGJ] (May 2020).

[150] *Id.* at 7.

[151] *See generally* Xiaoguang Wang et al., *Abstract: Built Environment and Driving Outcomes: The Case for an Integrated GIS/GPS Approach*, 5 INT'L J. APPLIED GEOSPATIAL RSCH. 11 (2014) ("This study demonstrates a segment-based approach to integrate GIS and GPS data to address questions about the connections between the built environment and travel behaviors. . . . This demonstration showed that fusing GPS and GPS data provides spatial intelligence which can be used to address planning, traffic safety, and transportation related issues.").

[152] *See, e.g.*, *About Our Data*, WEATHER UNDERGROUND, https://www.wunderground.com/about/data [https://perma.cc/3YHL-E35R] ("U.S. current conditions data comes from 180,000+ weather stations across the country including: Almost 2,000 Automated Surface Observation System (ASOS) stations located at airports throughout the country. . . . Over 250,000 Personal Weather Stations (PWS's) . . . . Stations are put through strict quality controls and observations are updated as often as every 2.5 seconds. Over 26,000 weather stations that are part of the Meteorological Assimilation Data Ingest System (MADIS) which is managed by the National Oceanic and Atmospheric Administration (NOAA)." (emphasis omitted)).

It might be wondered whether introducing factors into the statistical analysis removes much of the simplicity of the Average Conduct Conception approach. Probably not. So long as there is sufficient background data to show the statistical impact for each of those factors in combination, the processing should be able to determine averages or majorities.

[153] *See* Seth Noble, *6 Hidden Bottlenecks in Cloud Data Migration*, INFOWORLD (Apr. 18, 2018), https://www.infoworld.com/article/3268954/6-hidden-bottlenecks-in-cloud-data-migration.html [https://perma.cc/P2NN-EM9K].

[154] *See* Aaron Tan, *A Peek into the Future of Storage*, COMPUTERWEEKLY.COM (Sept. 15, 2020), https://www.computerweekly.com/feature/A-peek-into-the-future-of-storage [https://perma.cc/UX8H-RH4N].

To some readers, this analysis will seem too complex for machines to handle. For example, they might wonder how the machine could satisfactorily determine that the averages it derives come from comparable circumstances. They should remember, however, that the way in which a negligence case proceeds should lighten the load somewhat. The jury makes its breach determination only after the parties have selected and submitted evidence to it and, thereafter, the jury has found the relevant facts.[155] Moreover, the moving party will put forth a theory regarding the act or omission that served as the actual and proximate cause of the breach.[156] Consequently, the breach analysis is significantly narrowed, now with a much smaller number of variables under consideration. Of course, the adversary system incentivizes litigants to move the battle to the most favorable terrain, and an automated approach to the Average Conduct Conception might simply shift the grunt work toward a battle of the expert witnesses over which variables ought to matter. But it is at least conceivable that judicial gatekeeping of scientific evidence could drastically limit the role of variable selection in the breach determination itself.[157]

Other readers might have the inverse concern, wondering why complex machine processing might be necessary at all when most streets are already subject to bright-line speed limits. If we seek to expedite and lower costs, why not use those instead? Although I do not deny that this might be an option, courts have been loath to let violation or observation of speed limits, alone,

---

[155] *See, e.g.*, Douglass v. Irvin, 531 N.E.2d 1214, 1218 n.1 (Ind. Ct. App. 1988), *vacated*, 549 N.E.2d 368 (Ind. 1990) ("Should a duty be found to exist, the issues of whether there was a breach of duty and whether its breach was the proximate cause of injury are matters to be determined by the fact finder after weighing the evidence adduced at trial."); Kurz v. Kuhn, 223 N.W. 412, 413 (Wis. 1929) ("The question, of whether or not the defendant was negligent, can be determined only by the application of some standard of care to the conduct of the defendant Kuhn, as disclosed by the evidence. The jury must first find the facts, and to those facts apply the instructions of the court as to the law.").

[156] *See, e.g.*, Rahaman v. J.C. Penney Corp., No. N15C-07-174, 2016 Del. Super. LEXIS 258, at *4 (Del. Super. Ct. May 4, 2016) ("Pursuant to Superior Court Civil Rule 9(b), a plaintiff alleging negligence must state with particularity the circumstances constituting negligence. 'When pleading negligence, plaintiffs have to meet the heightened standard of Rule 9(b), and must specify a duty, a breach of the duty, who breached the duty, what act or failure to act caused the breach, and the party who acted.'"(quoting Rinaldi v. Iomega Corp., No. 98C-09-064, 1999 Del. Super. LEXIS 563, at *23 (Del. Super. Ct. Sept. 3, 1999))); Pope v. Hancock Cnty. Rural Elec. Membership Corp., 937 N.E.2d 1242, 1246–47 (Ind. Ct. App. 2010); Scampone v. Highland Park Care Ctr., L.L.C., 57 A.3d 582, 596 (Pa. 2012) ("Generally, to state a cause of action for negligence, a plaintiff must allege facts which establish the breach of a legally recognized duty or obligation of the defendant that is causally connected to actual damages suffered by the plaintiff.").

[157] *See generally* Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579 (1993) (announcing a rule that trial judges must act as a safeguard against admission of unreliable expert scientific evidence in cases before them). As mentioned in note 7 of this Article, one conceit of this Article is that we can assume for the purposes of analysis that other aspects of the torts case will remain human. Of course, if complete automation were desired, this hand-off to humans would be unavailable and might increase the overall costs of the Average Conduct Conception.

dictate the question of breach. Although some state courts have held that traveling in excess of a speed limit constitutes a breach under operation of negligence per se, others have not, and generally state courts refuse to accept that simply traveling under the speed limit is reasonable per se.[158] As a result, a person's driving speed is frequently just another piece of evidence under consideration during application of the RPPT: "Exceeding the prima facie speed limit . . . is not negligence per se" but "negligen[ce] in so doing depends upon all of the attendant circumstances,"[159] and "if the actual and potential hazards existing at any particular place on a highway require a speed of less than [the legal limit], then the operator should reduce the speed of his automobile at that place to whatever is reasonable and prudent under the conditions."[160] Machines can engage in statistical techniques to account for the conditions that matter to courts; they might control for environmental hazards in determining the conduct from which to derive an average.

Even if we allow vehicle cases to be handled under an Average Conduct Conception—and that is not entirely clear—there are many areas of life for which we have not yet amassed sufficient background data to arrive at a useful figure for Average Conduct. To choose but one illustrative example, there is very little data available regarding the operation of amusement park rides (an activity likely governed by the RPPT)[161] even though injuries therefrom are not unusual.[162] Suppose that a group of amusement park operators have been tracked in detail, perhaps because a single company decided to collect the data for performance reviews. The sampled data might not yield sufficiently reliable averages for the amusement park field as a whole; generally speaking, the smaller the sample size compared to the whole, the higher the margin of error and, therefore, the greater the likelihood that the plaintiff's conduct falls within

---

[158] 4 LOUIS R. FRUMER & MELVIN I. FRIEDMAN, PERSONAL INJURY: ACTIONS, DEFENSES, DAMAGES § 12.59 (Matthew Bender & Co., rev. ed. 2021) ("The violation of such a statute is negligence per se in some states and evidence of negligence in others . . . . However, compliance with the posted maximum speed limit does not prove due care." (footnotes omitted)).

[159] Hiott v. Bishop, 137 S.E.2d 780, 784 (S.C. 1964) (emphasis omitted) (citation omitted).

[160] Thomas v. Barnett, 131 S.E.2d 818, 824 (Ga. Ct. App. 1963).

[161] *See* Harlan v. Six Flags Over Ga., Inc., 699 F.2d 521, 524 (11th Cir. 1983) (holding the operation of "The Wheelie" ride at Six Flags to ordinary care in a negligence suit).

[162] *See* Paul Mose, *Wet 'n Wild: When Water Rides Should Be Subject to the Highest Duty of Care*, 63 KAN. L. REV. 787, 813–14 (2015); Gabrielle Russon, *Universal Tested Fixes to Punga Racers Water Slide as Guest Injuries Piled Up, Court Records Say*, ORLANDO SENTINEL (Aug. 5, 2020), https://www.orlandosentinel.com/business/tourism/os-bz-volcano-bay-punga-follow-20200804-dequmqxsnvaatbwmj7zoy5waby-story.html [https://perma.cc/PE73-E8H6] ("At least 115 people were hurt on Punga Racers with minor injuries such as scrapes to nosebleeds and others suffering whiplash during 13 months from Volcano's Bay grand opening in late May 2017 to mid-July 2019, earlier court documents showed.").

that margin even if it is below the stated average.[163] In such cases, it could be hard for the court to be confident that the plaintiff breached.

Moreover, when data collection is highly incomplete, there is a greater risk that averages for one context will be forced onto a context that is heterogenous, leading to external invalidity.[164] In other words, the captured average from that sampling might lack the capacity to be usefully applied to a number of different amusement park rides or to a number of different rider densities. There might be scanty data regarding the frequency with which a Tilt-A-Whirl operator looks up from a control panel to evaluate riders' wellbeing, even if there is plenty of data regarding the gaze of Pirate Ship operators. Similar problems are likely to arise when industries or contexts have actively resisted tracking.[165]

And even where there is a willingness to be tracked, some conduct is so obscure that a patchwork is the very best data collectors could hope for, at least until the proliferation of wearable or implanted data collectors. For example, in 2006 the Michigan Court of Appeals considered whether the manner in which a child mowed a lawn breached ordinary care after a person was allegedly hit by debris jettisoned by the lawnmower.[166] Although there is a chance that lawn mower paths will be tracked and coded into a dataset—there are already GPS trackers in robotic lawnmowers(!)[167]—important aspects of lawn mowing that do not involve devices are likely to arise and be material to a breach inquiry.

---

[163] WILLIAM MENDENHALL, III ET AL., INTRODUCTION TO PROBABILITY AND STATISTICS 313 (Michelle Julet ed., Brooks/Cole, Cengage Learning 14th ed. 2013) ("In a statistical estimation problem, the accuracy of the estimate is measured by the margin of error or the width of the confidence interval, both of which have a specified reliability. Since both of these measures are a function of the sample size, specifying the reliability and accuracy allows you to determine the necessary sample size." (emphasis omitted)).

[164] *See, e.g.*, Jorge Faber & Lilian Martins Fonseca, *How Sample Size Influences Research Outcomes*, DENTAL PRESS J. ORTHODONTICS, 2014 July–Aug., at 27, 27–29 ("Very small samples undermine the internal and external validity of a study."); Elizabeth Tipton et al., *Implications of Small Samples for Generalization: Adjustments and Rules of Thumb*, 41 EVALUATION REV. 472, 472–505 (2016) ("In small random samples, large differences between the sample and population can arise simply by chance and many of the statistics commonly used in generalization are a function of both sample size and the number of covariates being compared.").

[165] *Cf.* Natalie Kitroeff & Jessica Silver-Greenberg, *Airlines Refused to Collect Passenger Data That Could Aid Coronavirus Fight*, N.Y. TIMES, https://www.nytimes.com/2020/03/31/business/coronavirus-airlines-contact-tracing-cdc.html [https://perma.cc/TDA4-RR29] (Apr. 29, 2020) ("Public health officials have been pushing airlines for years to gather more traveler data, but airlines have balked, citing cost and time.").

[166] Campbell v. Kovich, 731 N.W.2d 112, 116 (Mich. Ct. App. 2006).

[167] Ashlee Clark Thompson, *Robotic Lawn Mowers Get Voice Assistant, GPS Upgrades*, CNET (Mar. 23, 2018), https://www.cnet.com/news/robotic-lawn-mowers-getting-amazon-alexa-google-assistant-gps-upgrades/ [https://perma.cc/4XLC-EAMB].

Indeed, the Michigan Court of Appeals noted that a youth lawnmower had inspected the grass beforehand in concluding that there was no breach.[168]

To be sure, a simple closure rule could be used to improve reliability. Courts could decide, for example, that conduct below the stated average yet within the margin of error does not meet the burden under RPPT. Or they could place the burden to produce a suitably reliable average (as determined by a preset battery of statistical tests) on the moving party. Such rules, however, risk allowing careless, cavalier, or uncaring conduct to go unpunished simply because it is unique. If that were to occur too often, this approach might be viewed as too unreliable by legislators.

Thus, it is likely that such a system would have holes, though those holes would likely shrink as the tracking data of day-to-day life expands.

2. Compatibility

Although the Average Conduct Conception is somewhat promising from a computational standpoint, it suffers from problems of legal compatibility. In the 1920s, Francis Bohlen, drafter of the First Restatement of Torts, famously wrote the following:

> Were the "reasonable man" identical with the average man and were the question what the average conduct of mankind under similar circumstances is, the question would be purely one of fact—of what is or exists—though involving an enormously extended inquiry as to the conduct of all other men or a great number of other men under similar circumstances. But the "reasonable man" is not the average man.[169]

There are no state model jury instructions that invoke "average" conduct, and only one describes the conduct of the "average person."[170] Instead, model jury instructions commonly state that "[t]he law does not say what a reasonably careful person using ordinary care would or would not do under such circumstances. That is for you to decide."[171] This is not to suggest that evidence regarding average conduct is not relevant to the breach determination, it is likely that most states would admit such evidence.[172]

---

[168] *See Campbell*, 731 N.W.2d at 115–16 (considering the youth's inspection in the breach analysis).

[169] FRANCIS H. BOHLEN, STUDIES IN THE LAW OF TORTS 603–04 (1926).

[170] Kelley & Wendt, *supra* note 135, at 625–80, 671 (South Carolina).

[171] *Id.* at 608 (quoting MODEL CIV. JURY INSTRUCTIONS 10.02 (MICH. SUP. CT. 1981)).

[172] *See* Cucinella v. Weston Biscuit Co., 265 P.2d 513, 515, 520–21 (Cal. 1954) (providing a split decision with majority finding no error on admission of evidence of average speed for determination of pedestrian's contributory negligence, but with dissenting judge strenuously arguing "[i]t has been held numerous times by this Court and by appellate courts of this state that the question of negligence

Nor does average conduct appear to match how jurors actually interpret the RPPT. Neal Feigenson's studies of juror behavior indicate that jurors follow a holistic approach, taking in facts, including those that do not appear to be relevant, in assessing whether an actor failed to do what others within that culture usually do.[173] These results are consistent with Tobia's experiments. He asked subjects to indicate a quantitative value in connection with over a dozen items, such as the number of weeks to wait for a criminal trial or the number of times that a person calls his or her parents in a month, varying whether they were asked to assess the "average," "ideal," or "reasonable" value for those items.[174] In one of the three experiments, subjects' values for "average" were not a significant predictor of the "reasonable" value, but other values, such as the value that subjects gave for "ideal," or a combination of "ideal" and "average" values, were significant predictors.[175] In the remaining two experiments, "average" was a significant predictor, but so were "ideal" and a combination of "ideal" and "average," and this combination had greater explanatory power than "average."[176]

Thus, an automated RPPT in the mold of the Average Conduct Conception would likely meet considerable resistance even if it were to produce reliable, representative averages. Though it has been discussed by legal luminaries, it has never achieved mainstream acceptance. For this to be an attractive choice, then, it would have to be because of its technological strengths. All things considered, the Average Conduct Conception giveth and taketh away: there are important areas in which reliable averages could be found with feasi-

---

. . . is not to be determined by what others did or did not do at the time and place under the particular facts and circumstances then and there confronting them").

[173] NEAL FEIGENSON, LEGAL BLAME: HOW JURORS THINK AND TALK ABOUT ACCIDENTS 17 (2000).

[174] Tobia, *supra* note 136, at 320–25.

[175] *Id.* at 323.

[176] *Id.* at 321, 326. I would like to note that Jaeger's recent and impressive series of experiments provided support for the notion that laypeople's understanding of reasonableness is linked to the percentage of people who perform the conduct in question. Jaeger, *supra* note 141, at 931–32 (finding in four experiments that subjects "who were told that 90% of people in the defendants' position would have avoided injuring the plaintiff found the defendant negligent 77.3% of the time" but subjects "who were told that 10% of people would have avoided injuring the plaintiff . . . found the defendant negligent only 50.5% of the time"). Importantly, Jaeger's experiments were designed to determine whether "economic" or "empirical" views of reasonableness received more support among participants in the study, and to determine whether, among empirical views, participants were more likely to adopt an "aspirational" or "average" view of reasonableness. *Id.* at 910–33. Although his results regarding economic views will be important when I consider the Learned Hand Test, they are of limited utility in this section because Jaeger understandably did not draw the distinction between average conduct, the conduct of the average person, or (in that vein) conduct that has formed a convention due, in part, to its commonness. As a consequence, his results do not indicate whether the Average Conduct Conception receives more support than its closest rival under consideration here, the Conventionalist Conception.

ble background data costs, but it suffers from gaps and lack of legal support that affect its soundness as a legal approach.

B. The Learned Hand Test

Another conception of the RPPT is economic cost-benefit testing or, more popularly, the Learned Hand Test (Hand Test).[177] Tobia characterizes this conception as a prescriptive theory because it has a normative justification for the results provided.[178] To be sure, the Hand Test has been embraced as an ideal for efficiency or welfare maximization and as a practical reform, but a small number of scholars have maintained that the Hand Test is the very best description of the RPPT in torts doctrine.[179]

The test is well-known: breach should be determined by an assessment of whether the precaution that allegedly should have been taken by the defendant was more burdensome than the utility it would have yielded from the *ex ante* perspective, and this is measured by the probability that loss would occur times the size of that loss.[180] In short, the burden must be less than the expected value of liability costs. Thus, the Hand Test is a bright-line rule.

1. Background Data and Reliability

The Hand Test is a mathematical formula—a promising starting point for computerization. Much like the Average Conduct Conception, the Hand Test poses no processing challenge once numerical values are plugged into it. So long as those values are reliable, the implementation of the formula will not introduce new statistical or computational errors. Indeed, decades-old computers could handle it with aplomb. In distinguishing the Hand Test from the Average Conduct Conception, then, the question is how gathering and processing background data regarding probabilities of injury, liability costs from injury, and the cost of taking the necessary precaution differ from doing the same with respect to average conduct.

Real-world application of the Hand Test has yielded some insight into its weaknesses. Identifying the specific values of its variables has proven difficult.[181] In rare cases, the answer under the test is obvious, making the precision

---

[177] United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947).

[178] Tobia, *supra* note 136, at 303.

[179] *See, e.g.*, WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW 85–86 (1987); *cf.* Alex Stein, *The Domain of Torts*, 117 COLUM. L. REV. 535 (2017) (offering a twist on the Hand Test, limiting the descriptive theory to instances in which the benefits of harm-causing activity are public).

[180] *See Carroll Towing Co.*, 159 F.2d at 173.

[181] Peter Z. Grossman et al., *Uncertainty, Insurance and the Learned Hand Formula*, 5 L., PROBABILITY & RISK 1, 2–3 (2006) ("Before the fact of an accident, individuals often do not know, even

of the individual variables less important.[182] Nevertheless, as Judge Richard A. Posner stated four decades ago in *McCarty v. Pheasant Run, Inc.*, ordinary cases do not lend themselves to clear results and, thus, face problems of obscurity:

> Ordinarily . . . the parties do not give the jury the information required to quantify the variables that the Hand Formula picks out as relevant. That is why the formula has greater analytic than operational significance. Conceptual as well as practical difficulties in monetizing personal injuries may continue to frustrate efforts to measure expected accident costs with the precision that is possible, in principle at least, in measuring the other side of the equation—the cost or burden of precaution.[183]

The analysis of probability in the formula is essentially an actuarial assessment of the likelihood of events, such as the one suffered by the plaintiff (and those that might be avoided with the measure) and the affixation of price to the precaution in question.[184] Assuming that the implementation of the automated Hand Test would include arriving at figures for those variables, the computer program would require a tremendous amount of background data to perform that task for each and every situation.

Perhaps existing technology is up to the challenge. There are over twenty-five thousand actuaries in the United States alone.[185] They hold a tremendous amount of information, perhaps enough to power the Hand Test if courts were somehow granted access to it.[186] Over a decade ago, before machine learning gained a foothold in the workplace, scholars Peter Grossman, Reed Cearley, and Daniel Cole asserted that actuarial work in the insurance industry "provides much of the information that courts need to apply the marginal Learned

---

within a rough approximation, the probability that they will have an accident. Nor do they know the likely harm should an accident occur. That is to say, they do not possess the information the Learned Hand formula requires them to possess in order to perform the requisite *ex ante* calculations.").

[182] *See, e.g.*, R.I. Hosp. Tr. Nat'l Bank v. Zapata Corp., 848 F.2d 291, 295 (1st Cir. 1988) (demonstrating an instance where the measure taken would have yielded no protection from loss at all).

[183] 826 F.2d 1554, 1557 (7th Cir. 1987).

[184] *See* Grossman et al., *supra* note 181, at 2–3.

[185] *See Occupational Outlook Handbook: Actuaries,* U.S. BUREAU OF LAB. STAT., https://www.bls.gov/ooh/math/actuaries.htm [https://perma.cc/Z96A-ZLXK] (Apr. 9, 2021).

[186] This would not be easy. Most actuaries are employed by insurance companies, which value actuarial work so highly that they consider it a trade secret. *Id.*; Insurent Agency Corp. v. Hanover Ins. Co., No. 16-CV-3076, 2020 U.S. Dist. LEXIS 2565, at *24 (S.D.N.Y. Jan. 8, 2020) (holding that actuarial data are trade secrets); Xavian Ins. Co. v. Marsh & McLennan Cos., No. 18cv8273, 2019 U.S. Dist. LEXIS 65067, at *2 (S.D.N.Y. Apr. 16. 2019) (same); Goshawk Dedicated Ltd. v. Am. Viatical Servs., L.L.C., No. 05-CV-2343, 2007 U.S. Dist. LEXIS 82210, at *3–4 (N.D. Ga. Nov. 5, 2007) (same); Destiny Health, Inc. v. Conn. Gen. Life Ins. Co., 39 N.E.3d 275, 280 (Ill. App. Ct. 2015) (same).

Hand formula in negligence cases."[187] Today, machine learning techniques have been adopted by the insurance industry to improve the accuracy of actuarial work.[188] Although that industry continues to debate the utility and value of machine learning approaches (which can use brute force processing and big data[189] in unsupervised learning without the burdens of traditional inference assumptions) compared to bedrock actuarial techniques (which are bound by traditional inference assumptions),[190] there is evidence that the former will win the day eventually, if it has not already.[191]

In that vein, legal scholars have expressed optimism that big data can be tapped to create more accurate norms regarding a machine-generated standard of optimal care. Omri Ben-Shahar and Ariel Porat believe that machines using big data can overcome information costs in creating a personalized standard of care based on a variety of individualized factors.[192] Although they do not consider whether big data could power an objective rule, such as the Hand Test,

---

[187] Grossman et al., *supra* note 181, at 1. The authors go on to say the following about the capacity to update information bearing upon the risk of liability:

> Further, these processed [insurance] claims provide the industry with a continual stream of information about the probability curve surrounding the liability variable, *L*, of the Learned Hand formula, which permits continual updating of [the expected value of the harm]. That information, which is not possessed by the individual injurers and victims because of diseconomies of scale, increases the general foreseeability of the harm. In other words, the insurance industry has the necessary information to reasonably anticipate the potential harm for a selected type of accident.

Id. at 16.

[188] *See* Alexander Cherry, *Artificial Intelligence and the Actuary of the Future*, REUTERS EVENTS, https://www.reutersevents.com/insurance/analytics/artificial-intelligence-and-actuary-future [https://perma.cc/3FSK-3SSB] ("Cast your eye down your newsfeed and one thing is clear: Artificial Intelligence (AI) has arrived in insurance. From natural-language processing and chatbots for claims resolution to big data and algorithms in the actuarial back office, it seems there is barely a cog in the venerable insurance machine that AI will leave untouched."). *See generally* 1 BIG DATA FOR INSURANCE COMPANIES (Marine Corlosquet-Habart & Jacques Janssen eds., 2018) (illustrating the profound impact of big data and machine learning upon the insurance industry).

[189] I follow the technologist definition, "Big data is high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation." *Gartner Glossary: Big Data*, GARTNER, http://www.gartner.com/it-glossary/big-data/ [https://perma.cc/58Q8-EPQU].

[190] Warren Franklin, *Machine Learning Algorithm vs. Actuarial Science . . . Who Will Win?*, TOWARDS DATA SCI. (Aug. 5, 2016), https://towardsdatascience.com/machine-learning-algorithm-vs-actuarial-science-who-will-win-b203f31145ce [https://perma.cc/763M-TFCP].

[191] *See* Guojun Gan & Emiliano A. Valdez, *An Empirical Comparison of Some Experimental Designs for the Valuation of Large Variable Annuity Portfolios*, 4 DEPENDENCE MODELING (SPECIAL ISSUE: RECENT DEVS. QUANTITATIVE RISK MGMT.) 382, 382–400 (2016) (arguing that machine learning provides speed and accuracy when evaluating multiple annuities in a portfolio simultaneously compared to actuarial techniques).

[192] Omri Ben-Shahar & Ariel Porat, *Personalizing Negligence Law*, 91 N.Y.U. L. REV. 627 (2016).

they assert that it could power an even more ambitious subjective regime that relies on similar variables, such as one that "seeks to distinguish people according to their tendencies to create risks and their capabilities to prevent them."[193] In this regard, they are confident that machines can identify particularized risks of harm. In another article, Lior Strahilevitz joins Porat in addressing big data's capacity to generate particularized assessments of price, claiming that "[f]irms could use what they know about their customers to provide them with personalized default terms and prices in contracts that are determined at the time a contract is entered into and which any customer could see before she executes the contract."[194]

A significant portion of the useful[195] information stored under the umbrella of big data is the collection of data regarding purchases and purchase price. There are over 112 million Americans who are users of Amazon's Prime subscription service alone, which is more than half the number of adults in the country, and Prime members have spent an average of fourteen hundred dollars on Amazon per year.[196] There can be little doubt that the acceleration of online purchasing during the COVID-19 pandemic has only inflated these numbers.[197] Even five years ago, the Obama Administration raised red flags about harnessing big data to increase price discrimination at the individual level, warning, "Big data allows companies to collect more information about customers and use it to create new kinds of measures, raising the likelihood differential pricing will become more common and more personalized over time."[198] The same AI that is used to predict consumer behavior and optimize pricing could be used to identify costs for taking a protective measure at the critical moment in time. Although most of these databases are inaccessible to the gov-

---

[193] *Id.* at 679.

[194] *See* Ariel Porat & Lior Jacob Strahilevitz, *Personalizing Default Rules and Disclosure with Big Data*, 112 MICH. L. REV. 1417, 1440 (2014).

[195] Most stored data is redundant or otherwise useless. *Veritas Global Databerg Report Finds 85% of Stored Data Is Either Dark or Redundant, Obsolete, or Trivial (ROT)*, VERITAS (Mar. 15, 2016), https://www.veritas.com/news-releases/2016-03-15-veritas-global-databerg-report-finds-85-percent-of-stored-data [https://perma.cc/G32E-FGF4].

[196] Stephanie Chevalier, *Average Spending of Amazon Prime and Non-Prime Members 2015–2 2019*, STATISTA (July 7, 2021), https://www.statista.com/statistics/304938/amazon-prime-and-non-prime-members-average-sales-spend/ [https://perma.cc/7WBM-4ASB]; Don Reisinger, *Amazon Prime's Numbers (and Influence) Continue to Grow*, FORTUNE (Jan. 16, 2020), https://fortune.com/2020/01/16/amazon-prime-subscriptions/ [https://perma.cc/Q6FZ-CB4R].

[197] Matt McFarland, *Amazon Thrived During the Pandemic. These Drivers Say They Paid the Price*, CNN (June 3, 2021), https://www.cnn.com/2021/06/03/tech/amazon-dsp-delivery/index.html [https://perma.cc/E676-NXM9] (reporting that "Amazon purchases ballooned in 2020—it delivered 2.3 billion more of its own packages in 2020 compared to 2019").

[198] Jason Furman & Tim Simcoe, *The Economics of Big Data and Differential Pricing*, THE WHITE HOUSE PRESIDENT BARACK OBAMA: BLOG (Feb. 6, 2015), https://obamawhitehouse.archives.gov/blog/2015/02/06/economics-big-data-and-differential-pricing [https://perma.cc/R7A2-BCUZ].

ernment, it is easy to imagine that courts might someday gain access, whether through discovery, party submission, or other means.

There will be both technical and social challenges, however. As to the former, the same problems with data-related bottlenecking under the Average Conduct Conception exist here as well.[199] As to the latter, aspects of the Hand Test are difficult to assess in dollar terms regardless of the quality or quantity of data, such as the value of human life.[200]

Having discussed existing resources, we are ready to return to the question of how the Hand Test compares to the Average Conduct Conception.

Considering background knowledge, despite having more variables, it is not clear that the Hand Test conception is any better or worse than the Average Conduct Conception. There exist impressive data collections for both, though they are largely in the possession of private companies. And there are likely to be areas for which little data is collected until tracking technology becomes more widespread.

The bigger distinction between the two approaches is reliability. Recall that when conduct is infrequent, the statistical reliability of an average for that conduct tends to go down.[201] Under the Hand Test, by contrast, the uncommonness of an accident is built into the analysis: infrequency is represented in lower values for the probability variable.[202] Thus, the negative impact of infrequency on reliability is potentially mitigated—if something is infrequent, the impact of price reliability goes down as well because the formula multiplies that price by a much smaller number.[203] It might be objected that infrequency

---

[199] *See* discussion *supra* Part II.A.

[200] Kenneth W. Simons, *The Hand Formula in the Draft Restatement (Third) of Torts: Encompassing Fairness as Well as Efficiency Values*, 54 VAND. L. REV. 901, 924–25 (2001) ("If too low a value is placed upon human life, the negligence standard will be too forgiving. If too high a value is placed upon human life, the negligence standard will be too stringent. Juries and judges, however, are ill-equipped to make these controversial evaluations, and the process of making these judgments on a case-by-case basis threatens to cause similar fact patterns to be treated quite inconsistently." (footnote omitted)).

[201] *See* discussion *supra* Part II.A.1.

[202] *See* United States v. Carroll Towing Co., 159 F.2d 169, 173 (2d Cir. 1947).

[203] Comparing the Average Conduct Conception and the Hand Test is a bit tricky because they do not measure identical things, but if we can imagine an example in which they do, it might be helpful to illustrate the point. Imagine that a case involves the failure to pay for a safeguard, but the safeguard is seldom taken. For example, if something is infrequent (suppose it only happens to one out of 100,000 people engaged in an activity), that infrequency might make it harder to assess the cost of the safeguard that is under consideration reliably (suppose it introduces a margin of error of plus or minus $10,000). Under the Hand Test, the range of error introduced is divided by the denominator of probability—and infrequent events have larger denominators (in the hypothetical, $10,000/100,000 which is, if kept in terms of dollars, exactly ten cents)—whereas the Average Conduct Conception must accept the margin of error as is (in the hypothetical, $10,000). Perhaps it is better to say it this way: when a safeguard is particularly rare, the deck is generally stacked in favor of the defendant such that the margin of error is less likely to lead to error on the breach determination.

is still the problem, particularly because the Average Conduct Conception is not the only rival here. Infrequent events can, at one time, seem low risk due to their rarity, but, at a later date when they become more frequent, be properly understood as high risk. There might be a lag before the true riskiness can be incorporated into the test. For example, the sudden popularity of the hoverboard during the Christmas season of 2015 gave rise to so many accidents that it effectively doubled the annual number of injuries associated with devices of that sort.[204] Importantly, however, the Hand Test is about available evidence from the *ex ante* perspective,[205] so the possibility of a change in frequency revealing that people have mistakenly evaluated risk is of no import in the analysis.

To be clear, this does not eliminate a weakness that is also suffered by the Average Conduct Conception—namely, that some areas of life are not tracked at all, so an automated assessment of risk or of the cost of the precaution will have a high risk of error or non-answer. As with the Average Conduct Conception, however, these areas of life will likely shrink over time.

2. Compatibility

Turning to legal compatibility, there can be little question that the Hand Test is a marginalized conception of the RPPT under law, despite being somewhat more popular than the Average Conduct Conception.[206] Although aspects of the test have become enshrined as requirements for design or warning de-

---

[204] *See* Brianna L. Siracuse et al., *Hoverboards: A New Cause of Pediatric Morbidity*, 48 INJURY: INT'L J. CARE INJURED 1110, 1112 (2017) (showing an increase of 400% for injuries associated with powered skateboards/scooters between November and December of 2015, as well as a similar increase compared to that same period in the prior four years, as well as an average annual increase of 200% compared to the preceding four years).

[205] *See* Guido Calabresi, Remarks of Hon. Guido Calabresi, Address at Tort Law in the Shadow of Agency Preemption, Symposium of the New York University Annual Survey of American Law (Feb. 27, 2009), *in* 65 N.Y.U. ANN. SURV. AM. L. 435, 440 (2010) ("We all know that supposedly in torts, under the old Learned Hand test, liability for negligence is based on what a reasonable person should have known at the time the accident-causing event took place."); Thomas C. Galligan, Jr., *Strict Liability in Action: The Truncated Learned Hand Formula*, 52 LA. L. REV. 323, 325 (1991) ("One is negligent, per the Hand formula, when B < P x L. B is the burden of avoiding the risk; P is the *ex ante* probability the risk will materialize in injury; and, L is the gravity of the risk if it materializes in injury.").

[206] *See* Gilles, *supra* note 142, at 1015–18 ("The initial puzzle lies in the gap between the authoritative blackletter status of the Hand Formula and the standard instructions given to juries in negligence cases. . . . Yet, rather than telling juries to balance the costs and benefits of greater care, courts ordinarily instruct them to determine whether the actor behaved as a 'reasonably prudent person' would have under the circumstances. Even on appeal, many courts make surprisingly little use of cost-benefit analysis in reviewing negligence cases. . . . Some scholars claim that these practices demonstrate that the actual meaning of negligence in American law is defined by a reasonable person standard that marginalizes or even supplants the Hand Formula." (footnote omitted)).

fects in products liability cases in many states,[207] courts generally do not ask jurors to use it in cases of negligence.[208] The exception is Louisiana, which embraced the test in *Dobson v. Louisiana Power & Light Co*.[209] A handful of other states embrace the doctrine of changing expected legal duties according to the amount of foreseeable risk, but they do not set out a cost-benefit test.[210] That is, there are higher duties for riskier activities in those jurisdictions, but there is no requirement that the Hand Test be used.[211]

Moreover, there is little evidence that the Hand Test is how jurors understand the RPPT. Twenty years ago, Kip Viscusi performed a survey experiment of 500 juror-eligible Americans and found they did not do well in following the Hand Test when making negligence judgments.[212] He concluded that "[t]he underlying theme throughout these results is . . . that jurors make errors that are inconsistent with the usual law and economics principles [like the Hand Test]" and frequently pay little attention to a company's assessment of risk level or cost per life saved.[213] More recently, Christopher Jaeger performed a series of experiments in which he sought to investigate how much layperson negligence determinations were swayed by either cost-benefit analysis or by statistics regarding the percentage of people who would have undertaken a safety measure to avoid accident.[214] Each time, higher percentages of accident avoidance correlated significantly with decreases in the likelihood that laypeople would find negligence but no such correlation was found with increased results under a Hand Test analysis.[215] Jaeger explained, "Participants did not seem to care whether precautions were cost-justified under the Hand Formula. Participants found the defendant negligent 66.5% of the time when the precautions were cost-justified, and 63.3% of the time when the precautions were not."[216]

Judges and scholars have attributed states' unwillingness to adopt the Hand Test to a number of things: (1) that it is too difficult for fact-finders to

---

[207] *See* 2 CARY STEWART SKLAREN, PRODUCTS LIABILITY § 11.03 (Matthew Bender & Co., rev. ed. 2021) (discussing the acceptance of risk-utility and linking it to negligence standard).

[208] *See* Steven Hetcher, *The Jury's Out: Social Norms' Misunderstood Role in Negligence Law*, 91 GEO. L.J. 633, 648–49 (2003) (discussing how the fairly frequent mention of the Hand Test by appellate courts is not nearly as important as the near absence of its mention by trial courts).

[209] 567 So. 2d 569 (La. 1990).

[210] Kelley & Wendt, *supra* note 135, at 607 (Pennsylvania and Utah).

[211] *Id.*

[212] *See* W. Kip Viscusi, *Jurors, Judges, and the Mistreatment of Risk by the Courts*, 30 J. LEGAL STUD. 107, 135 (2001).

[213] *Id.*

[214] Jaeger, *supra* note 141, at 931–32.

[215] *Id.* at 932.

[216] *Id.* (footnote omitted).

assess price, harm, and probability;[217] (2) that jurors would find cost-benefit assessment distasteful and engage in jury nullification;[218] (3) that values are fatally difficult to quantify or are incommensurate;[219] and (4) that economic rationality is not a demand of tort law.[220] It would serve to briefly address how automation of the RPPT might respond to those concerns.

The first concern is not particularly weighty in this analysis so long as machines can perform the test more reliably. Assuming that background data are sufficient to give reliable values, there is little doubt that a math-powered machine will outperform a juror on that score.

The second concern, too, is to be accorded little weight here; one virtue of machines is their incapacity to violate their own programming, so there is no risk of nullification. That will do nothing to change *public* reaction to the distastefulness of its application. For instance, the test can impose lighter burdens on actors who put the lives of poorer people at risk because injuries to that population can be compensated for lower cost, making the liability cost variable smaller.[221] To be sure, this is a problem that exists elsewhere in the torts system—it owes much to the manner in which assessments for damages

---

[217] *See* Jonathan J. Koehler, *Train Our Jurors*, *in* HEURISTICS AND THE LAW 303, 305, 310–11 (G. Gigerenzer & C. Engel eds., 2006) (discussing juror problems with understanding instructions on assessments of probability); Gilles, *supra* note 142; D.H. Kaye & Jonathan J. Koehler, *Can Jurors Understand Probabilistic Evidence?*, 154 J. ROYAL STAT. SOC'Y 75, 76–80 (1991); B. Michael Dann et al., Testing the Effects of Selected Jury Trial Innovations on Juror Comprehension of Contested mtDNA Evidence 4 (Dec. 30, 2004) (unpublished report) (https://www.ojp.gov/pdffiles1/nij/grants/211000.pdf [https://perma.cc/Z9T9-TLFC]) (highlighting jury errors on assessing probability evidence in DNA cases across six studies).

[218] *See* Steven Hetcher, *Non-Utilitarian Negligence Norms and the Reasonable Person Standard*, 54 VAND. L. REV. 863, 876 (2001).

[219] *See* A. MITCHELL POLINSKY, AN INTRODUCTION TO LAW AND ECONOMICS 126 (1983) ("[C]ritics [of economic analysis] have suggested that because of the difficulty of quantifying certain costs or benefits, the economic analyst will tend to substitute his own subjective values for these items, and therefore the analysis may simply 'confirm' his prior beliefs.").

[220] *See generally* Gregory C. Keating, *Reasonableness and Rationality in Negligence Theory*, 48 STAN. L. REV. 311 (1996) (claiming that social contract theory fits better with negligence doctrine than does economic rationality).

[221] Tsachi Keren-Paz, *Egalitarianism as Justification: Why and How Should Egalitarian Considerations Reshape the Standard of Care in Negligence Law?*, 4 THEORETICAL INQUIRIES L. 275, 316–17 (2003) ("It should be noted that there are two separate reasons why the Hand formula as traditionally applied by law and economics leads to regressive results. The first reason is that since the expected loss of the rich is higher, the same risk might be deemed negligent if it targets the rich, but non-negligent if it targets the poor. In such a case, the rational potential tortfeasor would risk the poor; he would be deemed non-negligent; and he therefore would not need to compensate the poor for the harm resulting from his action. This results in a systematic regressive redistribution of wealth. The second reason for the regressive result derives from compensation rules rather than liability rules. Once liability is established (be it based on negligence or strict liability), the rich victim receives more compensation than the poor victim does, due to the principle of full compensation.").

awards are made[222]—but the Hand Test builds this feature into the threshold for breach, exacerbating the problem.[223]

The third concern is unlike the first two in an important respect. Although the rigidity of an automated Hand Test is fairly viewed as an asset with respect to the first two arguments, the inverse is true here: the mechanical nature of the test increases the risk that it will seek to commensurate that which many believe to be incommensurable, leading to controversial outcomes. Learned Hand, himself, doubted that the formula could be used mechanically. In *Conway v. O'Brien*, he described the likelihood of injury as "practically not susceptible of any quantitative estimate" and the cost of the measure taken and of liability exposure as "generally not [susceptible of any quantitative estimate], even theoretically" because "a solution always involves some preference, or choice between incommensurables."[224] It is true that Hand could scarcely have imagined the amount of data and the speed of data processing that exists today, but that does not remove the problem with respect to certain classes of injury, such as the previously discussed example of loss of human life.[225]

The fourth concern is not eased by automation either. Although scholars, particularly those who self-identify as operating in the field of law and economics, argue that the Hand Test *ought* to be the way to determine breach, few scholars maintain that it *is* the way juries determine breach.[226]

It is fair to say, then, that although automation could make the Hand Test more viable, it will be difficult to convince lawmakers that it is an extant directive in American tort law, it might bring about unpopular outcomes, and it might run counter to more dominant rival conceptions.

C. Conventionalism

Another candidate conception is the Conventionalist Conception.[227] Thereunder, the test for determining satisfaction of the RPPT is grounded in emergent conventions of conduct in the community or jurisdictional area.[228]

---

[222] *See* John B. Attanasio, *The Principle of Aggregate Autonomy and the Calabresian Approach to Products Liability*, 74 VA. L. REV. 677, 741–42 (1988).

[223] Keren-Paz, *supra* note 221.

[224] 111 F.2d 611, 612 (2d Cir. 1940), *rev'd* 312 U.S. 492 (1941).

[225] *See* discussion *supra* Part II.A.1.

[226] *See, e.g.*, Alexander B. Lemann, *Coercive Insurance and the Soul of Tort Law*, 105 GEO. L.J. 55, 83 (2016) ("The descriptive claim that tort law is about efficiency thus often becomes a prescriptive claim that tort law should be tweaked in some way to make it more efficient. This rhetorical move has been used for a generation." (footnote omitted)).

[227] Tobia, *supra* note 136, at 303.

[228] *See* Patrick J. Kelley, *Who Decides? Community Safety Conventions at the Heart of Tort Liability*, 38 CLEV. ST. L. REV. 315, 364 (1990) [hereinafter Kelley, *Who Decides?*] ("Early negligence law can thus be seen as a brilliant attempt to retain customary or conventional norms as the basis for

Conventionalism is a philosophical doctrine that takes on many guises, so there is a risk that my discussion of Conventionalist Conceptions of the RPPT might be confused for theories regarding fundamental matters such as legality[229] or legal objectivity[230] that share the same label and some of the same features. In the context of RPPT, the Conventionalist Conception asserts there is a form of public agreement or consensus in the community that provides normative support for the conclusion that the duty of ordinary care was or was not breached.[231] On this account, the RRPT is a call to jurors to view the facts through the lens of the values that exist in their community, either historically, today, or some hybrid of the two.

The thorny thing about the Conventionalist Conception is that it is not simply a test that calls the interpreter to identify common conduct, so merely identifying whether there is a range of median behavior in a normal distribution is not enough. Inasmuch as it is conventionalist, the conception involves the complex task of looking for not only regularities of behavior, but a value system that views failure to conform to those regularities as worthy of criticism or punishment.[232] For example, a norm could become apparent from review of private conversation, pop culture items, and scholarship, and that norm might be a standard of criticism for certain categories of conduct. Although there continues to be philosophical debate about the rationality of supposing that conventions of this sort can create obligations or are otherwise reason-giving, [233] the resolution of that debate is of little consequence here so long as conventions are identifiable and the RPPT operates as an authoritative directive to render a decision under the applicable convention in cases of negligence.

---

tort judgments without transferring from the community to the judges the authority to define and change those norms. The system relied heavily on what we might call 'covering generalities' to refer to, but not to specify, the conventional norms. The system also used a procedure that left the ultimate judgment to the community-representing jury.").

[229] *See, e.g.*, JULES L. COLEMAN, THE PRACTICE OF PRINCIPLE: IN DEFENCE OF A PRAGMATIST APPROACH TO LEGAL THEORY 68 (2001); ANDREI MARMOR, POSITIVE LAW AND OBJECTIVE VALUES 4–5 (2001).

[230] *See, e.g.*, Owen M. Fiss, *Conventionalism*, 58 S. CAL. L. REV. 177, 191–96 (1985) (discussing a Conventionalist approach to constitutional interpretation).

[231] *See* Kelley, *Who Decides?*, *supra* note 228, at 318. ("First, there must be a social convention or coordinating norm intended to protect people like plaintiff from a particular hazard.").

[232] *Id.*; *see also* DAVID K. LEWIS, CONVENTION: A PHILOSOPHICAL STUDY 78 (1969); Richard H. McAdams, *Conventions and Norms: Philosophical Aspects*, *in* 4 INTERNATIONAL ENCYCLOPEDIA OF THE SOCIAL AND BEHAVIORAL SCIENCES 844, 844–49 (2d ed., Elsevier 2015).

[233] *See* Leslie Green, *Positivism and Conventionalism*, 12 CAN. J.L. & JURIS. 35, 35–52 (1999); Christopher P. Taggart, *How Can 'Positivism' Account for Legal Adjudicative Duty?*, 33 OXFORD J.L. STUD. 169, 179–82 (2013).

1. Background Data and Reliability

The computerization of the Conventionalist Conception poses challenges, not only with respect to the reliability of processing, but also with the collection of the background data that will inform that processing.

The good news for proponents of automation is that the Conventionalist Conception asserts that community values are social facts.[234] At least theoretically, then, the identification of the content of a community value is an act that can be done by non-evaluative processing of descriptive data. On this account, applying the RPPT does not require deep, context-sensitive moral evaluation, a task that is well beyond the capacities of today's computers.[235] The bad news is that these facts are still very challenging to find.

The formulation of an operation for the identification of community values is a difficult matter. Tort doctrine, itself, provides precious little guidance regarding protocols for convention identification,[236] but there are relevant resources in scholarship across disciplines. The most famous approach is from philosopher David Lewis, who sets forth necessary conditions for a convention to arise:

> A regularity R in the behaviour of members of a population P when they are agents in a recurrent situation S is a convention if and only if it is true that, and it is common knowledge in P that, in any instance of S among members of P,
>
> (1) everyone conforms to R;
>
> (2) everyone expects everyone else to conform to R;
>
> (3) everyone has approximately the same preferences regarding all possible combinations of actions;
>
> (4) everyone prefers that everyone conform to R, on condition that at least all but one conform to R;
>
> (5) everyone would prefer that everyone conform to R′, on condition that at least all but one conform to R′,

---

[234] This is not to suggest that all theories have this character. Indeed, scholars debate whether the standard calls for descriptive analysis or normative judgment. *See* Zipursky, *supra* note 111, at 2133.

[235] *See* discussion *infra* Part II.D.1.

[236] Torts doctrine does elaborate upon a close cousin of conventions—namely, customs. "A custom [in tort law] is a widespread and, for some courts, nearly universal practice[,]" according to torts scholar Kenneth Abraham. Kenneth S. Abraham, *Custom, Noncustomary Practice, and Negligence*, 109 COLUM. L. REV. 1784, 1788 (2009) (citing DAN B. DOBBS, THE LAW OF TORTS § 163, at 394 (2000)). This concept, which is little more than simple numerousness, is thinner than the concept of conventionalism that I will describe in this section, but I do discuss custom where relevant.

> where R′ is some possible regularity in the behaviour of members of P in S, such that no one in any instance of S among members of P could conform both to R′ and to R.[237]

Classic Lewis Conventionalism, although systematic, is overengineered for the RPPT. The first condition stands out for requiring that *everyone* conform. Commonness is an important aspect of the Conventionalist account of RPPT, but its proponents do not assert that the formation of a convention demands *complete* conformity. For instance, it is well-settled that evidence of what is usual or customary is admissible on the issue of negligence.[238] Torts scholar Kenneth Abraham, explained, "Although the courts rarely engage in an express headcount, discussions of the custom rule seem to me to presuppose that a practice must be followed by at least a majority of relevant actors in order to qualify as a custom."[239] For their part, philosophers have departed from Lewis on this score, claiming that he has set the bar too high.[240] Indeed, perfect conformity raises the question of why there would be a need to set up legal enforcement at all. In later decades, Lewis heard his critics and softened his approach, claiming that conventions must merely hold for "almost all" members of a "subpopulation."[241] Thus, it is fair to say that Conventionalist accounts do not generally require that conventions exist countrywide or across the entire population of a jurisdiction—the scope need only be the relevant community.[242]

The second aspect of Lewis's approach that seems incongruous with the formation of an RPPT convention is his insistence on people in the population accepting the convention on a set of strategic conditions. It should be obvious from conditions four and five that Lewis presumes a game-theoretic basis for conventions, wherein they arise out of rational strategy and result in stability through mutual-self-interested equilibrium.[243] In other words, Lewis's conditions serve both to describe the attributes of conventions and to ensure that those who create a convention have done so sensibly and strategically in coor-

---

[237] This concise articulation comes from JP Smit, *An Alternative to the Lewisian View of Conventions*, 46 STELLENBOSCH PAPERS LINGUISTICS 141, 141–42 (2016) (emphasis omitted) (citation omitted) (quoting LEWIS, *supra* note 232, at 76).

[238] *E.g.*, Ludman v. Davenport Assumption High Sch., 895 N.W.2d 902, 917 (Iowa 2017).

[239] Abraham, *supra* note 236, at 1788 n.9.

[240] Smit, *supra* note 237.

[241] David Lewis, *Convention: Reply to Jamieson*, 6 CANADIAN J. PHIL. 113, 116 (1976).

[242] Lorena Ramírez-Ludeña & Josep M. Vilajosana, *Introduction* to LEGAL CONVENTIONALISM 1–5 (Lorena Ramírez-Ludeña & Josep M. Vilajosana eds., 2019) (using Lewis as a point of departure in setting the scope for convention at the level of a given community and requiring the presence of recurring behavior, belief that existence of this behavior is a reason to follow it, and a set of expectations of these circumstances).

[243] *See* Robin P. Cubitt & Robert Sugden, *Common Knowledge, Salience and Convention: A Reconstruction of David Lewis' Game Theory*, 19 ECON. & PHIL. 175, 175–210 (2003).

dination. Philosophers have criticized Lewis's notion that conventions must arise from consciousness of a game or with an understanding of how the convention came to arise, [244] and it is difficult to see how these conditions would be necessary for interpreting the content of the RPPT. Considering the test, itself, breach in ordinary negligence does not turn on the reasons that someone conforms to reasonableness; rather, it turns on whether that person *acted* reasonably.[245] If the test, itself, does not demand particular motivations to act reasonably, we have little reason to suspect that the convention that undergirds our understanding of reasonableness demands it.

To be clear, this does not mean that we should dispense with the requirement that members of a community recognize there is a convention and conform to it. For our purposes, a convention should still require that people regularly perform an action in a context where there is widespread acceptance that failing to perform the action is a justification for criticism and does not accord with people's preferences.[246] These conditions cannot be excised from the test because they serve as the justification for punishment under law: the reasoning is that because the defendant was a member of that community, they should have been aware that her conduct would be a violation of a prevailing convention.

Returning to the question of what machine operations will be necessary to effectuate a Conventionalist Conception of the RPPT, we have boiled it down to three steps: the machine must (1) determine the relevant community; (2) determine that a majority of actors in that community take the precaution at issue; and (3) operate in an environment where a majority of the members of that community expect and prefer others to take that precaution as well.

Thus, a machine will first need a method to determine the borders of the relevant community. For humans, this is often obvious, such as when the dispute concerns activities performed exclusively by a certain occupation. But other times the lines of demarcation will be blurry. The development of reliable, machine-powered community identification has been a focal point in the

---

[244] Tyler Burge, *On Knowledge and Convention*, 84 PHIL. REV. 249, 252 (1975) ("Parties to a convention are frequently confused about the relevant ends (the social functions of their practice); they are often brought up achieving them and do not know the origin of their means . . . .").

[245] RESTATEMENT (SECOND) OF TORTS § 463 cmt. b (AM. L. INST. 1965) ("Negligence is conduct which creates an undue risk of harm to others. Contributory negligence is conduct which involves an undue risk of harm to the person who sustains it."); *id.* at ch.12, topic 4, scope note ("In order that either an act or a failure to act may be negligent, the one essential factor is that the actor realizes or should realize that the act or the failure to act involves an unreasonable risk of harm to an interest of another, which is protected against unintended invasion."); Mark P. Gergen, *Negligent Misrepresentation as Contract*, 101 CALIF. L. REV. 953, 1011 (2013) ("Negligence is agnostic about what makes conduct unreasonable.").

[246] In this regard, the work of H.L.A. Hart on the conventional basis for the Rule of Recognition is helpful. H.L.A. HART, THE CONCEPT OF LAW 94–95 (2d ed. 1994).

field of network science for decades,[247] and several techniques have emerged that provide useful results. Network science is the study of interactions (links) between entities (nodes), and the communities that emerge from their data processing are often visualized through clusters of interconnected nodes on a graph.[248] Scientists have been successful in disentangling overlapping communities or identifying borderless communities that humans cannot perceive on their own.[249] In law, scholars have employed these techniques to identify communities within the law professoriate.[250]

This is not to suggest that automated community detection is already here. Important challenges remain: there does not appear to be a scientific consensus on the definition of community; network science is at an early stage of development in modeling community change; there are limits to scalability with more complex real-world systems (and those systems are likely plentiful);[251] and researchers still generally hand-select the variables that they use when seeking to identify a network.[252] As a consequence, we are still far away from an all-purpose, self-contained community detector, despite the fact that network science has made great strides in the last two decades. Importantly, this aspect of the Conventionalist Conception, even with future innovations, will require data regarding a broad array of activities so that communities can be identified. Moreover, further innovation is needed to create reliable community

---

[247] *See* Nacim Fateh Chikhi et al., *A New Algorithm for Community Identification in Linked Data*, *in* KNOWLEDGE-BASED INTELLIGENT INFORMATION AND ENGINEERING SYSTEMS: KES 12TH INTERNATIONAL CONFERENCE PART I, at 641 (Ignac Lovrek et al. eds., 2008) ("Since [the] late nineties, identification of web communities has received much attention from researchers.").

[248] *See* FILIPPO MENCZER et al., A FIRST COURSE IN NETWORK SCIENCE 1–4 (2020).

[249] *See, e.g.*, Till Hoffmann et al., *Community Detection in Networks Without Observing Edges*, SCI. ADVANCES, Jan. 22, 2020 (detecting communities with undetectable borders using statistical modeling and discussing similar work in climate change, finance, and neuroscience).

[250] *See* Daniel Martin Katz et al., *Reproduction of Hierarchy? A Social Network Analysis of the American Law Professoriate*, 61 J. LEGAL EDUC. 76 (2011).

[251] Anna D. Broido & Aaron Clauset, *Scale-Free Networks Are Rare*, NATURE COMMC'NS, Mar. 4, 2019, at 1, 7–8, https://www.nature.com/articles/s41467-019-08746-5.pdf [https://perma.cc/NMZ6-C8T7] (finding empirical support for rarity of real-world networks that are "scale-free," meaning that they are organized similarly at small scale and at large scale); Petter Holme, *Rare and Everywhere: Perspectives on Scale-Free Networks*, NATURE COMMC'NS, Mar. 4, 2019, at 1, 1–3, https://www.nature.com/articles/s41467-019-09038-8.pdf [https://perma.cc/8KQH-NQNB] (commenting on importance on Broido and Clauset's article and concluding it highlights the debate in network science over the definition of "scale-free," between those who view them as Platonic ideals and those who view them as real-world objects).

[252] *See* Santo Fortunato, *Community Detection in Graphs*, 486 PHYSICS REPS. 75, 161 (2010) ("Everybody has his/her own idea of what a community is, and most ideas are consistent with each other, but, as long as there is still disagreement, it remains impossible to decide which algorithm does the best job and there will be no control on the creation of new methods. . . . It means designing practical examples of graphs with communities, and, in order to do that, one has to agree on the fundamental concepts of community and partition.").

identification in complex environments. Thus, even this first step brings significant background data and reliability costs. In a parallel to the Average Conduct Conception, however, the plaintiff's theory of duty and causation in the case might narrow the realm of possible communities somewhat, thereby lightening the load for machines.

Assuming that a relevant community has been identified, there remain two processing tasks under the Conventionalist Conception: identify that a majority in a relevant population (1) perform the actions in question and (2) expect and prefer others to do so as well. Because the first is largely the same as the Average Conduct Conception, I will focus on the second.

Unlike assessments discussed under alternative conceptions (e.g., average conduct, price, or probability of accident), the determination of expectations, preferences, and criticism requires the capacity to gain insight into mental states. The scope of conduct covered by ordinary negligence is wide, so this is no easy task.

The law has long presumed that humans have the capacity to determine the mental states of others,[253] and the primary mode of determination is to interpret the statements of the person examined.[254] In the foreseeable future, the internet will continue to be the biggest repository of accessible data that evinces community values in the history of mankind. But the vast majority of data on the internet is unstructured, meaning data that does not follow a readily interpretable format, such as a column-row database.[255] Most of that unstructured data consists of statements in natural language, items like: emails, text files, social media posts, websites, audio files, and video files.[256]

---

[253] Alan C. Michaels, *Acceptance: The Missing Mental State*, 71 S. CAL. L. REV. 953, 999 (1998) ("The criminal law frequently hinges liability on subtle distinctions about the actor's state of mind. Was it her purpose? Did she know? Was she reckless? Reliance on these subjective mental states for liability assumes that juries can accurately determine what the actor was actually thinking.").

[254] *Cf.* RICHARD A. EPSTEIN, THE CLASSICAL LIBERAL CONSTITUTION: THE UNCERTAIN QUEST FOR LIMITED GOVERNMENT 438 (2014) ("One central challenge in First Amendment law is the extent to which organizations that engage in various speech activities, broadly conceived, should be subject to the general law of the land. This is clearly unproblematic in connection with the general rules against force and fraud, for which speech acts are used as evidence of intent in such crimes as murder, rape, arson, and theft. The legal system could not operate if the external evidence of these mental states was systematically excluded from evidence, which of course it is not.").

[255] *See* Bernard Marr, *What Is Unstructured Data and Why Is It So Important to Businesses? An Easy Explanation for Anyone*, FORBES (Oct. 16, 2019), https://www.forbes.com/sites/bernardmarr/2019/10/16/what-is-unstructured-data-and-why-is-it-so-important-to-businesses-an-easy-explanation-for-anyone/ [https://perma.cc/Q3HV-5E8G].

[256] *See id.* (listing these items as the most common examples); Dick Weisinger, *Unstructured Data: Using Analytics to Make Sense of Dark Data's Secrets*, FORMTEK (June 29, 2014), https://formtek.com/blog/unstructured-data-using-analytics-to-make-sense-of-dark-datas-secrets/ [https://perma.cc/J98P-BALA] (reporting on Bloomberg Vault's global head, Harald Collet, as saying, "Eighty percent

Identifying collective values in a particular context from internet data will require natural language processing. It will be up to AI to determine from that unstructured data whether people in the community have taken a critical position regarding the conduct in question. Within NLP, the AI technology that is likely to be most useful in a Conventionalist Conception is sentiment analysis (SA), which seeks to use machines to process natural language in order to determine the emotional state of the author.[257] With SA, the machine attempts to identify the emotional or attitudinal state of the author of the text under examination—e.g., that they are pleased with the subject they are talking about, are critical, are saddened, etc.[258] Although SA is relatively young as a discipline, it has made considerable progress.[259]

Importantly, its success is determined in part by the degree to which it can reliably analyze the semantics and pragmatics of text. In a recent study, a group of computer scientists tortured top SA software by making semantic and pragmatic manipulations to film reviews, finding them both to be effective, especially pragmatic manipulations.[260] For example, the scientists changed a sentence with positive sentiment, "Unfolds with the creepy elegance and carefully calibrated precision of a Dario Argento horror film," to a sentence of negative sentiment, "Unfolds with all the creepy elegance and carefully calibrated precision of a Jim Carrey comedy film."[261] The programs performed poorly under pragmatic and semantic alterations, failing 80% and about 70% of the time, respectively; syntactical alterations created failure rates that were slightly above 60%.[262]

As NLP becomes more semantic and pragmatic, sentiment analysis performance will improve. In recent years, state-of-the-art sentiment analysis has relied upon machine-learned algorithms.[263] When reinforced with examples of ground truth sentiment at the individual level, we can expect a machine-learning-powered algorithm at least to guess sentiment correctly more often

---

of [unstructured] data is generated by humans in the form of documents, emails, and recorded phone calls and is typically harder for employees to manage" (alteration in original)).

[257] *See* BING LIU, SENTIMENT ANALYSIS: MINING OPINIONS, SENTIMENTS, AND EMOTIONS 1 (2015).

[258] *Id.* In this paper, I do not distinguish between Sentiment Analysis and Emotion Analysis.

[259] Michael A. Livermore et al., *Computationally Assisted Regulatory Participation*, 93 NOTRE DAME L. REV. 977, 1004–05 (2018) (discussing the progress of sentiment analysis).

[260] Taylor Mahler et al., *Breaking NLP: Using Morphosyntax, Semantics, Pragmatics and World Knowledge to Fool Sentiment Analysis Systems*, *in* ASS'N FOR COMPUTATIONAL LINGUISTICS, PROCEEDINGS OF THE FIRST WORKSHOP ON BUILDING LINGUISTICALLY GENERALIZABLE NLP SYSTEMS 33–39 (2017), https://aclanthology.org/W17-54.pdf [https://perma.cc/L65L-X3HR].

[261] *Id.* at 36.

[262] *Id.* at 37–38 & fig.5.

[263] *See* Shervin Minaee et al., *Deep-Sentiment: Sentiment Analysis Using Ensemble of CNN and Bi-LSTM Models*, CORR (2019), https://arxiv.org/pdf/1904.04206.pdf [https://perma.cc/5A84-RZM2].

over time.[264] Of course, identifying the sentiment of an author is a different task than determining the collective sentiment of a population regarding a category of action. Needless to say, the degree of handcrafting required to create and, especially, to maintain the system will cut into its utility.

One saving grace for pragmatic advancement could be the rapidly expanding collection of data regarding the non-verbal aspects of context, which improve the reliability of interpretations of the statements made in that context and, additionally, increase access to mental states.[265] Facial expressions, skin conductivity, temperature, heart rate, blood pressure, location, and other physiological measures tend to increase the accuracy of our determinations of mental states.[266] As the Internet of Things rapidly expands, more physiological information is placed onto the internet and might eventually be accessed by the machine-learning-powered NLP programs.[267]

Lastly, even our soft-Lewis approach could suffer from gaps in which no convention has arisen. As with other approaches, closure rules, such as placing the burden on the plaintiff to identify a sufficiently reliable convention, offer a sort of fix, but overreliance on closure rules could lead to unjust or unpopular results.[268] Of course, it is possible, that human juries will arrive at unjust or unfair results when deciding under similar circumstances, and it is difficult to know at this point in time whether machines will surpass human performance in this regard.

All of this is to say that if technology continues on its current trajectory and courts are granted access to sufficient data, then it is theoretically possible that an automated Conventionalist Conception of the RPPT could develop and lead to satisfactory outcomes. Legislators would have to wait a long time[269] before the prospect of funding such a project would be appealing, and even then, it will be colossally expensive to gain sufficient background data and to develop reliable processing techniques. Compared to the rival conceptions we have already discussed, this approach not only requires assessments of conformity that are similar to (and just as costly as) the Average Conduct Concep-

---

[264] *See id.* (stating that vast amount of labeled data has helped bring deep-learning models used for sentiment analysis to state-of-the-art levels).

[265] *See, e.g.*, Lin Shu et al., *A Review of Emotion Recognition Using Physiological Signals*, SENSORS, June 28, 2018, at 1, 33.

[266] *See id.*

[267] *See* V. Jagadeeswari et al., *A Study on Medical Internet of Things and Big Data in Personalized Healthcare System*, HEALTH INFO. SCI. & SYS., Dec. 2018, at 1, 1 (2018). To be clear, I'm not suggesting that this is a good thing!

[268] *See* Benjamin C. Zipursky, *Sleight of Hand*, 48 WM. & MARY L. REV. 1999, 2033 (2007) ("Of equal importance, conventionalist accounts do not adequately explain the central role of the prototype of the reasonably prudent person, and they do not explain how the reasonably prudent person standard can function if community convention does not offer any decisive answer.").

[269] *See* Cambria & White, *supra* note 76, at 51.

tion, but it additionally requires complex community identification and NLP techniques that are not currently reliable enough to do the job.

2. Compatibility

Kelley and Wendt have persuasively argued that the Conventionalist Conception is the interpretive framework of the RPPT *most* supported by model jury instructions across the country:

> The instructions seem to call on the jury to determine whether the defendant's conduct, which resulted in harm to the plaintiff, was a private injustice to the plaintiff. So it is only in relation to the community's standard of what conduct the defendant owed to the plaintiff under the circumstances of the case that the jury is called on to make a community moral judgment. That judgment is not a free-floating moral judgment, but a precise determination of whether, in light of the community's preexisting coordination patterns, the plaintiff could reasonably have expected the defendant to have acted differently in order to protect the plaintiff from harm . . . .
>
> . . . .
>
> . . . The language points the jury to preexisting standards of conduct that the plaintiff could reasonably expect from the defendant . . . . The most persuasive explanation of the negligence standard, and the jury's role in applying that standard, identifies the source of that preexisting standard in the safety conventions of the community and the associated expectations of the plaintiff.[270]

There is also evidence that this conception is a fair characterization of how jurors behave. Feigenson's field research yielded the finding that jurors have a tendency to rely upon cultural norms in assessing defendant behavior.[271] Steven Hetcher's analysis of the empirical literature led him to conclude "that there is good reason to suppose that juries do not engage in Hand Test normative processing, but instead draw from their diverse array of everyday norms and customs when providing concrete substance to the abstract reasonable person standard to render a decision on the issue of negligence."[272] Tobia, in interpreting the results of his experimental studies, offered the possibility that reasonableness determinations might differ from ideal or average judgments in part because of the role that community values play in assessments of statisti-

[270] Kelley & Wendt, *supra* note 135, at 621–22.

[271] FEIGENSON, *supra* note 173, at 165 ("Thus, the jurors' decision reflects a confluence of various cultural norms but a common pattern of inferring blame from the transgression of those norms.").

[272] Hetcher, *supra* note 208, at 646–47.

cally common behavior.[273] Jaeger echoes this sentiment. After his experiments provided conflicting evidence regarding whether data showing aspirational conduct or data showing average conduct were more relevant to layperson determinations, Jaeger concluded, "Legal scholars have sometimes been critical of the idea that the reasonable person standard is informed by observations and beliefs about what other people would do" but "[m]y findings indicated that lay participants put more weight on information about community customs than many tort theorists would expect and much less weight on cost-justification than many tort theorists would expect."[274] Recall that a Lewisian-AI determination of conventions could capture what he calls the "hybrid view" of reasonableness: it involves an assessment of both statistical frequency and shared standards of criticism.[275]

Still, none of these authors would go so far as to say that the Conventionalist Conception is the only permissible way to understand the RPPT from an internal or external point of view. In other words, although it receives the *most* support, it would be wrong to suggest that other conceptions are non-starters.

It could be the breadth and flexibility accorded to RPPT interpreters in a Conventionalist Conception that makes it most consistent with existing law and practice. It doesn't rule much out. Functionally, the standard comports with the wide-ranging approach to determining liability that many jurors appear to take.[276] It would be unwise to assume, however, that variety in juror conduct alone indicates that they are cognitively following a conventionalist interpretation of RPPT. In other words, the mere fact that jurors do not appear to be following a single, mechanical approach to the RPPT does not entail that conventionalism is the winner.

There are at least two more problems that could be exacerbated with an AI-powered Conventionalist Conception RPPT. First, conventions might be at odds with notions of legality or morality: racism, sexism, homophobia, and other disgusting notions are not categorically out of bounds with the Conventionalist Conception, even if they are ruled out under substantive law.[277] Second, as discussed, algorithm-powered AI has the capacity to mask biases, making them all the more pernicious.[278]

---

[273] Tobia, *supra* note 136, at 332.

[274] Jaeger, *supra* note 141, at 932, 948.

[275] Tobia, *supra* note 136, at 305–12, 335–41.

[276] *See* Shari Seidman Diamond & Jessica M. Salerno, *Empirical Analysis of Juries in Tort Cases*, *in* RESEARCH HANDBOOK ON THE ECONOMICS OF TORTS 414, 414 (Jennifer Arlen ed., 2013).

[277] *Cf.* Tobia, *supra* note 136, at 342 (identifying risk of such immoralities with statistical notions of reasonableness).

[278] *See* discussion *supra* note 117.

Summing up, the Conventionalist Conception is more costly than both the Average Conduct Conception and the Hand Test with regard to background data and reliability, but it is the most compatible with existing law.

D. Moral Approaches

Yet another candidate for the RPPT are Moral Conceptions. On this account, the RPPT directs jurors to engage in some form of moral reasoning.[279]

Seminal American torts theorist, Francis Bohlen, after dismissing the Average Conduct Conception, captured the spirit of morality-driven interpretations of the RPPT:

> [The reasonable man] is an ideal creature, expressing public opinion declared by its accredited spokesman, whether court or jury, as to what ought to be due under the circumstances by a man who is not so engrossed in his own affairs as to disregard the effect of his conduct upon the interests of others. He may be called a personification of the court or jury's social judgment. The factor controlling the judgment of the defendant's conduct is not what is, but what ought to be.[280]

Thus, under a Moral Approach Conception the RPPT is an instance in which moral principles—whether they are principles culled from notions of equality,[281] corrective justice,[282] virtue ethics,[283] or some other similar domain—have been incorporated into the law,[284] operating as a demand that those who apply it engage in a moral evaluation of the instant case. Moral Conceptions have some appeal, at least compared to the other conceptions I have discussed, be-

---

[279] *See* Tobia, *supra* note 136, at 304.

[280] BOHLEN, *supra* note 169, at 604 (emphasis omitted).

[281] *See, e.g.*, ARTHUR RIPSTEIN, EQUALITY, RESPONSIBILITY, AND THE LAW 53 (paperback ed. 2001).

[282] *See, e.g.*, Jules L. Coleman, *Legal Theory and Practice*, 83 GEO. L.J. 2579, 2603–04 (1995).

[283] *See, e.g.*, Heidi Li Feldman, *Prudence, Benevolence, and Negligence: Virtue Ethics and Tort Law*, 74 CHI.-KENT L. REV. 1431, 1431–32 (2000).

[284] *See, e.g.*, Mark P. Gergen, *The Jury's Role in Deciding Normative Issues in the American Common Law*, 68 FORDHAM L. REV. 407, 408 n.1 (1999) ("[T]he common form of jury instruction on the standard of care in negligence (the reasonably prudent person standard) is an example of a standard that allows the decision-maker to choose moral criteria."); John V. Jacobi, *Fakers, Nuts, and Federalism: Common Law in the Shadow of the ADA*, 33 U.C. DAVIS L. REV. 95, 115 (1999) ("The common-law reasonably prudent person standard, however, did not derive from logical or semantic purism, but from a pragmatic application of the moral fault principle."); Timothy Macklem, *Provocation and the Ordinary Person*, 11 DALHOUSIE L.J. 126, 135 (1987) ("Moreover, in the law of negligence and in other areas of the criminal law, the reasonable person test indicates an ethical standard . . . ."); Glanville Williams, *Provocation and the Reasonable Man*, 1954 CRIM. L. REV. 740, 742 ("[H]ow can it be admitted that that paragon of virtue, the reasonable man, gives way to provocation?").

cause they operate as a check on certain forms of evil, insofar as that can be widely acknowledged.

1. Background Data and Reliability

The Moral Approach Conception is the most computationally demanding test we have discussed. There are those who believe that a moral reasoning machine is impossible or unforeseeably distant because moral reasoning is inextricably linked to features that we associate with humans or other living creatures.[285] This is not unreasonable; as there is support for the positions that our moral reasoning is intertwined with our consciousness and emotions[286] and that technologists will struggle to imbue machines with those features or their functional equivalents.[287] These arguments are not so convincing, however, that we should stop our analysis here, particularly because it is not the case that all Moral Approach Conceptions assert that consciousness and emotion are necessary conditions. Besides, a small number of people believe that AI might have already achieved consciousness![288]

---

[285] *See, e.g.*, John R. Searle, *Minds, Brains, and Programs*, 3 BEHAV. & BRAIN SCIS. 417, 417–424 (1980) (articulating the well-known argument against impossibility of machine understanding).

[286] *See, e.g.*, NEIL LEVY, CONSCIOUSNESS AND MORAL RESPONSIBILITY 86 (2014) (using neuroscience to support the claim that consciousness plays an integrative function that significantly increases our capacity to reason and, ultimately, be morally responsible and stating that "[g]enuine flexibility of behavior and of reasoning—and therefore the extent to which human beings are able to approximate to rational, domain-general, reasoners—is dependent on consciousness"); Jonathan Haidt, *The New Synthesis in Moral Psychology*, 316 SCI. 998, 998–1002 (2007) (discussing the primary role of emotions in moral reasoning).

[287] *See, e.g.*, Drew McDermott, *Artificial Intelligence and Consciousness*, *in* THE CAMBRIDGE HANDBOOK OF CONSCIOUSNESS 117, 146 (Philip David Zelazo, Morris Moscovitch & Evan Thompson eds., 2007) ("Consciousness stems from the structure of the self-models that intelligent systems use to reason about themselves. . . . Frustratingly, we won't be able to create systems and test hypotheses against them in the foreseeable future, because real progress on creating conscious programs awaits further developments in enhancing the intelligence of robots. There is no guarantee that AI will ever achieve the requisite level of intelligence, in which case this chapter has been pretty much wasted effort." (emphases omitted)); Bernd Carsten Stahl, *Information, Ethics, and Computers: The Problem of Autonomous Moral Agents*, 14 MINDS & MACHS. 67, 79 (2004) ("A further narrowing of the [Moral Turing Test (MTT)] to cognitivist ethics would appear to increase the computer's chances by ruling out emotional or intuitive appeals that computers would presumably have difficulties with. Given all these limitations, one could expect computers to pass the MTT. While this is basically an empirical question, it nevertheless seems improbable to me that a computer will succeed in this any time soon.").

[288] *See* George Musser, *Consciousness Creep*, AEON (Feb. 25, 2016), https://aeon.co/essays/could-machines-have-become-self-aware-without-our-knowing-it [https://perma.cc/54K6-SV4R] (arguing that our tests for consciousness need revision and that conscious AI might have happened already); James A. Reggia et al., *Beliefs Concerning the Nature of Consciousness*, 22 J. CONSCIOUSNESS STUD. 146, 158 (2015) (surveying university students and finding that 3% believe that contemporary computers are conscious).

It will be helpful to start with programs that are already in existence. Several computer scientists and philosophers have offered computerized approaches to moral decision-making.[289] The most complete and transparent approaches offer hand-crafted rules or algorithms for narrow situations. For example, Derek Leben offered a programmable algorithm based on the moral reasoning of philosopher John Rawls for autonomous vehicles that face trolley problem situations.[290] Whether the algorithms that result from these efforts truly embody the complexity of the philosophy from which they arose can be questioned,[291] but there is a more obvious shortcoming for our purposes. Generally speaking, closed-system approaches like static algorithms can produce satisfactory results only in narrow circumstances.[292]

From a background data standpoint, fixed codifications of moral reasoning are inexpensive (assuming that the codified moral approach was readily knowable), making background data costs low. But these success stories are likely to be rare. It is uncontroversial to assert that moral disagreement is widespread,[293] so there is reason to believe that any satisfactory handcrafted approach will have a rather narrow scope of application. Such approaches are, therefore, not reliable enough to be used in the RPPT context, given its breadth. What might be needed is something closer to a multi-purpose moral deliberator.

Although machine-learning could theoretically offer a solution, efforts thus far have not yielded reliable results. Techno-optimistic ethicists Wendell Wallach and Colin Allen discussed its viability in connection with developing a virtue ethics-based program:

---

[289] *See, e.g.*, Susan Leigh Anderson & Michael Anderson, *A Prima Facie Duty Approach to Machine Ethics and Its Application to Elder Care*, *in* AAAI PUBL'NS, WORKSHOPS AT THE TWENTY-FIFTH AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE: HUMAN-ROBOT INTERACTION IN ELDER CARE (2011), https://www.aaai.org/ocs/index.php/WS/AAAIW11/paper/view/3812 [https://perma.cc/6BBR-A7Q5] (presenting a prima facie ethics approach); WENDELL WALLACH & COLIN ALLEN, MORAL MACHINES: TEACHING ROBOTS RIGHT FROM WRONG (2009) (presenting a virtue ethics approach); Derek Leben, *A Rawlsian Algorithm for Autonomous Vehicles*, 19 ETHICS & INFO. TECH. 107, 107–15 (2017) (presenting a Rawlsian approach); Thomas M. Powers, *Prospects for a Kantian Machine*, IEEE INTELLIGENT SYS., July/Aug. 2006, at 46, 46–51 (presenting a Kantian approach).

[290] *See* Leben, *supra* note 289, at 107–15.

[291] *See* Geoff Keeling, *Against Leben's Rawlsian Collision Algorithm for Autonomous Vehicles*, *in* 44 STUD. IN APPLIED PHIL., EPISTEMOLOGY AND RATIONAL ETHICS, PHILOSOPHY AND THEORY OF ARTIFICIAL INTELLIGENCE 2017, at 259, 259–72 (Vincent C. Müller ed., 2018) (disagreeing with Leben's interpretation of Rawls's original position).

[292] *See* Anderson & Anderson, *supra* note 289, at 3 ("Since we wanted to focus on the critical problem of discovering a decision principle required for a machine to implement a prima facie duty ethical theory, in the process establishing a prototype solution to the problem, we constrained the task as much as possible.").

[293] *See, e.g.*, RICHARD ROWLAND, MORAL DISAGREEMENT 6 (2021) ("There is clearly a lot of moral disagreement.").

> It is interesting and suggestive to note the similarity between Aristotelian ethics and [neural network machine learning], and the possibility that character might emerge from a [neural network] of how the brain works. Given that virtues are context-sensitive, the power of [neural network machine learning] to unite virtue theory and particularism is attractive. However, existing [neural network] systems are a long way from tackling the kind of complex learning tasks one associates with moral development. The challenge of implementing virtues within a neural network remains a formidable one.[294]

The scope of the injuries that fall under the umbrella of ordinary negligence is so large[295] that handcrafting is unlikely to yield a scalable tool.

Even if there exist moral imperatives that naturally structure into programmable rules, those approaches might not correspond to the approach demanded under the Moral Conception of the RPPT. Because there is no single moral approach to the RPPT or account of moral reasoning that emerges as the consensus pick among legal theoreticians,[296] let alone among judges or jurors, I will consider the leading alternatives.

It is somewhat popular for legal theorists to conceive of moral reasoning as an originative act in which reasoners rely on creativity or imagination to select and balance multiple values in arriving at a result.[297] Richard H. Fallon, Jr. offers a vivid account:

> Much practical moral reasoning requires a kind of instinctive, frequently unconscious sizing up of the facts of a situation, including others' psychological states and their likely perception of particular

---

[294] WALLACH & ALLEN, *supra* note 289, at 123.

[295] *See* Brian P. Dunigan & Jerry J. Phillips, *Comparative Fault in Tennessee: Where Are We Going, and Why Are We in This Handbasket?*, 67 TENN. L. REV. 765, 788–89 (2000) (describing negligence as encompassing "infinite degrees of human behavior"); *cf.* Gergen, *supra* note 245, at 1008 (describing a court's broad capacity to create causes of action under the umbrella of negligence based on an "all-things-considered" assessment).

[296] *Compare* Seana Valentine Shiffrin, *Inducing Moral Deliberation: On the Occasional Virtues of Fog*, 123 HARV. L. REV. 1214, 1224–25 (2010) (emphasizing connection between standards and moral deliberation and stating "citizens must engage in legal interpretation by engaging with the underlying purposes of law"), *with* Tomiko Brown-Nagin, *A Critique of Instrumental Rationality: Judicial Reasoning About the "Cold Numbers" in* Hopwood v. Texas, 16 L. & INEQ. 359, 414 (1998) ("Moral reasoning requires courts to consider the nation's social problems and the social meaning of legal rulings in all their complexity."), *and* W. Bradley Wendel, *Institutional and Individual Justification in Legal Ethics: The Problem of Client Selection*, 34 HOFSTRA L. REV. 987, 1002 (2006) ("[M]oral reasoning requires stepping back from our reasoning process and deciding whether we are willing to endorse our reasons from the standpoint of a description under which we value ourselves . . . .").

[297] *See, e.g.*, STUART HAMPSHIRE, MORALITY AND CONFLICT 29–31 (1983) (emphasizing the imaginative dimensions of moral reasoning); Jeff Todd, *The Poetics and Ethics of Negligence*, 50 CAL. W. L. REV. 75, 80, 120–21 (2013) (viewing the RPPT as a "poetic device").

> acts or gestures as kind, supportive, funny, creative, insightful, and so forth. In my experience, there are many non-philosophers whose lives exhibit shining moral excellence, not because they are profound conceptual reasoners, but because of their capacities to make insightful and sometimes imaginative psychological and empirical judgments about the effects of possible actions and about ways to transform moments or lives for the better. They perceive possibilities, and means for realizing those possibilities, that others—sometimes including first-rate philosophers—do not.[298]

The notion that moral reasoning is imaginative, occasionally unconscious, and insightful is in line with the modern philosophical trends that reject rule-driven conceptions in favor of particularist, contextual ones. As philosopher Catrin Misslehorn explains:

> Understanding morality as a kind of calculus was quite popular in the history of moral philosophy. . . . Yet, there has been more recently a trend towards challenging this way of thinking. The opponents of this view have a radically different picture of the human mind and morality. Dyed-in-the-wool particularists believe that human morality cannot be captured by a set of rules, however rough they may be.[299]

Misselhorn notes that this conception is abhorrent to existing methods of computation: "Such an extremely context-sensitive capacity is very hard to implement in an artificial system, even with the help of a bottom-up approach [such as those that derive morality from patterns in data or upon models that mimic evolution]."[300] The more responsive to context a program needs to be, the more data it needs in order to identify reliable context-specific algorithms. And even if that were possible, it would be difficult to confirm that the algorithms mimic the moral reasoning process demanded by the RPPT, as opposed to following an inscrutable, counter-intuitive process-based statistical correlation.[301]

A less artistic version of the moral reasoning process simply demands that the interpreter act as a moral agent.[302] This requires some degree of autonomy in deciding right and wrong, as well as acting upon reasons that justify those

---

[298] Richard H. Fallon, Jr., *Is Moral Reasoning Conceptual Interpretation*?, 90 B.U. L. REV. 535, 544 (2010).

[299] Catrin Misselhorn, *Artificial Morality. Concepts, Issues and Challenges*, 55 SOC'Y 161, 168 (2018).

[300] *Id.*

[301] *See generally* Brian Sheppard, *Warming Up to Inscrutability: How Technology Could Challenge Our Concept of Law*, 68 U. TORONTO L.J. 36, 48–49 (2018) (explaining the ways that algorithmic outputs can be incomprehensible).

[302] Misselhorn, *supra* note 299, at 161.

decisions according to the decider.[303] In the legal context, some scholars have described jurors as moral agents.[304]

Faster processing and increased access to relevant background data are, on their own, unlikely to give rise to a machine that can engage in moral reasoning, so understood.[305] This generation of AI is good at following a set program, but it is not yet capable of using a rich and versatile method to identify moral conduct or to choose among various ethical approaches. In short, it struggles to be autonomous. Machine learning is, at its basic core, a statistical approach that assesses progress based on how well the program increases predictive power or other maximization goals; it ordinarily gets better as it gets more feedback data and further adjusts its statistical approach.[306] We have a long

---

[303] Vicky Charisi et al., *Towards Moral Autonomous Systems*, CoRR (2017), https://arxiv.org/abs/1703.04741 [https://perma.cc/HD3B-KKJD] ("[Moral agency] means . . . capable of understanding concepts of right and wrong. The debate of whether an artificially created entity can be a moral agent is still far from settled. Wallach and Allen hint at the idea that how ethically sensitive an artificial system can be depends on how able of autonomous action it is." (emphasis and citations omitted)); Misselhorn, *supra*, note 299, at 163 ("The first dimension involves the idea that agents are in some sense self-originating sources of their doings. This condition has been formulated in different ways, for instance, by saying that actions may not be determined by external factors, involve some sort of flexibility, or are under the control of the agent. The second dimension that being an agent involves is the capacity to act for reasons.").

[304] *See, e.g.*, John R. Allison, *Combinations of Decision-making Functions, Ex Parte Communications, and Related Biasing Influences: A Process-Value Analysis*, 1993 UTAH L. REV. 1135, 1219 ("One important distinction between single-decider views and jury views that may overcome many of the risks associated with jury views is the fact that a jury includes multiple independent intellectual and moral agents."); Mitchell Keiter, *From* Apprendi *to* Blakely *to* Cunningham*: Popular Sovereignty Enters the Courtroom*, 34 W. ST. U. L. REV. 111, 140 (2007) ("American law vindicates the perception of the individual as a rational moral agent in her role as a voter, who may select candidates and the laws they should to enact (including sentencing provisions), in her role as a defendant, who may choose how to conduct her own defense, in her role as a juror . . . ."); Robert F. Schopp, *Verdicts of Conscience: Nullification and Necessity as Jury Responses to Crimes of Conscience*, 69 S. CAL. L. REV. 2039, 2107 n.173 (1996) ("The juror, as an independent moral agent, must consider all morally relevant reasons to act, but reasons that carry normative force only in the juror's comprehensive doctrine are not relevant to the conventional public morality.").

[305] *See* John Sullins, *Information Technology and Moral Values*, *in* STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Edward N. Zalta ed., 2018), https://plato.stanford.edu/entries/it-moral-values/#MorParInfTec [https://perma.cc/ATP2-V5VD] ("While scholars recognize that we are still some time from creating information technology that would be unequivocally recognized as an artificial moral agent, there are strong theoretical arguments in suggesting that automated moral reasoning is an eventual possibility and therefore it is an appropriate area of study for those interested in the moral impacts of information technologies."). I should mention that quantum computing shows promise for huge leaps in calculation speed and pattern recognition, but it is not yet clear that it will break the bounds of purpose-built machines into practical computing. *See* Richard Versluis, *Here's a Blueprint for a Practical Quantum Computer*, IEEE SPECTRUM (Mar. 24, 2020), https://spectrum.ieee.org/computing/hardware/heres-a-blueprint-for-a-practical-quantum-computer [https://perma.cc/HD68-JCZY].

[306] *See* Matthew Stewart, *The Actual Difference Between Statistics and Machine Learning*, TOWARDS DATA SCI. (Mar. 24, 2019), https://towardsdatascience.com/the-actual-difference-between-statistics-and-machine-learning-64b49f07ea3 [https://perma.cc/DER8-TVEV].

way to go before a system can reliably escape the strictures of its initial goal and turn back on itself, changing its own fundamental project as it learns more about human goals or values even though they were not pre-programmed.[307] We might be farther still from a machine acting autonomously and reflecting upon or seeking to justify a moral position with reasons.[308]

A promising technique in this regard is inverse reinforcement-learning (IRL), in which a computer is programmed to observe conduct and help delineate or infer its hidden values or goals.[309] This is distinct from traditional reinforcement learning, in which the goal is posited and learning is in service of furthering that goal.[310] We should not expect this technique to lead to complex realizations of social goods anytime soon, however.[311] And machine learning has thus far made more progress when operating for a pre-programmed goal than when operating to learn new goals and self-revise.[312] But like other species of machine learning, IRL's performance is contingent upon access to copious amounts of relevant and exploitable data of human moral decision-making, making it expensive from a background data standpoint.[313]

Yet another approach under the Moral Approach umbrella are rights-based theories. Benjamin Zipursky, describing commonalities among rights-based theorists such as Gregory Keating, Arthur Ripstein, Ernest Weinrib, and to some extent himself, explained that "individuals are entitled, as a matter of political morality, to a substantial level of respect and vigilance for their physical integrity—as well as their property."[314] At first glance, these rights-based theories appear not to emphasize the moral character of jury deliberation, which, if true, would make them friendlier to the mechanical approaches at which machines excel. Indeed, Zipursky concedes that it bears a superficial resemblance to the Hand Test "because it accommodates both the need for ac-

---

[307] *See* Natalie Wolchover, *Artificial Intelligence Will Do What We Ask. That's a Problem*, QUANTA MAG. (Jan. 30, 2020), https://www.quantamagazine.org/artificial-intelligence-will-do-what-we-ask-thats-a-problem-20200130/ [https://perma.cc/EVB5-EKX3].

[308] *See* Misselhorn, *supra* note 299, at 164 ("An important aspect of moral responsibility is the capacity to deliberate about one's moral reasons, and current artificial systems cannot do this.").

[309] *See* Sven Collette et al., *Neural Computations Underlying Inverse Reinforcement Learning in the Human Brain*, 6 ELIFE 1, 1–2 (Oct. 30, 2017), https://elifesciences.org/articles/29718#info [https://perma.cc/8F3A-BY69] (discussing the recent formalization of inverse reinforcement-learning).

[310] *See id.*

[311] Wolchover, *supra* note 307 (discussing the lack of expectation that machine understanding will be good anytime soon).

[312] *Id.*

[313] *See* Alexandre Gonfalonieri, *Inverse Reinforcement Learning: Introduction and Main Issues*, TOWARDS DATA SCI. (Dec. 26, 2018), https://towardsdatascience.com/inverse-reinforcement-learning-6453b7cdc90d [https://perma.cc/6USB-4XAJ] (explaining that IRL is unique because it can potentially extract useful data from any digital record in an unsupervised or semi-supervised manner).

[314] Zipursky, *supra* note 268, at 2030.

tion and the need to be free of harm."[315] But he goes on to explain that, when properly understood, it defies simple formulation "because it does not aggregate well-being."[316] In other words, jurors must set the line for liability between the right not to be physically injured and the right not to be unduly restricted in action; but this is not operationalizable as a formula.

Although it is possible to handcraft programs with rules of defeasibility that might resolve conflicts between rules,[317] it is difficult to see how one might encode the borderlines of the competing rights without some way to weight or otherwise quantify those values. I cannot rule out the possibility that a model that is friendly to if-then structures could be derived; yet it is somewhat clearer that such a technique is unlikely to be reliable in the wide variety of contexts that courts ordinarily encounter.

Despite the differences in the Moral Approaches discussed here, they are all quite expensive, requiring either a paradigm shift in computing or, more modestly, far-off innovations in machine-learning or other artificial intelligence techniques. In short, this is the most expensive approach from the standpoint of background data and reliability.

2. Compatibility

Turning to legal compatibility, it should be obvious from the variety of approaches already discussed that Moral Approach Conceptions are popular among academics. They do not find much support in the instructions provided to jurors, however.

Although the RPPT uses evaluative language, it does not command, in the content of the standard itself, that jurors make their own moral judgment; rather, the typical instruction directs them to consult a reasonable person *standard*. Kelley and Wendt describe it thusly:

> Each of the recurring critically important phrases in these pattern instructions, ordinary care and the conduct of a reasonably prudent or a reasonably careful person, seem to refer to a preexisting standard. That standard does not seem to be so much a moral as a social standard, based on the actual conduct of one who exercises ordinary care for the safety of others. This social, rather than moral, nature can be seen, as well, by focusing on what these instructions do not

[315] *Id.* at 2031.
[316] *Id.*
[317] *See* discussion *supra* note 131.

> contain. There is no mention of fault, improper conduct, excuses, good or bad, right or wrong.[318]

Perhaps this is unsurprising. As the analysis thus far has shown, scholarly Moral Approach Conceptions exhibit a high degree of variation, and it is reasonable to surmise that this diversity of opinion would not have arisen had the courts directed jurors to engage in a particular type of moral reasoning.[319]

Insofar as we want a machine that emulates what human jurors *actually* do, there is empirical evidence that jurors engage in moral reasoning during deliberation. Just as Feigenson's research supports Conventionalist approaches by providing evidence that cultural norms are a factor in juror assessments of liability, it additionally supports the Moral Approach by providing evidence that "jurors [also] thought about responsibility in a personalized and moralized way."[320] Likewise, in two experiments, psychologists John Bernard, Robert Cohen and Michael Lupfer observed that juries composed of people with higher-level moral reasoning acquitted a defendant although mixed juries or juries composed of people with low-level moral reasoning hung.[321] These results are similar to those in an experiment performed by psychologists Ken Rotenberg, Maureen Hewlett, and Catherine Siegwart, in which they found that people with higher-level moral reasoning were more influential in jury deliberation.[322] There is little evidence that jurors engage in one of the detailed, scholarly accounts described above, but it is probable that jurors frequently tap into their senses of right and wrong and that doing so materially affects case outcomes. In other words, there is evidence that jurors act as moral agents.

This is not surprising. Theorists have suggested that the RPPT passes the buck, leaving the juror to rely on their own personal sense of right and wrong.[323] But that is a far cry from the RPPT constituting an authoritative legal dictate to use a moral approach, even if it is statistically likely that jurors will use the RPPT to engage their senses of right and wrong as a factual matter.

---

[318] Kelley & Wendt, *supra* note 135, at 620 (emphasis omitted).

[319] That said, the directives in jury instructions, although conventionalist in nature, lack the specific detail of, say, Kelley's more elaborate conventionalist accounts. So, we must be careful not to unfairly hold the detail of Moral Approaches against them.

[320] *See* Neal R. Feigenson, *Accidents as Melodrama*, 43 N.Y.L. SCH. L. REV. 741, 770–71 (1999).

[321] John L. Bernard et al., *The Influence of Juror's Level of Moral Reasoning and the Nature of Closing Arguments in Determining the Verdict in a Civil Case: A Report of Two Experiments*, 9 LAW. & PSYCH. REV. 93, 97–98 (1985).

[322] Ken J. Rotenberg et al., *Principled Moral Reasoning and Self-Monitoring as Predictors of Jury Functioning*, 20 BASIC & APPLIED SOC. PSYCH. 167, 172–73 (1998).

[323] *Cf.* John Gardner, *The Many Faces of the Reasonable Person*, 131 LAW Q. REV. 563, 568 (2015) (describing the position that RPPT incorporates non-law, such as moral principles, into the law by passing the buck).

The question here is whether a Moral Approach is a bigger deviation from existing *law* than its rival conceptions. The above analysis supports the conclusion that a legislator would face greater resistance in pushing for a Moral Approach than for a Conventionalist Conception, all other things being equal.

In sum, the Moral Approach is a poor performer in almost all dimensions. Its demands are likely beyond existing technology, leaving us to guess how future technology might deal with background data availability or reliability. Furthermore, it does not receive strong support in American jury instructions. There is evidence, however, that juror deliberation is significantly normative and to some extent, driven by a juror's personal sense of morality, as a de facto matter.

E. The Deflationary Conception

Having read the analysis thus far, one could not be blamed for wondering whether the existence of rival conceptions is evidence of a deeper truth: there just isn't a significant, legally dictated RPPT conception to be found. Although Conventionalism is most consistent with jury instructions, that is but one way to understand the meaning of the RPPT. Scholars have arrived at different conceptions after analyzing appellate decisions,[324] relying on political theory,[325] or using other methodologies.[326] To quote Gertrude Stein, maybe there really is "no there there."[327] Perhaps juries, finding little guidance, are engaging in unmoored, will-o-the-wisp decision-making. Their conduct might appear to be dictated by legal principles, but maybe it is really a byproduct of some combination of influences that does not comport with any of our law-grounded conceptions thus far.

I call this the "Deflationary Conception." Despite being something of an anti-conception, it might nevertheless offer a computerizable approach, one through which predictions of legal and factual determinations are reliable enough to meet our expectations of legality regardless of the way in which those determinations are made. Proponents of this approach would care about the meaning of the RPPT only insofar as it is necessary to predict how jurors actually decide cases under the standard.

The closest analogue to this view is the so-called "Prediction Theory of Law," which eschews grand theorizing about legal guidance from an internal

---

[324] *See, e.g.*, Ronald J. Allen & Ross M. Rosenberg, *Legal Phenomena, Knowledge, and Theory: A Cautionary Tale of Hedgehogs and Foxes*, 77 CHI.-KENT L. REV. 683, 708 (2002).

[325] *See, e.g.*, Gregory C. Keating, *Rawlsian Fairness and Regime Choice in the Law of Accidents*, 72 FORDHAM L. REV. 1857, 1862–80 (2004) (using a Rawlsian conception of fairness to interpret reasonableness in negligence and strict liability).

[326] *See, e.g.*, Stephen G. Gilles, *The Emergence of Cost-Benefit Balancing in English Negligence Law*, 77 CHI.-KENT L. REV. 489, 501–04 (2002) (looking at English tort law).

[327] GERTRUDE STEIN, EVERYBODY'S AUTOBIOGRAPHY 289 (1937).

point of view.[328] Holmes is credited with Prediction Theory. And although he might have espoused a different view of the RPPT in other areas of his writing, he took a broadly deflationary approach in *The Path of the Law*:

> The law talks about rights, and duties, and malice, and intent, and negligence, and so forth, and nothing is easier, or, I may say, more common in legal reasoning, than to take these words in their moral sense, at some state of the argument, and so to drop into fallacy.[329]

Instead, he argues, we would benefit by taking a positivistic stance and conceiving of the law as a prediction of how courts behave.[330]

This prediction theory is not specifically about the RPPT, but the breach standard is a fair target: thereunder, one could dismiss the notion that the content of the law is a specific call to community, economic, moral, or mathematical principle and assert, instead, that its meaning has no importance beyond assisting us in predicting what jurors or courts will do under the circumstances of a case.

1. Background Data and Reliability

Statistically-grounded predictions are AI's forte. Since at least Segal and Spaeth's first major work,[331] there has grown an impressive empirical literature regarding improved methods for predicting case outcomes, many of which use rather limited data but yield results that beat human experts.[332] Harnessing supervised machine learning techniques has ratcheted predictions even higher.[333] Currently, state-of-the-art methods incorporate reliable scores assigned to the

---

[328] *See* Michael S. Green, *Prediction Theories of Law and the Internal Point of View*, 51 SAN DIEGO L. REV. 921, 922–23 (2014) ("Under prediction theories, not only is the internal point of view not necessary for law, internal legal statements are morally suspect.").

[329] OLIVER WENDELL HOLMES, JR., THE PATH OF THE LAW 8 (Floating Press ed. 2009) (1897).

[330] *Id.* at 9 ("The prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law."). By "positivistic," I mean only that Holmes is viewing law as a social fact. I do not mean to suggest that Prediction Theory is emblematic of legal positivism. To the contrary, many legal positivists reject Prediction Theory insofar as it is taken to be constitutive of the concept of law. *See, e.g.*, HART, *supra* note 246, at 91.

[331] *See generally* JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL (1993) (using statistical methods to improve prediction of Supreme Court cases).

[332] *See* Michael Abramowicz & Emerson H. Tiller, *Citation to Legislative History: Empirical Evidence on Positive Political and Contextual Theories of Judicial Decision Making*, 38 J. LEGAL STUD. 419, 420–21 (2009) (summarizing major additions to the predictive model).

[333] *See* Daniel Martin Katz et al., *A General Approach for Predicting the Behavior of the Supreme Court of the United States*, 12 PLOS ONE 1, 4 (Apr. 12, 2017), https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174698 [https://perma.cc/L9JW-RSYN] (using random forests learning in a supervised context and improving on the general level of prediction demonstrated by prior work but with the power of being applied out-of-sample to the entire past and future of the Court, not a single term).

facts of previous cases, judges, and other relevant details, which are used to train the machine under human supervision with the help of machine-learning programs.[334] For over a decade, technologists have been keen to develop methods for predicting damages awards and settlement amounts in negligence cases; Lex Machina, a legal analytics company, claims, "With this data, for the first time, lawyers can predict the behaviors and outcomes that different legal strategies will produce."[335] And there is an emerging cottage industry of (arguably disturbing)[336] software that scrapes the internet or other sources to set prices for settlement or provide data about potential individual jurors regarding biases, strengths, and risks.[337]

A project of this nature is reliant on a massive amount of data regarding the circumstances that bear upon jury decision-making. Because it limits its focus to jurors, however, the universe of relevant data is potentially a lot smaller than what would be necessary for the construction of a versatile moral deliberator, convention finder, cost-benefit analyzer, or average assessor. Moreover, that universe has a finite limit because, when the process of automating the RPPT is complete, there will no longer be new jury determinations to review.

Although background data costs might be lower under the Deflationary Conception than under its rivals, it is an open question whether jurors exhibit patterned enough behavior when subjected to the demands of the RPPT to yield strong and reliable predictions of juror behavior. If so, then the Deflationary path to automation might be comparatively short. If not, the Deflationary project might be doomed. Even under the former scenario, however, there is the risk that performance will worsen over time.

This requires a bit of explanation. A good starting place is to imagine how the Deflationary Conception would operate if it were to reach its theoretical upper bound. Under this thought experiment: imagine that scientists, using similar techniques to those in existence today, develop an algorithm that predicts breach determinations under the RPPT with near one hundred percent accuracy. Suppose further that rather than predicting outcomes, courts begin to

---

[334] *The Supreme Court Database*, WASH. U. L., http://scdb.wustl.edu/ [https://perma.cc/SH35-ECUD].

[335] *About Us: What We Do,* LEX MACHINA, https://lexmachina.com/about/ [https://perma.cc/QAV9-R4NU].

[336] Todd Feathers, *This Company Is Using Racially-Biased Algorithms to Select Jurors*, VICE (Mar. 3, 2020), https://www.vice.com/en_us/article/epgmbw/this-company-is-using-racially-biased-algorithms-to-select-jurors [https://perma.cc/Q6K7-4YU7] ("Momus Analytics' predictive scoring system is using race to grade potential jurors on vague qualities like 'leadership' and 'personal responsibility.'").

[337] *See, e.g.*, MOMUS ANALYTICS, https://momusanalytics.com/ [https://perma.cc/VS6T-NQJR]; VOLTAIRE, https://voltaireapp.com/ [https://perma.cc/7TXN-JLGN].

use it to decide all new negligence cases. Under those circumstances, the court could be reasonably confident that the algorithm is reaching the same outcomes that human jurors would have reached had they been given the chance. At least initially.

The world is constantly changing, but the machine's predictions rely upon old data regarding the way that dynamic outside-of-the-jury-box factors impact the semi-random assembly of people who actually sat within the jury box. Replacing jurors with machines would make it impossible to capture post-automation data of juror responses to new, post-automation circumstances. Over time, its absence would likely diminish the predictive power of the machine, so long as the machine seeks to predict how new jurors at that point in time would respond to the circumstances of a case. Following the old trope about judges,[338] it might be the case that juror RPPT decisions are influenced by what a juror ate for breakfast. The development of a new breakfast cereal might have a novel impact on jury decision-making, but the machine would struggle to discover that connection when no data about jurors who had eaten it exist. The machine could simulate jurors, but its prediction would have to be based on prior juror behavior and the behavior of non-jurors.

Of course, this is just an intuition pump. It is unlikely that we will be able to predict juror behavior with that level of accuracy in the foreseeable future (if ever), but the dynamic is the same. Suppose that we can someday predict juror behavior at lower, but satisfactory, levels of accuracy. Even then, the accuracy of prediction will drop over time if jurors are displaced by machines, leading to increasing reliability problems.

Perhaps that weakness does not matter to Deflationists. If the content of the law does not guide the juror, why should we quibble over what *new* jurors would do? If the way jurors have actually behaved before automation was an incomprehensible mix of factors gathered from past decades of data, then why would a different incomprehensible mix gathered from current data better meet the demands of a rule of law, welfarism, or some other public value? These are not easy questions to answer, especially if we presume that the decision to automate the RPPT has already been made. After all, a significant number of people insist that we interpret our Constitution using the murky, collective understanding of people who lived more than two centuries ago.[339]

---

[338] *See generally* Dan Priel, *Law Is What the Judge Had for Breakfast: A Brief History of an Unpalatable Idea*, 68 BUFF. L. REV. 899 (2020).

[339] *See, e.g.*, Neil M. Gorsuch, *Why Originalism Is the Best Approach to the Constitution*, TIME (Sept. 6, 2019), https://time.com/5670400/justice-neil-gorsuch-why-originalism-is-the-best-approach-to-the-constitution/ [https://perma.cc/H69T-4M5H] ("Living constitutionalists often complain we can't know the original understanding because the document's too old and cryptic. Hardly. We figure out

We have thus far been focused on the problem of waning reliability, but there is the other side of the coin to consider: fixing the data to a particular period of history would lower background data costs over time. Without new data to consider, there would be little need to make significant background data expenditures once automation took place. It would vastly limit the amount of data necessary and would limit collection costs to the period before it launches.

Compared to other conceptions, then, the realm of relevant background data is very likely smaller than it would be under Conventionalist or Moral Conceptions, and, over time, it could be dwarfed by Average Conduct Conception or the Hand Test.

2. Compatibility

Little need be said about the profound incompatibility of the Deflationary Conception. No jury instructions take this approach. How could they? Juries do not engage in the absurdity of deciding based solely upon the best prediction of how they would decide.

The Prediction Theory is a one-size-fits-all approach to legality. It does not matter whether the most determinative factor is the text of the RPPT or something outside of the courtroom, or whether jurors are conscious of its influence or not. The only concern is predictive power. For example, the Deflationary Conception could lead to the bizarre conclusion that femininity of a male voice is, to some extent, built into the law of Supreme Court cases because, according to a recent study, male advocates were more likely to win Supreme Court cases when they are perceived as less masculine based on a speech sample of less than three seconds.[340] Needless to say, no jury instructions take this approach. Moreover, it is a conception that will not likely appeal to legislators because it gives no privileged place to the content of their laws.

The Deflationary Conception is political kryptonite. It has the highest compatibility costs of any approach discussed thus far.

To summarize, the Deflationary Account is the nuclear option, one that might shift the finish line closer to existing machine performance levels (due to high marks on both processing and background data) and make automation more feasible, but at the expense of dearly held notions of legality. Presumably, legislators are among the population of the legally faithful.

---

the original meaning of old and difficult texts all the time. Just ask any English professor who teaches Shakespeare or Beowulf.").

[340] Daniel Chen et al., *Perceived Masculinity Predicts U.S. Supreme Court Outcomes*, 11 PLOS ONE 1, 10 (Oct. 13, 2016), https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0164324 [https://perma.cc/R9AJ-K999].

F. Comparing the Alternatives

The foregoing analysis has revealed no clear winner for the legislator looking to maximize *both* technological feasibility and legal compatibility; some tests present fewer costs with respect to background data and reliability, but these approaches do not perform as well on compatibility.

The Deflationary and Average Conduct Conceptions are technologically attractive, but they are at the outer limits of legal compatibility. The former is founded on assumptions that are hostile to the widely accepted notion that jurors accept the RPPT as a legitimate standard for guiding their decision-making on breach. Moreover, the Average Conduct Conception fails to capture the fact that the RPPT invokes a more robust standard than numerical commonness. The Hand Test is a middling performer on most scores, but it is worse on compatibility than Conventionalist Conceptions and, as a matter of actual practice, the Moral Approach Conception. For its part, the Conventionalist Conception does best on permissibility but poorly on background data. Finally, the Moral Conception does quite poorly on technology scores and is not particularly impressive on compatibility.

Figure 1 shows a simplified version of their relationships to each other. Although it does not indicate the degree of separation between them, my hope is that it can serve as a useful rule of thumb. The asterisk marks the deep uncertainty regarding whether the Deflationary Conception can achieve or maintain reliability.

Figure 1: Simplified Ranking of RPPT Conceptions Along Technological and Legal Lines (1 is the top ranking)

	Average	L. Hand	Convention	Moral	Deflation
Background Data	2	3	4	5	1
Reliability	3	2	4	5	1*
Compatibility	4	3	1	2	5

Without a decisive winner in all categories, we are left to wonder whether legislators will be willing to accept the tradeoffs in adopting a cheaper, computer-friendly conception of the RPPT; whether they would be willing to wait for technological innovation that makes possible the automation of more legally compatible conceptions; or whether they will just sit the whole thing out.

The legislators who seek automation and for whom compatibility is a secondary concern will likely be most attracted to the Hand Test, and to a lesser extent the Average Conduct Conception, for the foreseeable future. Although the feasibility of creating satisfactory automated versions of those tests within that period is by no means clear, the data that it relies upon is commonly tracked in the automotive, insurance, and retail industries. Where there are gaps, the Hand Test is more reliable than the Average Conduct Conception.

As time goes on, however, automation of the Conventionalist Conception could become more appealing, particularly to those who value compatibility. Background data costs are likely to fall as trackers get more ubiquitous, storage gets cheaper, and data retrieval gets faster. The data will get better, it will get more plentiful, and it will get easier to process. The Hand Test and Conventionalist approaches will be clear beneficiaries as they presently face high background data costs and are most likely to use AI techniques that harness big data.

As background data costs fall, so too might reliability costs. Some of the techniques we have described become more reliable as they receive more or better data. For example, programs that seek to identify patterns in order to derive a function for RPPT compliance have weaker reliability if they lack enough data to identify correlations of high significance. All of the approaches highlighted here other than, perhaps, the deep AI versions of the Moral Approach are primarily based on techniques of that sort. As the gap of technological feasibility between those approaches shrinks, the Conventionalist Conception's advantage on compatibility is more likely to be determinative. Contrariwise, this dynamic suggests that those super-advanced Moral Approaches could be disfavored for a long period of technological development. All of this assumes, of course, that the law of breach does not significantly change in favor of a particular conception in the interim.

III. Structural Values of the Reasonably Prudent Person Test and the Fixed Costs of Automation

Thus far, the analysis has been framed by conceptions of the legal meaning of the RPPT. This was a sensible framework because an important part of legislation is considering how following the content of a norm, like the RPPT, will change the conduct of those subject to it, making a desired state of affairs more likely to occur. Lawmakers will have to settle on a conception of that content before they can decide how to automate it. It turns out that automation presents a tradeoff between computer-friendliness and complex, values-centered interpretation. In other words, analysis of the RPPT automation reveals a menu of incommensurate, imperfect compromises as stakeholders negotiate the dynamics of variable legislative cost. But what about fixed cost? Is there anything that would be lost or hurt by automation regardless of the conception of the RPPT legislators choose? The analysis below will reveal two: participatory and deliberative democracy. To get there, however, it will first discuss the surprising degree of freedom that jurors enjoy when they make breach determinations under the existing system.

A. The Puzzling Outcome-Centric Approach

One of the fascinating things about the breach determination for ordinary negligence is that it is the legislative hot potato of modern American civil law.

The RPPT is a pivotal determination in a massive number of cases: it is the one aspect of the tort case for ordinary negligence that determines whether the defendant committed a wrong, and ordinary negligence is the most commonly raised tort claim.[341] It is a meaningful and frequent opportunity to control the coercive power of the state to redress carelessly caused injuries. One might guess that lawmakers or other stakeholders would relish the opportunity to structure the resolution of so many tort claims, and the apportionment of so much money.

But that guess would be wrong. State legislatures have largely avoided supplementing or tinkering with the RPPT.[342] Judges have largely followed suit, saying very little about how to apply it in their jury instructions.[343] Judges in several states give model instructions that are designed to resist the notion of adding flesh to the RPPT skeleton, saying language like "the law does not say how the negligence standard applies, rather that it is for the jury to decide, based upon the facts in the case."[344]

Even when the judges are forced to face the hot potato through pre-trial motions or appeal, they have imposed standards that functionally keep the deciding power in the hands of the juror. For example, during summary judgment on the question of breach, courts set the bar high, giving it to the jury unless "there is no genuine dispute as to any material fact and the movant is entitled to judgment as a matter of law."[345] And once the jury has decided, the court

---

[341] Young v. Brown, 658 So. 2d 750, 753 (La. Ct. App. 1995) ("Negligence is the most common basis for tort liability."); Gregory M. Dexter, *Tort Liability for Golf Shots: Time to Reject the Recklessness Standard and Respect the Rules of Golf*, 9 DEPAUL J. SPORTS L. & CONTEMP. PROBS. 1, 41 (2012) ("[N]egligence is the most common standard for assessing liability in modern tort law . . . ."); F. Patrick Hubbard, *"Sophisticated Robots": Balancing Liability, Regulation, and Innovation*, 66 FLA. L. REV. 1803, 1819 (2014) ("Negligence is the most common form of 'fault' in tort law and is often defined by reference to a 'reasonable person,' whose conduct is, by definition, never negligent."); Arnold W. Reitze, Jr. & Marie Bradshaw Durrant, *State and Regional Control of Geological Carbon Sequestration (Part I)*, 41 ENV'T L. REP. 10348, 10373 (2011) ("Negligence is the most common cause of action in the tort system."); Eliot T. Tracz, *Half Truths, Empty Promises, and Hot Coffee: The Economics of Tort Reform*, 42 SETON HALL LEGIS. J. 311, 314 (2018) ("Negligence is the most common tort claim . . . .").

[342] *See* Jason M. Solomon, *Juries, Social Norms, and Civil Justice*, 65 ALA. L. REV. 1125, 1142–71 (2014) (analyzing the history of American tort law and arguing how a jury-centered approach to interpreting reasonableness has not been inevitable).

[343] *See* Kelley & Wendt, *supra* note 135, at 608. There are exceptions, of course, such as Holmes' much-maligned railroad crossing rule in *Baltimore & O. R. Co. v. Goodman*. 275 U.S. 66 (1927).

[344] Kelley & Wendt, *supra* note 135, at 608.

[345] *Cf.* FED. R. CIV. P. 56(a).

holds itself to the deferential standard of "clear error."[346] Even in cases where a plaintiff has violated a statute and jurors are instructed to follow the doctrine of negligence per se, courts nevertheless allow juries to disregard the instruction if they conclude that the violation is excusable after assessing that the defendant made a "reasonable" effort to comply or made a "reasonable" mistake regarding its applicability.[347] At so many junctures, the jurors maintain control of the breach determination perhaps because, well, they can't pass it to anyone else!

It is difficult to explain away legislative and judicial delegation of the RPPT to jurors as mere laziness or desk clearing. It is not obvious that having juries handle the determination spares judges any labor at all—it probably does the opposite by making jury trials more likely to occur. And it is hard to believe that judges are hoping to shift public blame to jurors, particularly when it is conceivable that jurors would still assess damages[348] and, therefore, bear much of the risk of public condemnation anyway.

If the only legislative goal were to maximize the likelihood that case *outcomes* under the RPPT correspond to the legislators' preferred state of affairs, then the hot potato strategy would make little sense. Jurors are accorded too much deference and freedom for that to be the case. Something nobler must be going on.

B. Solving the Puzzle with Structural Values

Several scholars have claimed that the prominent role that jurors play in tort law is an effort to promote the value of participatory democracy.[349] Participatory democracy is the notion that democracies require direct, individual participation by citizens in governmental decisions and policies that affect their lives.[350]

---

[346] *See, e.g.*, Furry v. United States, 712 F.3d 988, 992 (7th Cir. 2013) ("Ordinarily, breach . . . [is] reviewed for clear error.").

[347] *See* John C.P. Goldberg, *Inexcusable Wrongs*, 103 CALIF. L. REV. 467, 484–85 (2015).

[348] *See, e.g.*, Chesapeake & Ohio Ry. v. Arrington, 101 S.E. 415, 423 (Va. 1919), *abrogated by* John Crane, Inc. v. Jones, 650 S.E.2d 851 (Va. 2007) ("The law wisely leaves the assessment of damages, as a rule, to juries, with the concession that there are no scales in which to weigh human suffering . . . .").

[349] *See, e.g.*, ALBERT W. DZUR, PUNISHMENT, PARTICIPATORY DEMOCRACY, & THE JURY 14 (2012); Vikram David Amar, *Jury Service as Political Participation Akin to Voting*, 80 CORNELL L. REV. 203, 206 (1995); Joshua Kleinfeld, *Manifesto of Democratic Criminal Justice*, 111 NW. U. L. REV. 1367, 1390–91 (2017); Martin A. Kotler, *Social Norms and Judicial Rulemaking: Commitment to Political Process and the Basis of Tort Law*, 49 KAN. L. REV. 65, 98 (2000); Alexandra D. Lahav, *The Jury and Participatory Democracy*, 55 WM. & MARY L. REV. (CIV. JURY AS POL. INST. SYMP.) 1029, 1031–32 (2014); William E. Nelson, *Political Decision Making by Informed Juries*, 55 WM. & MARY L. REV. (CIV. JURY AS POL. INST. SYMP.) 1149, 1150–51 (2014).

[350] Akhil Reed Amar, *The Bill of Rights as a Constitution*, 100 YALE L.J. 1131, 1187–90 (1991).

Akhil Amar argues that the idea was highly salient during the period of our Constitution's ratification, with prominent legal thinkers characterizing the role of the jury as an opportunity for the people to have a share of control in the judiciary; they perceived it as a better opportunity for direct representation than the legislature.[351]

In recent history, courts have periodically echoed the sentiment that jurors serve this function. In *Taylor v. Louisiana*, the Supreme Court stated that broadly representative community participation in criminal juries is "not only consistent with our democratic heritage but is also critical to public confidence in the fairness of the criminal justice system."[352]

Although an interest in participatory democracy might explain why the jury is given the task of determining breach, it does not necessarily explain why the jury is given so little guidance in how to do it. Why not narrow the parameters?

One attractive possibility is that the openness of the RPPT maximizes the feeling in jurors that they are doing more than simply following a set of instructions. On this account, the minimal guidance of the RPPT turns out to be a strength. It allows jurors to be quasi-legislators within the limits set by the official legislature. So, although insofar as a jury is not allowed to usurp the legislature,[353] the RPPT gives them ample room to place their own personal stamp on the way the law is applied in a particular case.

But this solution opens another challenge: if this is about granting jurors an opportunity to influence the content of the law, then why narrow their deliberation through a norm at all? In other words, why have the RPPT rather than an even more open-ended norm, such as: "Assess whether the defendant should be liable"?

The best explanation, in my opinion, is that the modest constraining power of the RPPT addresses an additional value—namely, deliberative democracy.

---

[351] *Id.* at 1140.

[352] 419 U.S. 522, 530 (1975); *see also In re* Japanese Elec. Prods. Antitrust Litig., 631 F.2d 1069, 1093 (3d Cir. 1980) (Gibbons, J., dissenting) ("The jury is a sort of ad hoc parliament convened from the citizenry at large to lend respectability and authority to the process."); United States v. Walker, 423 F. Supp. 3d 281, 291 (S.D. W. Va. 2017) (linking jury trial in criminal cases to participatory democracy); Cerrone v. People, 900 P.2d 45, 52 (Colo. 1995) (en banc) ("The General Assembly enacted section 13-71-103 to ensure that all qualified citizens have the opportunity to serve as jurors. No doubt the legislature intended to preserve the values of participatory democracy and the public confidence in the jury system by ensuring that individuals would not be excluded from jury service on the basis of invidious criteria.").

[353] *See, e.g.*, People v. Douglas, 680 N.Y.S.2d 145, 153 (Sup. Ct. 1998) ("The jury system is an effort to secure participatory democracy; democratic theory, however, does not countenance extralegal lawmaking. Political legitimacy suggests that the jury room is an improper setting to pass judgment on the wisdom of policy choices democratically determined by duly elected representatives.").

Deliberative democracy is the demand that citizens appeal to principals when they participate in politics.[354] That is, citizens must provide reasons for their political decisions that would not be rejected by other individuals who demand fair terms of cooperation when they take part in the governance of that same society.[355] The primary appeal of policy driven by the dictates of deliberative democracy is that it drives citizens, "to hear otherwise powerless or oppressed groups and, insofar as their arguments are reasonable, to grant them influence."[356] It further has the capacity to create something like a regulated marketplace of ideas: allowing the better argument to prevail in a context unconstrained by authoritative directives other than the baseline directive that discussants provide reasons that meet the aforementioned condition.[357] Thus, it aspires to legitimize democratic rule while providing ground for sound decision-making.[358]

The RPPT provides just enough structure for juries to meet the demands of deliberative democracy. The content of the RPPT is open-ended enough to keep them from being constrained by authoritative directives beyond the bare necessities, but it provides enough of a nudge to get them into a reasoned mindset. This function is plausible. In earlier work,[359] I described the idea that legal standards could serve as "moral reminders," a phenomenon identified by Nina Mazar, On Amir, and Dan Ariely in their landmark article, *The Dishones-*

---

[354] AMY GUTMANN & DENNIS THOMPSON, WHY DELIBERATIVE DEMOCRACY? 3 (2004).

[355] *See id.* at 7 ("[W]e can define deliberative democracy as a form of government in which free and equal citizens (and their representatives), justify decisions in a process in which they give one another reasons that are mutually acceptable and generally accessible, with the aim of reaching conclusions that are binding in the present on all citizens but open to challenge in the future.").

[356] Andrew Knops, *Delivering Deliberation's Emancipatory Potential*, 34 POL. THEORY 594, 594 (2006).

[357] *See* Joshua Cohen, *Deliberation and Democratic Legitimacy*, *in* DEBATES IN CONTEMPORARY POLITICAL PHILOSOPHY: AN ANTHOLOGY 342, 347 (Derek Matravers & Jon Pike eds., 2003).

[358] There is a lot more that could be said about deliberative democracy, of course, but I hope that my brief discussion is not taken to suggest that academics uniformly agree that the concept has been widely accepted as a pillar of American democracy. They certainly do not. *See, e.g.*, RICHARD A. POSNER, LAW, PRAGMATISM, AND DEMOCRACY 107 (2003) (criticizing deliberative democracy "as purely aspirational and unrealistic as rule by Platonic guardians"). But the idea has spawned a vast literature, discussing its embodiment, or lack thereof, in American political institutions, as well as some rather ambitious proposals for reform in its spirit. *See, e.g.*, BRUCE ACKERMAN & JAMES S. FISHKIN, DELIBERATION DAY 17 (2004) (proposing a national holiday preceding presidential elections so that people can engage in structured deliberative debates); ETHAN J. LEIB, DELIBERATIVE DEMOCRACY IN AMERICA: A PROPOSAL FOR A POPULAR BRANCH OF GOVERNMENT 4 (2004) (calling for a branch of government composed of randomly selected citizens that adjudicates issues of public policy and enacts laws, subject to judicial review and possible veto by the executive and legislative branches).

[359] *See* Brian Sheppard, *Norm Supercompliance and the Status of Soft Law*, 62 BUFF. L. REV. 787, 849–56, 866 (2014).

ty of Honest People: A Theory of Self-Concept Maintenance.[360] There, the authors performed a series of experiments providing support for the notion that bringing items that have moral connotations, like the Ten Commandments, into attention can lower the probability of misconduct on subsequent, seemingly unrelated tasks.[361] The RPPT could, like other standards, serve as a moral reminder. Interpretation of the content of the RPPT might trigger a juror's "better angels," reminding them to focus upon the higher values that pertain to a case and potentially leading them to transcend their personal preferences regarding the outcome of the case.

To be sure, the RPPT does not direct jurors to do precisely what deliberative democracy demands. It does not tell them even to "act reasonably." But this is potentially a good thing, more consistent with the ethos of deliberative democracy.[362] Instead, the RPPT asks jurors to consider what "a reasonably prudent person would do;" it suggests stepping outside of the parameters of one's own identity, and it does so with a subtlety that could serve to maintain the autonomy of the juror. In the moral reminders experiment, subjects were not told that they must follow the dictates of the Ten Commandments, they were simply asked to list all that they could remember, and the effect was the same regardless of the apparent religiosity of the subject.[363] The RPPT is written such that jurors need to think in terms of *some* standard—some sort of way

---

[360] *See* Nina Mazar et al., *The Dishonesty of Honest People: A Theory of Self-Concept Maintenance*, 45 J. MKTG. RSCH. 633, 633 (2008).

[361] *Id.* at 635–36.

[362] Seana Shiffrin made a similar insight about standards, in general:

> Standards require that citizens deliberate about the moral properties of their interactions and work with more complex analogies. Because of their relative opacity, standards convey that moral reasoning requires deliberation and thoughtfulness. . . .
>
> . . . [T]he deliberation-inducing feature of standards plays a democratic role. By (partly) incorporating their purpose or rationale into their articulation, standards educate citizens about the underlying justifications and aims of law. Further, because, at least at first blush, standards do not admit of algorithmic interpretation, citizens must engage in legal interpretation by engaging with the underlying purposes of law. . . . These processes enable a richer form of democratic engagement and mutual understanding.

Shiffrin, *supra* note 296, at 1224–25.

[363] Mazar et al., *supra* note 360, at 635–36 ("The idea of the Ten Commandments recall task was that independent of people's religion, of whether people believed in God, or of whether they knew any of the commandments, knowing that the Ten Commandments are about moral rules would be enough to increase attention to their own moral standards and thus increase the likelihood of behavior consistent with these standards . . . . Note also that, on average, participants remembered only 4.3 of the Ten Commandments, and we found no significant correlation between the number of commandments recalled and the number of matrices the participants claimed to have solved correctly . . . . If we use the number of commandments remembered as a proxy for religiosity, the lack of relationship between religiosity and the magnitude of dishonesty suggests that the efficacy of the Ten Commandments is based on increased attention to internal honesty standards, leading to a lower tolerance for dishonesty (i.e., decreased self-concept maintenance threshold)." (citation omitted)).

to use real reasons to justify their position. This could increase the likelihood that they will engage in reasoned discussion rather than deciding—on the one hand, by bare, self-interested fiat or—on the other hand, from mechanical compliance with authoritative directives.

Importantly, a moral reminder function for the RPPT differs from a Moral Conception of the RPPT. The former operates to increase the rigor and sensitivity with which one interprets the RPPT; whereas the latter claims to be the one true interpretation of the RPPT and, thus, a legal command. A moral reminder gives an account of how the broad moralistic language of the RPPT could place a juror in the mindset that they must internally (or, among other jurors, in deliberation with them) articulate reasons along some sort of normative ground. It is not coercive or directed thinking in the way that a rule would be; therefore, it preserves autonomy, at least in principle. Thus, the RPPT's role does not turn on a particular understanding of the RPPT's meaning; it could very well lead a juror to adopt a mechanical approach to breach, such as the Average Conduct Conception.

I am not the first to link juries to deliberative democracy. Jeffrey Abramson, in his impactful 1994 book, *We, the Jury*, put forth the notion, stating, "Surely the jury has not survived all these centuries only to teach us that democracy is about brokering justice among irreconcilably antagonistic groups. I will argue for an alternative view of the jury, a vision that defends the jury as a deliberative rather than a representative body."[364]

It is fair to question whether the intention behind the development of the RPPT, insofar as there was one at all, was to effectuate the two values of participatory and deliberative democracy. I suspect not. Were the goal to get jurors to offer reasoned, principled bases for their decisions, there are more direct methods than using the RPPT. American courts could follow the example of the continental mixed courts[365] and require that jurors individually or collectively offer explanations for their decisions. The bare fact that structural dimensions of the RPPT might be explained by democratic values does not mean that, as a historical matter, it was why the RPPT took on the shape it now has. But this is beside the point: what matters is whether the RPPT *actually performs* this function and not whether it was designed to perform it.

---

[364] JEFFREY ABRAMSON, WE, THE JURY: THE JURY SYSTEM AND THE IDEAL OF DEMOCRACY 8 (First paperback ed., Harvard University Press 2000) (1994); *see also* Deborah Ramirez, *Affirmative Jury Selection: A Proposal to Advance Both the Deliberative Ideal and Jury Diversity*, 1998 U. CHI. LEGAL F. 161, 161–77 (presenting an alternative to Abramson that also conceives of the jury as a vehicle for deliberative democracy but includes a role for representation).

[365] Douglas G. Smith, *Structural and Functional Aspects of the Jury: Comparative Analysis and Proposals for Reform*, 48 ALA. L. REV. 441, 495 (1997).

There is reason to believe that it does. Feigenson observed in his field study that juries seek "total justice," in which they exhibit more concern for "making things come out right than with strictly following the relevant legal rules[,]" but their everyday habits of judgment and care usually yield results that are consistent with those of legal experts.[366] Although the reality sounds messier than the ideal, this supports the notion that the spirit of these democratic values is often embodied in the deliberations of actual juries. They care about justice, but they form their own pathways towards it. As mentioned, multiple studies provide evidence that the sophistication of moral reasoning influences jury outcomes.[367]

Not everyone agrees. Jason Solomon is skeptical that jury deliberation embodies this ideal,[368] citing evidence that juries suffer from uneven participation that favors already-advantaged groups, such as males and people with high occupational status.[369] Solomon is also concerned with erratic punitive damages assessments from juries, linking those results to problems in deliberation.[370]

The empirical literature regarding those awards, however, provides indirect support for the position that jurors engage in principled deliberation under moral reminders. Punitive damages awards are a useful parallel to breach determinations. They are governed by a standard that, like the RPPT, incorporates moralistic language such as, "[p]unitive damages may be awarded for conduct that is outrageous, because of the defendant's evil motive or his reckless indifference to the rights of others."[371] The landmark book, *Punitive Damages: How Juries Decide*, praises juror conduct under the punitive damages standard:

> Throughout the present research, we have been impressed by the serious and energetic manner in which citizens performed the difficult legal judgment tasks that are demanded by the punitive damages decision. The many systematic patterns of behavior that we observed are convincing evidence of the jurors' conscientiousness. Nonetheless, the legally required decision tasks often seemed to exceed their individual and social capacities. The decision task is not well defined by the jury

---

[366] FEIGENSON, *supra* note 173, at 16–18, 104–11.

[367] Bernard et al., *supra* note 321, at 97–98; Rotenberg et al., *supra* note 322, at 172–73.

[368] Solomon, *supra* note 342, at 1186–87.

[369] *Id.* at 1185 (first citing Erin York Cornwell & Valerie P. Hans, *Representation Through Participation: A Multilevel Analysis of Jury Deliberations*, 45 L. & SOC'Y REV. 667, 671 (2011); then citing Erin York & Benjamin Cornwell, *Status on Trial: Social Characteristics and Influence in the Jury Room*, 85 SOC. FORCES 455, 469 (2006)).

[370] *Id.* at 1128–29.

[371] RESTATEMENT (SECOND) OF TORTS § 908(2) (AM. L. INST. 1979); *see also* Daniel A. Crane, *Rules Versus Standards in Antitrust Adjudication*, 64 WASH. & LEE L. REV. 49, 91 n.170 (2007) (describing punitive damages instructions as relying on "open-ended" standards).

> instructions[, and] jurors are not provided with the necessary background information or experiences to make reliable judgments . . . .[372]

This finally brings us to an unavoidable cost of any approach to the automation of the RPPT: the value of the RPPT as a pathway for participatory and deliberative democracy, such as it is, would disappear if machines were to take on the role of deciding breach. There is no way to automate a function completely without taking human beings out of that function, regardless of the advancement of the technology.

Without knowing the particular interests of those who might consider automation, it is nevertheless important to understand the different ways that technology will enhance or diminish public values. There are no guarantees that a legislator would care about the RPPT's role in bringing about participatory or deliberative democracy. But for the legislator who *does* care about those values, automation of the RPPT eliminates one avenue for their realization.[373]

CONCLUSION

I am here neither to praise automation, nor to condemn it. Instead, I have taken a humbler role: I have attempted to map the vectors of cost and benefit if legislators someday seek to computerize a difficult but significant area of law. In doing so, I have flagged that legally plausible interpretations of marginal significance in tort law might, for a time, become more attractive because they are friendly to computerization. I have also provided reasons to doubt that certain interpretations will serve as the basis for automation. Lastly, I have tried to show that, although the Reasonably Prudent Person Test's interpretive malleability plays a crucial role in setting the cost of legal automation, there are structural aspects of negligence law beyond the text of the test that matter. Specifically, removing jurors from the determination of breach could extinguish celebrated opportunities for participatory and deliberative democracy. The value of those opportunities may be doubted, but we cannot question their essence: they are unavoidably human.

---

[372] CASS R. SUNSTEIN ET AL., PUNITIVE DAMAGES: HOW JURIES DECIDE 241 (2002).

[373] There might be competing interests that trump the democratic values in the eyes of that legislator, of course. *See* Kotler, *supra* note 349, at 134. But the elimination of the jury's role in breach should remain a factor of significance.