Individualized Suspicion in the Age of Big Data

Emily Berman[*]

ABSTRACT: Imagine that an algorithmic computer model known to be 80 percent accurate predicts that a particular car is likely to be transporting drugs. Does that prediction provide law enforcement probable cause to search the car? When generated by humans, courts have consistently regarded such evidence of statistical likelihood as insufficiently individualized to satisfy even the most permissive legal standards—a position that has generated decades of debate among commentators. The proliferation of artificial-intelligence-generated predictions—predictions that will be more accurate than humans' and therefore more tempting to employ—requires us to revisit this debate over use of probabilistic evidence with renewed urgency, and to consider its implications for the use of predictive algorithms. This Article argues that reliance on probabilistic evidence to establish the individualized suspicion required by the Fourth Amendment, regardless of that evidence's statistical accuracy—i.e., how likely it is that the predictions of criminal activity are correct—disregards fundamental interests that individualized suspicion is meant to protect, namely respect for human dignity, preservation of individual autonomy, and guarantees of procedural justice. So while accuracy is a necessary element of individualized suspicion findings, this Article contends that no level of statistical likelihood is sufficient. Further, it argues that careful consideration of these issues has become critically important in today's big data world, because the shortcomings that "analog" probabilistic evidence presents are even more pronounced in the context of predictive algorithms.

I. INTRODUCTION 464

II. INDIVIDUALIZED SUSPICION IN THEORY 471
  A. *THE INDIVIDUALIZED SUSPICION REQUIREMENT* 471

[*] Associate Professor of Law, University of Houston Law Center. The author would like to thank Dave Fagundes, Aziz Huq, James Nelson, and D. Theodore Rave, as well as participants in the University of Houston Law Center's faculty workshop, the Chapman University School of Law Junior Faculty Works-in-Progress Conference, and the Boston College Law School Junior Faculty Roundtable for helpful comments.

B. *INDIVIDUALIZED SUSPICION'S INCOHERENCE* 474

III. INDIVIDUALIZED SUSPICION AND PREDICTIVE ACCURACY: WHY PREDICTIVE ACCURACY IS NECESSARY BUT NOT SUFFICIENT 478
A. *THE PURPOSE OF THE INDIVIDUALIZED SUSPICION REQUIREMENT* 478
B. *PREDICTIVE ACCURACY AS A PROXY FOR INDIVIDUALIZED SUSPICION* 482
C. *THE ILLUSORY CONSTRAINING POWER OF NUMERICAL THRESHOLDS* 485
D. *THE CRUCIAL ROLE OF NON-PROBABILISTIC EVIDENCE* 487

IV. INDIVIDUALIZED SUSPICION AND ALGORITHMIC DECISION-MAKING 495
A. *ALGORITHMIC DECISION-MAKING* 496
B. *ALGORITHMS & INDIVIDUALIZED SUSPICION* 500

V. CONCLUSION 505

I. INTRODUCTION

Imagine that a law enforcement officer stops a vehicle because it rolled through a stoplight. When the officer runs the license plate through the computer in her squad car, it informs her that an algorithmic computer model predicts that the car is likely to be transporting illicit drugs. The predictive model is known to be accurate 80 percent of the time. Based on this information, the officer looks in the car's trunk—a search for which the Fourth Amendment requires probable cause. The probable cause standard is met when, based on "the factual and practical considerations of everyday life on which reasonable and prudent men . . . act," there is "a 'substantial basis for . . . conclud[ing]' that a search would uncover evidence of wrongdoing."[1] The search reveals the predicted drugs. Has the officer violated the Fourth Amendment, or does the highly accurate computer model's prediction satisfy the probable cause requirement?

This hypothetical presents a modern twist on an old debate about the role of statistical evidence in the legal system, which asks whether and how the law should treat purely probabilistic evidence.[2] Some scholars have long

1. Illinois v. Gates, 462 U.S. 213, 231, 236 (1983) (alteration in original) (quoting Jones v. United States, 362 U.S. 257, 271 (1960)). Probable cause requires "only the probability, and not a prima facie showing, of criminal activity." *Id.* at 235 (quoting Spinelli v. United States, 393 U.S. 410, 419 (1969)).

2. By probabilistic or statistical evidence, I refer to evidence that provides a statistical likelihood that some fact is true. The literature on questions related to statistical evidence is extensive, addresses both criminal-law and civil-law topics, and spans five decades. *See generally,*

argued that such evidence should be more widely employed.[3] Courts, however, have consistently characterized such information as too "generalized," insisting that "case-specific" or "individualized" evidence is required to satisfy even the most permissive legal standards,[4] notwithstanding the fact that exactly what it means for evidence to be "individualized," as opposed to generalized, has proven exceedingly difficult to articulate.[5]

This decades-old debate intersects with the contemporary discussion regarding the role of artificial intelligence in the legal arena, and, in particular, the question of whether and when it is appropriate to entrust legal decision-making to algorithms and computer models. Because such models have been implemented in numerous decision-making contexts already,[6] this

---

e.g., L. JONATHAN COHEN, THE PROBABLE AND THE PROVABLE (1977) (discussing the use of probability in the judicial process); FREDERICK SCHAUER, PROFILES, PROBABILITIES, AND STEREOTYPES (2003) (arguing that statistical evidence should be considered in the legal system); David Kaye, *The Paradox of the Gatecrasher and Other Stories*, 1979 ARIZ. ST. L.J. 101 (discussing the value of probability estimates); Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357 (1985) (exploring how probability affects public acceptance of jury verdicts); Charles R. Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187 (1979) [hereinafter Nesson, *Reasonable Doubt*] (discussing the ability of statistical evidence to assist in quantifying reasonable doubt); Michael S. Pardo, *The Paradoxes of Legal Proof: A Critical Guide*, 99 B.U. L. REV. 233 (2019) (challenging the view that standards of proof represent probabilistic thresholds and exploring the implications of that argument for civil litigation and criminal procedure); Mike Redmayne, *Exploring the Proof Paradoxes*, 14 LEGAL THEORY 281 (2008) (discussing the paradoxes that emerge if standards of proof are conceived of merely in terms of probability); Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971) (identifying dangers of using mathematical models in the legal process).

3. *See generally, e.g.*, SCHAUER, *supra* note 2 (arguing for using statistically sound generalizations); Ronald J. Bacigal, *Making the Right Gamble: The Odds on Probable Cause*, 74 MISS. L.J. 279, 295–304 (2004) (arguing that statistical evidence is just as valid as any other form of evidence).

4. *See* Jane Bambauer, *Hassle*, 113 MICH. L. REV. 461, 462 (2015) (conceding that a judge would reject a warrant application to search a home based on a statistical study indicating that 60 percent of the homes in that neighborhood have illicit drugs in them).

5. For example, is a reliable study showing that 60 percent of State College dorm rooms contain drugs individualized, because it specifically applies only to State College dorm rooms, or is it generalized, because it refers to all of the dorm rooms on the State College campus? *See id.* at 462–63 (positing a version of this hypothetical); *see also infra* Section III.B for a full discussion of the difficulties of distinguishing between generalized and individualized evidence.

6. Models currently decide whether arrestees should be granted bail, see Lauryn P. Gouldin, *Disentangling Flight Risk from Dangerousness*, 2016 BYU L. REV. 837, 867–71; determine who is eligible for a loan, see Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 18–30 (2014); identify credit card fraud, see Steven Melendez, *Algorithms Honed on Stock Trades Are Fighting Credit Card Fraud*, FAST COMPANY (Sept. 14, 2017), https://www.fastcompany.com/40465736/how-machine-learning-is-helping-cut-credit-card-fraud-cashshield [https://perma.cc/7V2Z-Y3PC]; suggest products to consumers, see Shauna Mei, *A.I. Can Help Us Make Quicker, Better Consumer Choices*, N.Y. TIMES (Dec. 5, 2016, 3:21 AM), https://www.nytimes.com/roomfordebate/2016/12/05/is-artificial-intelligence-taking-over-our-lives/ai-can-help-us-make-quicker-better-consumer-choices [https://perma.cc/Q8T9-87FA]; make medical diagnoses, see Cade Metz, *A.I. Shows Promise Assisting Physicians*, N.Y. TIMES (Feb. 11, 2019), https://www.nytimes.com/2019/02/11/health/artificial-intelligence-medical-

question has sparked vigorous debate about a variety of issues.[7] And while questions regarding when artificial intelligence should be employed and what safeguards must accompany it remain hotly contested, there is no dispute that its role in the legal system will only continue to grow.[8]

This proliferation of algorithmic predictions requires us to revisit the debate over use of probabilistic evidence with renewed urgency. This Article addresses the propriety of using such predictions in one specific context: the Fourth Amendment's individualized-suspicion requirements governing arrests, searches, and investigative stops.[9] It argues that statistical evidence,

---

diagnosis.html [https://perma.cc/E5W6-MVNL]; recommend where law enforcement resources should be deployed, see Chris Strohm, *Predicting Terrorism from Big Data Challenges U.S. Intelligence*, BLOOMBERG (Oct. 13, 2016, 4:00 AM), https://www.bloomberg.com/news/articles/2016-10-13/predicting-terrorism-from-big-data-challenges-u-s-intelligence [https://perma.cc/8XMG-JD9T]; and much more. *See also generally* Emily Berman, *A Government of Laws and Not of Machines*, 98 B.U. L. REV. 1277 (2018) (identifying multiple uses of predictive analytics in the national-security and law-enforcement contexts).

7. *See, e.g.*, BRUCE SCHNEIER, DATA AND GOLIATH: THE HIDDEN BATTLES TO COLLECT YOUR DATA AND CONTROL YOUR WORLD 273 (2015) (noting balancing of benefits to society versus costs to individuals involved in use of big data "is done for us by governments and corporations with their own agendas"); Steven M. Bellovin et al., *When Enough Is Enough: Location Tracking, Mosaic Theory, and Machine Learning*, 8 N.Y.U. J.L. & LIBERTY 556, 621–24 (2014) (describing possible violations of reasonable expectation of privacy from law enforcement's use of machine learning in location tracking); Kiel Brennan-Marquez, *"Plausible Cause": Explanatory Standards in the Age of Powerful Machines*, 70 VAND. L. REV. 1249, 1255–57 (2017) (arguing that the use of algorithms in law enforcement's decision-making threatens traditional criminal justice system values); Citron & Pasquale, *supra* note 6, at 4 (warning that "[b]ecause human beings program predictive algorithms, their biases and values are embedded into the software's instructions"); Andrew Guthrie Ferguson, *Predictive Policing and Reasonable Suspicion*, 62 EMORY L.J. 259, 313–25 (2012) (discussing the primary constitutional concerns implicated by predictive policing); Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*, 10 HARV. L. & POL'Y REV. 15, 17 (2016) (discussing the implications that arise from increased police surveillance made possible by big data tools); Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 695–96 (2017) (emphasizing the need for transparency and accountability in public policy decision-making based on algorithms); Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 U. PA. L. REV. 871, 925 (2016) (warning that inaccurate data and human error are two possible causes of error in using algorithms to automate findings of suspicion); Ric Simmons, *Quantifying Criminal Procedure: How to Unlock the Potential of Big Data in Our Criminal Justice System*, 2016 MICH. ST. L. REV. 947, 950 (discussing the challenges to incorporating big data tools into the criminal justice system effectively); Tal Z. Zarsky, *Transparent Predictions*, 2013 U. ILL. L. REV. 1503, 1506 (stating that law enforcement's use of predictive practices based on the analysis of personal information and data mining may result in biased, discriminatory processes that threaten privacy and autonomy).

8. *See, e.g.*, sources cited *supra* note 7.

9. To perform a search or make an arrest, law enforcement must establish probable cause, a standard that is satisfied when "the factual and practical considerations of everyday life on which reasonable and prudent men . . . act" provide a "'substantial basis for . . . conclud[ing]' that a search would uncover evidence of wrongdoing." Illinois v. Gates, 462 U.S. 213, 231, 236 (1983) (alteration in original) (quoting Jones v. United States, 362 U.S. 257, 271 (1960)). To make brief investigative stops, law enforcement must have reasonable suspicion, which requires "specific and articulable facts which, taken together with rational inferences [drawn] from those facts," would "warrant a [person] of reasonable caution in the belief" that the search was justified. Terry v.

irrespective of its accuracy in predicting the presence of criminal activity or the means by which it is generated, should not form the basis of individualized-suspicion determinations.[10] For while individualized suspicion is certainly intended as a tool to minimize erroneous investigative decisions—i.e., the stop, search, or arrest of innocent individuals—it also serves the less concrete, though no less fundamental, interests of safeguarding human dignity, preserving individual autonomy, and guaranteeing procedural justice. This Article contends that using probabilistic evidence undermines these interests. Moreover, it argues, this troubling aspect of probabilistic evidence is intensified when the predictions at issue are generated by algorithms. So while a certain level of predictive accuracy is certainly a *necessary* element of an individualized-suspicion finding, the primary contention of this Article is that no level of statistical likelihood is *sufficient* to meet that requirement.

The need for renewed examination of the debate over probabilistic evidence is particularly pressing in today's big data world for at least two reasons. First, while law enforcement may not yet be relying solely on statistical evidence to select targets of searches, arrests, or stops, the idea that they soon might is far from outlandish.[11] Already law enforcement uses statistical evidence in the form of predictive algorithms to determine things such as where to deploy personnel,[12] who is likely to be involved in gun violence,[13] and who should be freed on bail.[14] These predictive tools are generated through a form of artificial intelligence known as machine learning, which uses algorithms to analyze big data—enormous data sets—to develop

Ohio, 392 U.S. 1, 21–22 (1968) (quoting Carroll v. United States, 267 U.S. 132, 162 (1925)); *see also infra* Section II.A.

10. A handful of commentators have reached a similar conclusion regarding the use of predictive algorithms, but by narrowly focusing on the machine-learning context (rather than recognizing that the debate over use of algorithmic models is properly viewed as simply the latest version of the age-old debate over the use of probabilistic evidence), their analyses neither grapple with strong counterarguments that emerge from that larger debate, nor fully explore exactly why it is that statistical accuracy should not be considered sufficient to establish individualized suspicion. *See* Ferguson, *supra* note 7, at 262, 296–97; Rich, *supra* note 7, at 871, 893–900; Simmons, *supra* note 7, at 969–83. This Article avoids these pitfalls, situating the use of machine learning algorithms in the larger context of probabilistic evidence and explaining why, despite their potential for increased accuracy in decision-making, reliance on them remains inconsistent with a meaningful individualized-suspicion requirement.

11. *See* Rich, *supra* note 7, at 883 (arguing that law enforcement is likely to use predictive algorithms in this way in the near future).

12. *See* Simmons, *supra* note 7, at 954–58 (discussing various predictive policing programs).

13. *See* Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. 327, 384–85 (2015).

14. *See* Gouldin, *supra* note 6, at 837. Machine-learning algorithms are predictive algorithms that can analyze data regarding their own performance and make themselves more accurate as a result—in other words, they learn. *See* Berman, *supra* note 6, at 1285 (citing PETER FLACH, MACHINE LEARNING: THE ART AND SCIENCE OF ALGORITHMS THAT MAKE SENSE OF DATA 3 (2012)).

computer models that make predictions about the future.[15] While these predictive models are far from infallible—indeed, the literature addressing machine learning's weaknesses and potential flaws is substantial[16]—they are able to identify relationships among data that humans would never recognize.[17] As a result, they represent a predictive tool that promises to be more accurate than human decision-making, and one that is already being deployed in numerous contexts, including marketing, fraud detection, and medical diagnoses, just to name a few.[18] It is not a huge leap from there to a world where law enforcement uses algorithms to predict things like which individuals might be involved in criminal behavior. Thus, the predictive power of algorithms—and the resulting temptation to employ—renders the debate over the role of probabilistic evidence highly salient once again.

Second, basing individualized-suspicion determinations on predictive analytics follows from the view of many commentators that individualized-suspicion determinations should be tied to predictive evidence. While there are many variations around this theme, the core of these proposals is that the individualized suspicion requirement should be deemed met so long as the evidence available generates a sufficient likelihood that the proposed stop, search, or arrest will uncover criminal activity.[19] In all of these proposals, the

---

15. *See* JOHN D. KELLEHER ET AL., FUNDAMENTALS OF MACHINE LEARNING FOR PREDICTIVE DATA ANALYTICS 1 (2015) (describing predictive data analytics); *see also infra* Section IV.A (describing the machine-learning process). Systems that use computer models to make predictions are also sometimes labeled "predictive analytics," "predictive algorithms," or "predictive models." *See* KELLEHER ET AL., *supra*, § 1.1; David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653, 671 (2017).

16. *See, e.g.*, sources cited *supra* note 7; *see also, e.g.*, Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 714–29 (2016) (pointing out possible discriminatory results of machine-learning models); Jenna Burrell, *How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms*, BIG DATA & SOC'Y, Jan.–June 2016, at 1, 1–10 (identifying ways in which machine-learning algorithms are not transparent); Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1308 (2008) (arguing that we need a new concept of technological due process to enhance the transparency, accountability, and accuracy of rules embedded in automated decision-making systems); Ferguson, *supra* note 13, at 388–404 (reviewing positives and negatives of using data in traffic stops); Rich, *supra* note 7, at 871, 893–900 (explaining pitfalls of applying machine-learning methods to "government data with the purpose of identifying individuals likely to be engaged in criminal activity"); Simmons, *supra* note 7, at 969–83 (discussing risks inherent to predictive algorithms, such as racial bias, the use of forbidden factors like national origin, and preexisting biases in underlying data); Zarsky, *supra* note 7, at 1506 (stating use of predictive practices based on analysis of personal information and data mining by law enforcement may result in biased, discriminatory processes threatening privacy and autonomy).

17. *See* STUART RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH 807 (3rd ed. 2010).

18. *See* sources cited *supra* note 6.

19. Some seek to fix numerical thresholds for probable cause and reasonable suspicion. *See, e.g.*, Bacigal, *supra* note 3, at 338 (suggesting a "tiered model of the levels of certainty required for searches and seizures"); Erica Goldberg, *Getting Beyond Intuition in the Probable Cause Inquiry*, 17 LEWIS & CLARK L. REV. 789, 794 (2013); Simmons, *supra* note 7, at 999–1016; *see also infra*

crucial factor remains predictive accuracy—whether the probability of detecting criminal activity is sufficiently high (however that is defined) to justify government action.

These scholars' embrace of statistical evidence springs in large part from valid concerns regarding conscious or unconscious bias that historically have infected law enforcement decision-making. Individualized suspicion's current, nebulous definition confers an enormous amount of discretion on government officials and judges to determine when they believe the necessary threshold has been met.[20] Proposals based on predictive accuracy see the use of statistical evidence as a tool to cabin this discretion, thereby combating improper bias. At first blush, then, the arguments supporting use of probabilistic evidence appear convincing—they seem to promise better crime detection, fewer intrusions on the privacy of innocent individuals, and reduced bias in law enforcement decision-making.

The problem with arguments focused on the benefits of using probabilistic evidence is that they fail to take into account its costs. A closer examination of those potential costs reveals that any theory of individualized suspicion that focuses on predictive accuracy cannot vindicate the purposes of the individualized-suspicion requirement. Consider one reason for the requirement: limiting government discretion. It is perfectly reasonable to limit law enforcement discretion by denying the government the authority to perform stops, searches, or arrests until there is reason to believe such an action is warranted. Unfortunately, an individualized-suspicion requirement based on the accuracy of predictive tools will struggle to limit state discretion, especially when it comes to reasonable-suspicion determinations. Computer modeling and other analytical methods combine with the government's

---

notes 114–18 and accompanying text. Others contend that above a certain threshold of likelihood, the targets of stops or searches should be selected at random. For example, if there are 10 cars, each of which has a 35 percent likelihood of containing contraband in the trunk, law enforcement should randomly select which among those ten to search. *See* Bernard E. Harcourt & Tracey L. Meares, *Randomization and the Fourth Amendment*, 78 U. CHI. L. REV. 809, 834 (2011); *see also infra* notes 70–72, 92–95 and accompanying text. Perhaps the most common means of more precisely defining individualized suspicion is to see it as an exercise in cost-benefit analysis. A decision is sufficiently individualized, some scholars argue, when the benefits of acting on the information (e.g., conducting a search) outweigh the costs of the inevitable errors (e.g., searching an innocent person's home). *See, e.g.*, SCHAUER, *supra* note 2, at 126; Bambauer, *supra* note 4, at 465; Andrew E. Taslitz, *What is Probable Cause, and Why Should We Care?: The Costs, Benefits, and Meaning of Individualized Suspicion*, 73 L. & CONTEMP. PROBS. 145, 154 (2010); *see also infra* notes 73, 82, 96–105 and accompanying text. In other words, evidence that guarantees that the likelihood of true positives is sufficiently high and the likelihood of false positives is sufficiently low should be considered individualized, though the acceptable rates of true and false positives might vary depending on the gravity of the crime being investigated, see Taslitz, *supra*, at 172, or the intrusiveness of the government action in question, see Bambauer, *supra* note 4, at 486.

20. *See infra* Part II (discussing the difficulties in articulating a clear definition of "individualization").

ability to access massive amounts of data to allow government officials to develop statistical suspicion regarding vast swaths of the population.[21]

More importantly, however, the individualized-suspicion requirement furthers the additional, though perhaps less tangible, fundamental values of dignity, autonomy, and procedural justice. These goals will go under-protected if probabilistic accuracy becomes the touchstone of individualized suspicion. What preserves those values is not only whether a decision-maker ultimately gets the individualized-suspicion determination "right," but also the procedural mechanisms through which the determination is made. These mechanisms require something more than simple numerical arguments. Probabilistic evidence denies the procedural outlets for factual explanation and legal justification that are so critical to fulfilling the individualized-suspicion requirement's purposes. When relying on probabilistic evidence, the mechanisms that recognize individuals' uniqueness and provide them the opportunity be heard are diluted.[22] This, in turn, denies affected individuals the opportunity to meaningfully challenge the government's proffered justification and to insist that the individualized-suspicion requirement, however defined, has not been met *as to them.*

This is true with respect to all probabilistic evidence, but it is particularly salient in the predictive analytics context. Imagine, for example, that an algorithm analyzes a vast database of information about the past behavior of college students and generates a model that identifies with a predictable accuracy rate those individuals who are likely to have marijuana in their dorm rooms, thereby providing a basis on which to search these rooms. The model will make no causal claims. Rather it will identify correlations and indicate that a certain combination of characteristics, when present, predict at some rate of statistical probability that an individual is a marijuana user. If the prediction leads the government to search someone's dorm room, it will not be due to anything unique to that individual but rather to the fact that an algorithm has identified a specific combination of various characteristics he possesses or activities he has engaged in—perhaps things such as his age, his grades, or any other feature that the programmer provides and the algorithm finds relevant—as indicative of marijuana use. And because machine learning algorithms are nearly always unintelligible to humans,[23] the result is an even more mechanical, impersonal process than traditional probabilistic evidence, intensifying the concerns such evidence raises.

---

21. *See infra* Section IV.A.

22. Some commentators suggest that any deficiencies of this kind can be remedied by ensuring that a human remains the final decision-maker—what is sometimes called a "human-in-the-loop" requirement. *See, e.g.*, Ferguson, *supra* note 7, at 314–31; Simmons, *supra* note 7, at 1013–16. As Section IV.B *infra* will briefly note, however, this is not a cure-all.

23. *See, e.g.*, RUSSELL & NORVIG, *supra* note 17, at 707; *see also, e.g.*, FLACH, *supra* note 14, at 32; Berman, *supra* note 6, at 1288–90.

The Article proceeds in three parts. Part II will address the concept of individualized suspicion, identifying its doctrinal requirements and demonstrating that, as a conceptual matter, it fails to provide decision-makers with a clear idea of when those requirements are met. Part III will begin in Section III.A by presenting some of the proposed definitions of individualized suspicion that rest on the accuracy of predictive mechanisms. Section III.B will then demonstrate how the definitions provided in Section III.A cannot satisfy the underlying values from which individualized-suspicion requirements spring. Part IV will then turn to the modern means of predicting outcomes: algorithmic modeling. Section IV.A will briefly explain the process, with an emphasis on factors relevant to Section IV.B's argument that any use of the probabilistic output of computer models will not only share the flaws of the accuracy-based definitions discussed in Part III, but also will magnify them.

II. INDIVIDUALIZED SUSPICION IN THEORY

There are some actions that the government cannot undertake absent individualized evidence. In the Fourth Amendment context, this limitation manifests as the need to establish probable cause or reasonable suspicion. Section II.A will sketch out the Fourth Amendment's doctrinal requirements for individualized suspicion, and Section II.B will then explain why the doctrinal tests the courts have articulated actually provide very few definitive answers regarding when a government action is sufficiently individualized. Thus, individualized suspicion simultaneously constitutes a critical element underpinning much of Fourth Amendment law yet remains incoherent as a concept.

A. The Individualized Suspicion Requirement

Before intruding into protected spaces—your body, your home, your car's trunk—the government must establish individualized suspicion that rises to a particular standard.[24] For full-blown searches, this standard is probable cause, which exists when, based on "the factual and practical considerations of everyday life on which reasonable and prudent men, not legal technicians, act," there is a "'substantial basis for . . . conclud[ing]' that a search would

24. The Supreme Court has recognized a set of narrow exceptions to this rule when the police are acting not in the course of regular law enforcement but instead for some special need, see Skinner v. Railway Labor Executives' Ass'n, 489 U.S. 602, 620–21 (1989) (upholding mandatory drug testing for railroad employees); when interacting with a group that lacks full privacy protections, see Samson v. California, 547 U.S. 843, 857 (2006); or when conducting searches for which they have gotten consent, see Schneckloth v. Bustamonte, 412 U.S. 218, 247–49 (1973). In such circumstances, individualized suspicion usually "must be replaced with measures to protect against the state actor's unfettered discretion." *Samson*, 547 U.S. at 860–61 (Stevens, J., dissenting).

uncover evidence of wrongdoing."[25] For less intrusive actions, such as a "brief, investigatory stop," it is sufficient that an officer has "reasonable, articulable suspicion that criminal activity is afoot."[26]

The Supreme Court has repeatedly explained that, to establish either probable cause or reasonable suspicion, the government must show not only sufficient indicia of suspicion generally, but also that the available evidence points to criminal activity specific to the person being stopped or searched.[27] As the Court has clarified, suspicion is not sufficiently particularized unless the available evidence suffices to "raise a suspicion that the *particular individual* being stopped is engaged in wrongdoing."[28]

In satisfying this requirement, the government may "not base its judgments on stereotypes, assumptions, guilt-by-association, or other generalities."[29] This means that, "a person's mere propinquity to others independently suspected of criminal activity does not, without more, give rise to probable cause to search that person."[30] As one scholar puts it, "'[p]robable cause and reasonable suspicion require more than demographic probabilities. There must be something specific to the defendant to create the probability as to him (perhaps a furtive gesture, an informant's tip, excessive nervousness, etc.).'"[31] This means that, even if law enforcement knows that

25. Illinois v. Gates, 462 U.S. 213, 231, 236 (1983) (alternation in original) (first quoting Brinegar v. United States, 338 U.S. 160, 176 (1949); then quoting Jones v. United States, 362 U.S. 257, 271 (1960)).

26. Illinois v. Wardlow, 528 U.S. 119, 123 (2000).

27. Safford Unified Sch. Dist. No. 1 v. Redding, 557 U.S. 364, 376 (2009) (holding the body search of a student unconstitutional because there was no indication that this particular student was carrying pills in her underwear and "general background possibilities [that students often hide contraband in their underwear] fall short"); Maryland v. Pringle, 540 U.S. 366, 371 (2003) ("'[T]he substance of all the definitions of probable cause is a reasonable ground for belief of guilt,' and that the belief of guilt must be particularized with respect to the person to be searched or seized." (quoting *Brinegar*, 338 U.S. at 175)); United States v. Cortez, 449 U.S. 411, 418 (1981) (noting that particularized suspicion contains two elements, the first measuring suspicion and the second, ensuring that the suspicion raised is "the particular individual being stopped is engaged in wrongdoing"); Ybarra v. Illinois, 444 U.S. 85, 91 (1979) ("Where the standard is probable cause, a search or seizure of a person must be supported by probable cause particularized with respect to that person."); Wong Sun v. United States, 371 U.S. 471, 479 (1963) (requiring not only "reliability" but also "particularity" in evidence supporting probable cause); Tracey Maclin, *The* Pringle *Case's New Notion of Probable Cause: An Assault on* Di Re *and the Fourth Amendment*, 2003 CATO SUP. CT. REV. 395, 409–12 (2003–2004) (noting that the history of the Fourth Amendment supports the need for individualized suspicion).

28. *Cortez*, 449 U.S. at 418 (emphasis added); *see also* David A. Harris, *Particularized Suspicion, Categorical Judgments: Supreme Court Rhetoric Versus Lower Court Reality Under* Terry v. Ohio, 72 ST. JOHN'S L. REV. 975, 983–84 (1998) ("[T]his demand for specificity . . . is the central teaching of this Court's Fourth Amendment jurisprudence." (quoting Terry v. Ohio, 392 U.S. 1, 21 n.18 (1968))).

29. Taslitz, *supra* note 19, at 146.

30. *Ybarra*, 444 U.S. at 91.

31. Ferguson, *supra* note 7, at 299 (quoting Arnold H. Loewy, *Rethinking Search and Seizure in a Post-9/11 World*, 80 MISS. L.J. 1507, 1518 (2011)); *id.* at 298 ("Lower courts have upheld arrest warrants on DNA matches and other forensic science matches based on pure probabilities,

nine out of ten men on one block of Main St. are selling illegal drugs, that 90 percent likelihood of finding contraband does not, without more, establish reasonable suspicion or probable cause to stop or search any one of them.[32] Some commentators have argued in favor of quantifying doctrinal principles like probable cause,[33] but the Supreme Court has resisted the call to fix a "numerically precise degree of certainty."[34]

A defendant may generate individualized suspicion about herself by engaging in legal—though perhaps suspicious—behavior, such as wearing a heavy coat on a hot day. Or she may qualify through a combination of entirely innocuous actions—such as purchasing a plane ticket in cash to travel to Miami without checking luggage.[35] But in each instance, finding the requisite suspicion requires "that the police officer must observe conduct that gives her some reason to believe that the suspect is currently engaging in criminal activity"—it must be based on a person's actions—rather than "on who the person is."[36]

This is true of reasonable suspicion as well as probable cause. While reasonable suspicion is a lower bar to meet than probable cause, it requires "more than an 'inchoate and unparticularized suspicion or "hunch."'"[37] Rather, an officer must be able to articulate reasons, based on experience and observations, that give rise to suspicion. The issue, according to the Supreme Court is "whether the officer could 'point to specific and articulable facts which, taken together with rational inferences [drawn] from those facts,' would 'warrant a [person] of reasonable caution in the belief' that the action taken was appropriate."[38]

To be sure, the Court has handed down opinions whose individualized suspicion analysis appears anemic or conclusory.[39] Indeed, commentators

---

but there has never been a Supreme Court case in which the probability of crime explicitly has been used as the sole justification for a stop.").

32. In *Wong Sun v. United States*, 371 U.S. 471, 473 (1963), government agents arrested a man for narcotics trafficking based on information from an informant that "Blackie Toy," who ran a laundry on Leavenworth Street, was selling heroin. The agents executed a warrantless arrest of the proprietor of "Oye's Laundry" on Leavenworth, James Wah Toy. *Id.* at 474. The Supreme Court determined that the officers lacked probable cause to arrest Toy, because the information provided by the informant "merely invited the officers to roam the length of Leavenworth Street . . . in search of one 'Blackie Toy's' laundry." *Id.* at 480. Moreover, there was no indication that James Wah Toy and "Blackie Toy" were the same individual. *Id.* at 481.

33. *See infra* notes 92–108 and accompanying text.

34. Illinois v. Gates, 462 U.S. 213, 235 (1983); Bacigal, *supra* note 3, at 294.

35. *E.g.*, United States v. Sokolow, 490 U.S. 1, 1 (1989).

36. Simmons, *supra* note 7, at 985.

37. *Sokolow*, 490 U.S. at 7 (quoting Terry v. Ohio, 392 U.S. 1, 27 (1968)).

38. Harris, *supra* note 28, at 982 (alteration in original) (quoting *Terry*, 392 U.S. at 21–22).

39. *See* Maryland v. Pringle, 540 U.S. 366, 371 (2003); Illinois v. Wardlow, 528 U.S. 119, 124–25 (2000); *Sokolow*, 490 U.S. at 17–18 (Marshall, J., dissenting) (demonstrating that the courts have accepted as sufficient to establish individualized suspicion a litany of contradictory characteristics, rendering a finding of individualized suspicion too easy to justify); Maclin, *supra*

have harshly critiqued the doctrine governing *Terry* stops due to shortcomings they see in this regard.[40] Nevertheless, the Court has steadfastly maintained that findings of reasonable suspicion and probable cause are valid only upon identification of individualized factors pointing to the subject of the search or arrest.

B. Individualized Suspicion's Incoherence

Despite the courts' continued insistence on individualized suspicion, there is a significant debate with respect to whether there is, in fact, such a thing as individualized decision-making[41] or whether, in the end, all decision-making is based on generalizations of one kind or another.[42] This debate applies not only to the idea of individualized suspicion, but also to the question whether one form of generalization in particular—statistical or probabilistic information—can satisfy a formal legal requirement of individualized suspicion.[43] While the idea of individualized suspicion is contained solely within Fourth Amendment doctrine, the debate over whether probabilistic evidence provides information sufficiently individualized to meet legal requirements arises in other contexts as well. Whatever the context, the theoretical concern is the same—the difficulty in identifying a coherent definition of "individualized."

Professor Frederick Schauer has argued persuasively that individualization is impossible to define and that all decisions are based on some form of generalization.[44] He uses as an example a local ordinance banning ownership of pit bulls due to the increased prevalence of aggression against humans that breed evinces.[45] Pit bull lovers attack such ordinances as unjustified generalizations.[46] *Their* pit bull, some owners argue, displays nothing but friendliness and docility, and painting all pit bulls with the broad

note 27, at 409–16 (pointing out that the Court never explains why the individualized suspicion element is satisfied in *Maryland v. Pringle* and arguing that its reasoning is inconsistent with existing Supreme Court precedent); Taslitz, *supra* note 19, at 152–53.

40. Taslitz, *supra* note 19, at 210 ("[S]ometimes the cost of the most powerful individualized-suspicion requirement is too high[, but] departure from it should be reluctant, gradual, and based upon increasingly pressing need."); *see also* Harris, *supra* note 28, at 976–77.

41. *See* Taslitz, *supra* note 19, at 190–202.

42. SCHAUER, *supra* note 2, at 75; *see also* CHRISTOPHER SLOBOGIN, PRIVACY AT RISK: THE NEW GOVERNMENT SURVEILLANCE AND THE FOURTH AMENDMENT 68 (2007) (asserting that "the distinction between individualized and generalized suspicion is . . . meaningless"); Bambauer, *supra* note 4, at 472 ("[G]eneralizations can be more finely grained by adding variables, but the nature of the prediction does not change." (footnote omitted)); *id.* at 468–82 (considering and rejecting four ways that courts and commentators have conceptualized the individualization requirement); Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 GA. L. REV. 109, 154–56 (2017).

43. For a definition of "statistical evidence," see *supra* note 2.

44. SCHAUER, *supra* note 2, at 75.

45. *Id.* at 56–72.

46. *Id.* at 56–57.

brush of aggressiveness penalizes them and their pet unfairly.[47] Rather, they suggest, local officials concerned about aggressive dogs should base the ban not on breed, but on dog-specific indicia of aggressive tendencies.[48] Under such a regime, for example, perhaps nobody could keep as a pet a dog that has bitten a human.[49] But, Schauer argues, "dogs that have bitten humans" is itself a generalization likely to be both over- and under-inclusive—dogs that have not yet bitten anyone may do so in the future, and a dog that has bitten someone may have done so for arguably justifiable reasons, such as protecting its owner from an assault.[50] On this view, regardless of how specific a generalization becomes—imagine, for example, a rule that bars ownership of male pit bulls that have been trained to fight and have exhibited unprovoked aggressive behavior toward children—it remains a generalization, even if ultimately you come up with a description that matches only one individual dog.[51] There is thus no reason, Schauer argues, to refrain from relying on generalizations (so long as they are accurate).[52]

Similarly, some scholars have advocated greater use of a particular type of generalization—one that is often described as "naked statistical evidence," or evidence composed solely of statistical probability, which is in essence a generalization reduced to a numerical probability. Arguments supporting the validity of probabilistic evidence historically have played out in the context of evidentiary questions and burdens of proof, rather than in the language of search-and-seizure law. But the theoretical underpinnings of the idea are the same. In both instances, the question is whether or when a general likelihood can meet the requisite standard—e.g., "probable cause," "more likely than not," etc.—in an individual case.

To demonstrate, consider a couple of famous hypotheticals that usually arise outside the Fourth Amendment context but are relevant to the discussion nonetheless. First, there is the case of the Gatecrashers' Paradox.[53] In this hypothetical, there is a rodeo at a stadium that holds 1,000 spectators. The stadium is full, yet the organizers sold only 499 tickets. There is thus a greater than 50 percent likelihood that any one attendee failed to purchase a ticket. The second famous statistics-based hypothetical, the Blue Bus case, involves a traffic accident in which a bus drives a motorist off the road,

47. *Id.*

48. *Id.*

49. *Id.*

50. *Id.* at 67–69.

51. *Id.*

52. *Id.* at 72.

53. The Gatecrashers' paradox originates with L. Jonathan Cohen's, *The Probable and the Provable*, published in 1977. COHEN, *supra* note 2, at 49, 74–81; *see also* L. Jonathan Cohen, *Subjective Probability and the Paradox of the Gatecrasher*, 1981 ARIZ. ST. L.J. 627, 627–34. My limited inquiry into the question has not yielded an answer to the mystery of why a British philosopher would generate a hypothetical based on a rodeo. Why not a soccer (football), rugby, or cricket match?

injuring the driver, and then flees the scene.[54] The only thing the injured driver saw was that the bus was blue. There were no other witnesses, but the Blue Bus Company owns 80 percent of the blue busses in the city. In neither of these cases is there what is often referred to as "direct" evidence, or evidence specific to an individual rodeo attendee or the specific individual bus involved in the accident. Yet in both of these cases, the existing statistical evidence, standing alone, seems sufficient to satisfy the "preponderance of the evidence" standard required to prevail in a civil action—sometimes described as a greater than 50 percent likelihood.[55] Nevertheless, similar evidence is routinely deemed insufficient even to send a case to the jury.[56] The fact that there is no reason to suspect any one attendee, or any one blue bus, any more than any other is fatal to the case.[57] One can imagine these hypotheticals adapted to the Fourth Amendment context: Based on existing law, the police would not have probable cause to arrest for trespassing any one of the 1,000 rodeo attendees, knowing that 499 of them are innocent, despite the likelihood of each individual's guilt.[58] Nor would there be probable cause to arrest owners of the Blue Bus Company if it was alleged to have been criminally negligent in hiring unqualified drivers when there is a 20 percent chance that the Rainbow Bus Company (which also owns some blue buses) was responsible for the accident.[59]

Why are generalized or probabilistic showings, whether it is the Supreme Court's insistence on individualized suspicion or rules of evidence rejecting solely probabilistic proof, insufficient? Why does the use of such tools make so many of us uncomfortable?[60] Why is it that the legal system rejects the generalization—the statistical likelihood—and insists on individualized or "direct" evidence, even when the distinction between the two is difficult to define?

---

54. The Blue Bus hypo originates with Laurence Tribe's article in the 1971 volume of the *Harvard Law Review* entitled *Trial by Mathematics* and is loosely based on a Massachusetts Supreme Court case, *Smith v. Rapid Transit, Inc.*, 58 N.E.2d 754 (1945). Tribe, *supra* note 2, at 1340–41.

55. Interestingly, the English articulation of this standard is formulated as "a balance of the probabilities." SCHAUER, *supra* note 2, at 81.

56. Bacigal, *supra* note 3, at 295–96 ("[T]he law seems uncomfortable with relying wholly on base rates and making the leap from aggregate likelihood to a conclusion of probable cause in a specific case. 'Background evidence is considered somehow inferior to evidence that is individuating and specific to the case at hand.'" (footnote omitted)).

57. *See* Nesson, *Reasonable Doubt*, *supra* note 2, at 1196–97.

58. *See id.*

59. *See id.*

60. To be sure, there are some who do not find these uses of probabilistic information troubling. *See* Bacigal, *supra* note 3, at 306 (arguing that "one ten percent probability is as good as another"). Others find that choosing people at random when there is a one in ten chance that they committed a crime is worse than searching ten people who have been singled out based on non-mathematical evidence, knowing that nine of them are innocent. *See* Joseph D. Grano, *Probable Cause and Common Sense: A Reply to the Critics of* Illinois v. Gates, 17 U. MICH. J.L. REFORM 465, 496 (1984).

A partial answer is that, while generalizing may be impossible to avoid, not all generalizations are created equal. "Dogs who have bitten people" is a generalization, but it is a generalization that takes into account an individual dog's behavior rather than solely his breed, unlike the generalization referring only to "pit bulls." These two generalizations are fundamentally different when it comes to the concept of individualized suspicion as conceived in the Fourth Amendment context. One is based on what one *is*, while the other is based on what one *has done*. Professor Schauer treats such distinctions as merely differences in degree—some generalizations, he concedes, are more detailed than others.[61] But the difference between "no pit bulls" and "no dogs who have bitten people" is more than a difference in degree. One permits government action in the absence of any specific information about the affected individual, while the other requires some specific tie to the dog being singled out.[62]

Yet while requiring evidence regarding a particular individual or situation may be a necessary element of Fourth Amendment individualized suspicion, the presence of such an element is not inevitably sufficient. To illustrate this idea, consider *Richards v. Wisconsin*, which presented the question of when the police may dispense with the knock-and-announce rule that generally applies to the execution of searches.[63] The case involved illicit drugs, and the government argued that because drugs are so easy to dispose of given a warning, all drug-related arrests should qualify "for a per-se knock-and-announce exception."[64] The Supreme Court unanimously rejected this position on the grounds that it would overgeneralize and "encompass many situations that pose no risk of evidence destruction."[65] Instead, the Court said, the police must know that the arrestee is aware of their presence before entering without knocking. In such a circumstance there is a reason, *specific to the situation at hand*, raising a suspicion that evidence may be destroyed.[66] The Supreme Court's opinion did not dispense entirely with generalizations (drug-related arrests), but instead substituted a more specific generalization, one that distinguished the drug-related arrest in this case from others (drug-related arrests where the suspect is aware of the police presence).[67] Thus the

61. SCHAUER, *supra* note 2, at 55–65.

62. *See* Taslitz, *supra* note 19, at 156–60.

63. Richards v. Wisconsin, 520 U.S. 385, 392–94 (1997).

64. Bambauer, *supra* note 4, at 471.

65. *Id.*; Harris, *supra* note 28, at 1014 (noting that the Supreme Court "expressed strong misgivings about 'creating exceptions . . . based on the "culture" surrounding a general category of criminal behavior.'" (quoting *Richards*, 520 U.S. at 392)).

66. Bambauer, *supra* note 4, at 471.

67. *Id.* Professors Bernard Harcourt and Tracy Meares also recommend a rethinking of individualized suspicion determinations, recommending the use of what they call a "suspicion-sufficient randomization paradigm." *See* Harcourt & Meares, *supra* note 19, at 818. This system would first identify a group within which everyone has a sufficient level of suspicion, and then determine who among them to subject to a search or a stop on a purely randomized basis. *Id.* at

individualized fact that a case involves drugs is insufficient, but the individualized fact that a case involves drugs and the suspects are aware of the police's presence would be enough.

So, the legal concept of "individualized suspicion" is not entirely meaningless, but neither is it self-defining. It requires factors or evidence specific to an individual actor, but not just any individual-specific factor—it must be a factor (or combination of factors) that sufficiently differentiates that actor from a larger population to earn the label "probable cause" or "direct evidence." The question thus becomes which factors, and how many of them, must a generalization include before the state is justified in imposing burdens on individuals. That is to say, the question is not whether there is a difference between generalizations and individualization, but rather exactly what individualized facts do we, as a society, recognize as sufficient. In effect, whether one refers to such combinations of factors as "generalizations" or "individualized evidence" is therefore more legal conclusion than factual description.[68]

III. INDIVIDUALIZED SUSPICION AND PREDICTIVE ACCURACY: WHY PREDICTIVE ACCURACY IS NECESSARY BUT NOT SUFFICIENT

This Part makes the argument that predictive accuracy alone fails to provide a satisfying definition for individualized suspicion. Section III.A reviews the fundamental values underlying the individual suspicion requirement. Then, Section III.B lays out the arguments in favor of using the rate of predictions' statistical accuracy to satisfy this requirement. Section III.C will then demonstrate why these intuitively appealing arguments are ultimately unconvincing. Finally, Section III.D will explain the crucial role that non-probabilistic evidence plays in vindicating the individualized-suspicion requirement's purposes.

A. *THE PURPOSE OF THE INDIVIDUALIZED SUSPICION REQUIREMENT*

The Supreme Court's persistent demand for individualized suspicion is not arbitrary. Rather, its purpose is to restrain government discretion, thereby promoting multiple important interests and values central to the American

---

814–16. The issue should be whether there is *enough* suspicion, not whether it is individualized. This proposal does not, however, dispense with the individualization requirement entirely. What factors must intersect for an individual to fall within Harcourt and Meares' "suspicion-sufficient group?" Likely they envision something more akin to "drug-related arrests + awareness of police," than to "drug-related arrests" alone. So, when it comes to that suspicion-sufficient group, it might be said that there is, in fact, individualized suspicion for each member of the group.

68. *See* Harcourt & Meares, *supra* note 19, at 842 (arguing that individualized suspicion is just a placeholder for the conclusion that the stop or search is "reasonable" and has no independent meaning); *see also* Bacigal, *supra* note 3, at 303 ("The difference between unacceptable speculation and reasonable inference is not a logical distinction, but a legal judgment.").

legal system.[69] In assessing the scope of the Fourth and Fifth Amendments, the Supreme Court has noted that "the Bill of Rights was added to the original Constitution in the conviction that too high a price may be paid" for law enforcement and that "other social objects of a free society should not be sacrificed" for that goal.[70] In other words, law enforcement must be prohibited from operating in ways that impair basic values—such as privacy, dignity, and autonomy—that have historically been described as "sacred,"[71] "inviolate,"[72] or "indefeasible"[73] rights, principles of humanity and civil liberty,[74] and whose protection is an independent good. These values reflect a rejection of strictly efficacy-based decision-making in favor of a constitutional regime that, at least in certain circumstances, is willing to sacrifice some efficiency, security, or accuracy.[75]

69. Camara v. Mun. Court of City & Cty. of San Francisco, 387 U.S. 523, 528 (1967); *see also, e.g.*, Delaware v. Prouse, 440 U.S. 648, 653–54 (1979) ("The essential purpose of the proscriptions in the Fourth Amendment is to impose a standard of 'reasonableness' upon the exercise of discretion by government officials . . . in order 'to safeguard the privacy and security of individuals against arbitrary invasions.'" (footnote omitted) (citations omitted)); Dunaway v. New York, 442 U.S. 200, 214–15 (1979) ("Nothing is more clear than that the Fourth Amendment was meant to prevent wholesale intrusions upon the personal security of our citizenry."); United States v. Chadwick, 433 U.S. 1, 11 (1977) ("[A] fundamental purpose of the Fourth Amendment is to safeguard individuals from unreasonable government invasions of legitimate privacy interests . . . ." (footnote omitted)), *abrogated by* California v. Acevedo, 500 U.S. 565 (1991); Brinegar v. United States, 338 U.S. 160, 180 (1949) (Jackson, J., dissenting) (insisting that Fourth Amendment rights are "in the catalog of indispensable freedoms" because "[u]ncontrolled search and seizure is one of the first and most effective weapons in the arsenal of every arbitrary government"); Wolf v. Colorado, 338 U.S. 25, 27–28 (1949) ("The security of one's privacy against arbitrary intrusion by the police—which is at the core of the Fourth Amendment—is basic to a free society."); Anthony G. Amsterdam, *Perspectives on the Fourth Amendment*, 58 MINN. L. REV. 349, 417 (1974) ("A paramount purpose of the fourth amendment is to prohibit arbitrary searches and seizures as well as unjustified searches and seizures."); M. Blane Michael, *Reading the Fourth Amendment: Guidance from the Mischief that Gave It Birth*, 85 N.Y.U. L. REV. 905, 921 (2010) ("[T]he mischief that gave birth to the Fourth Amendment was the oppressive general search . . . . The lesson from this mischief is that granting unlimited discretion to [government officials] inevitably leads to incursions on privacy and liberty . . . ."); Barry Friedman & Cynthia Benin Stein, *Redefining What's "Reasonable": The Protections for Policing*, 84 GEO. WASH. L. REV. 281, 316–17 (2016) ("It has long been a common consensus that the Fourth Amendment guards against the evil of arbitrary government rummaging in people's lives.").

70. Feldman v. United States, 322 U.S. 487, 489 (1944).

71. Weeks v. United States, 232 U.S. 383, 391 (1914).

72. *Feldman*, 322 U.S. at 490.

73. Boyd v. United States, 116 U.S. 616, 630 (1886) (insisting that the Fourth Amendment protects an "indefeasible right of personal security[,] personal liberty, and private property"); *see also* Osborn v. United States, 385 U.S. 323, 350 (1966) (Douglas, J., dissenting).

74. Mapp v. Ohio, 367 U.S. 643, 656–57 (1961) ("We find that . . . the freedom from unconscionable invasions of privacy and the freedom from convictions based upon coerced confessions do enjoy an 'intimate relation' in their perpetuation of 'principles of humanity and civil liberty.'" (quoting Bram v. United States, 168 U.S. 532, 543–44 (1897))).

75. Other areas where fundamental principles outweigh accuracy are the use of the exclusionary rule as a Fourth Amendment remedy, see United States v. Leon, 468 U.S. 897, 907 (1984) (cautioning against indiscriminate use of the exclusionary rule because it impedes "the

Privacy is one of those fundamental rights "basic to a free society."[76] The Fourth Amendment's privacy guarantees apply whenever the Federal government invades an area in which individuals have a reasonable expectation of privacy.[77] In extending the right to privacy to the states through the Due Process Clause of the Fifth Amendment, the Supreme Court has deemed privacy to be one of the rights "implicit in 'the concept of ordered liberty'" and forming part "of human rights enshrined in the history and the basic constitutional documents of English-speaking peoples."[78] Barring this zone of privacy from government intrusion preserves individual dignity[79] and the ability to engage in "intimate activity associated with . . . 'the privacies of life.'"[80] The necessary respect for the right to privacy—and the concomitant benefits of a zone of dignity and intimacy—requires rules designed to narrow the possibility that government officials will intrude on this zone without good reason. Individualized suspicion is designed to differentiate from the general public those people the government has a reason to stop or search, making it less likely that innocents will be affected. When the government may take action only against individuals for whom it has individualized suspicion to suspect criminal activity, it reduces the likelihood of false positives, or subjecting law-abiding citizens to the "hassle"—i.e., dignitary harms—of unjustifiable stops and searches.[81] Individualized-suspicion requirements thus safeguard fundamental individual rights by imposing constraints on government powers.

Insisting upon a particularized evidentiary basis to act also demonstrates respect for individual autonomy. Demanding that any intrusion into our

---

criminal justice system's trust-finding function"); and the recognition of evidentiary privileges, such as lawyer-client, doctor-patient, spousal, or priest-penitent privilege, see Jaffee v. Redmond, 518 U.S. 1, 9 (1996) (noting that testimonial privileges are justified by "public good transcending the normally predominant principle of utilizing all rational means for ascertaining truth" (quoting Trammel v. United States, 445 U.S. 40, 47 (1980))).

76. Wolf v. Colorado, 338 U.S. 25, 27 (1949).

77. *See, e.g.*, Rakas v. Illinois, 439 U.S. 128, 143 (1978) (holding that the determinative question is "whether the person who claims the protection of the Amendment has a legitimate expectation of privacy in the invaded place" (citations omitted)); United States v. Chadwick, 433 U.S. 1, 11 (1977) (asserting that the "purpose of the Fourth Amendment is to safeguard individuals from unreasonable government invasions of legitimate privacy interests"), *abrogated by* California v. Acevedo, 500 U.S. 565 (1991); Katz v. United States, 389 U.S. 347, 357–59 (1967).

78. *Wolf*, 338 U.S. at 27–28 (citation omitted); *see also* Feldman v. United States, 322 U.S. 487, 489–90 (1944) (noting that the Fourth and Fifth Amendments are "intertwined" in order "to maintain inviolate large areas of personal privacy").

79. Schmerber v. California, 384 U.S. 757, 767 (1966) ("The overriding function of the Fourth Amendment is to protect personal privacy and dignity against unwarranted intrusion by the State.").

80. Dow Chem. Co. v. United States, 476 U.S. 227, 236 (1986) (quoting Oliver v. United States, 466 U.S. 170, 180 (1984)) (holding that aerial photography of a corporation's industrial complex did not implicate the Fourth Amendment because "[t]he intimate activities associated with family privacy" do not take place there).

81. *See* Bambauer, *supra* note 4, at 464.

privacy is premised on things we actually have done—as opposed to who we are or with whom we associate—reflects an appreciation of our need to develop a unique sense of self, while reliance on generalizations engenders conformity, rather than individual flourishing.[82] Because "we experience invasions of privacy as assaults on our identity," an individualized suspicion requirement avoids invasions of our constitutionally protected zone of privacy by demanding "a significant level of justification" based upon an "individual's actions and behavior," rather than upon "mere conjecture," "generalizations or surmise."[83] Thus, even though drug dealers tend to have drugs in both their cars and their houses, for example, finding drugs in an individual's car does not provide probable cause to search his house for drugs—the generalization regarding drug dealers is insufficiently tailored to any one particular case.[84] Not only must there be a concrete justification for government action, but that justification "must be based on [an] individual's actions and behavior."[85] Thus before your privacy or liberty are infringed upon, you must have sufficiently differentiated yourself through your actions such that we, as a society, feel justified in taking more intrusive steps to determine whether you are engaged in nefarious activity.[86]

Individualized suspicion also promotes fairness, and perhaps as importantly the perception of fairness, in the legal process—i.e., procedural justice.[87] Procedural justice demands that government actors follow the rules, that their reasons for doing so are transparent, and that citizens have an opportunity to challenge and receive an explanation for the government's exercise of its power against them. The presence of sufficient indicia of

---

82. *See* Andrew E. Taslitz, *Police Are People Too: Cognitive Obstacles to, and Opportunities for, Police Getting the Individualized Suspicion Judgment Right*, 8 OHIO ST. J. CRIM. L. 7, 66–67 (2010) [hereinafter Taslitz, *Police Are People Too*]; Taslitz, *supra* note 19, at 187 ("Privacy is . . . one way by which we express our need for individualized justice: for being judged for who we really are.").

83. Taslitz, *supra* note 19, at 187–90.

84. *See* United States v. Hodson, 543 F.3d 286, 293 (6th Cir. 2008) ("[S]tanding alone, a high incidence of child molestation by persons convicted of child pornography crimes may not demonstrate that a child molester is likely to possess child pornography." (quoting United States v. Adkins, 169 F. App'x 961, 967 (6th Cir. 2006))); State v. Kahn, 555 N.W.2d 15, 19 (Minn. Ct. App. 1996).

85. Taslitz, *supra* note 19, at 189.

86. *See, e.g.*, Davis v. Mississippi, 394 U.S. 721, 722–28 (1969); Carroll v. United States, 267 U.S. 132, 153–54 (1925) (holding that, during prohibition, "[i]t would be intolerable and unreasonable if a prohibition agent were authorized to stop every automobile on the chance of finding liquor, and thus subject all persons lawfully using the highways to the inconvenience and indignity of such a search"); Thomas K. Clancy, *The Role of Individualized Suspicion in Assessing the Reasonableness of Searches and Seizures*, 25 U. MEM. L. REV. 483, 535 (1995).

87. Tom R. Tyler & E. Allan Lind, *Procedural Justice*, *in* HANDBOOK OF JUSTICE RESEARCH IN LAW 84 (Joseph Sanders & V. Lee Hamilton eds., 2001) ("People are . . . interested in the . . . message the process used to handle their problem communicates about their standing as full participants in the society. People want to believe that third parties care about their concerns, consider their arguments, and try to be fair to them—symbols of particularistic attention."); Taslitz, *supra* note 19, at 175–79.

procedural fairness is important because "[p]eople are . . . interested in the . . . message the process use[s] to handle their problem" and what that message "communicates about their standing as full participants in the society."[88] Tyler and Lind note that "[p]eople want to believe that third parties care about their concerns, consider their arguments, and try to be fair to them—symbols of particularistic attention."[89] Legal decisions are meant to reach conclusions people will accept. Imposing standards such as probable cause insists that the process is—and is seen as—non-arbitrary.[90] These fundamental values—privacy and dignity, autonomy, procedural justice —underlie individualized suspicion requirements, rendering respect for such requirements essential to our legal system and the rule of law.

B. Predictive Accuracy as a Proxy for Individualized Suspicion

Because the point at which evidence goes from being "generalized" to being "individualized" is essentially in the eye of the beholder, and because the existing individualized suspicion doctrine fails in the eyes of many commentators to sufficiently constrain or channel the exercise of government power, several commentators have sought systematic ways to define objectively what constitutes individualization. While different theories have been offered, each of them relies upon statistical probabilities as a proxy for whether a decision is individualized. One common approach is to advocate that courts assign a numerical value to standards like probable cause—perhaps 33 percent likelihood—and have courts determine that evidence indicating a likelihood at or above that point establishes probable cause.[91]

Others seek to replace individualized suspicion with something else entirely. When it comes to reasonable suspicion, for example, Professors Tracy Meares and Bernard Harcourt propose what they call a "suspicion-sufficient randomization paradigm."[92] This system would first identify a group for whom there is sufficient suspicion regarding each individual member, and then pick among them on a purely randomized basis.[93] As with the idea of a

88. Tyler & Lind, *supra* note 87, at 84; *see also* Maclin, *supra* note 27, at 574 n.114 (citing evidence that people are more reluctant to help police when they think police are abusing their discretion); Taslitz, *Police Are People Too*, *supra* note 82, at 11 ("Repeated police errors can also undermine public trust, leading citizens to avoid reporting crimes, sharing information with law enforcement, or cooperating with police requests.").

89. Tyler & Lind, *supra* note 87, at 84.

90. *See* Nesson, *Reasonable Doubt*, *supra* note 2, at 1195–97 (discussing the impracticability of defining precise statistical measurements that constitute a sufficient likelihood of guilt to convict); Andrea Roth, *Safety in Numbers? Deciding When DNA Alone Is Enough to Convict*, 85 N.Y.U. L. REV. 1130, 1165 (2010) (arguing that our insistence that jurors are personally convinced before condemning a defendant rather than allowing them to bet on guilt helps preserve perception that the system values individual dignity).

91. *See supra* text accompanying note 19.

92. Harcourt & Meares, *supra* note 19, at 818.

93. *Id.* at 814–18.

fixed numerical value, the issue would be whether there is enough suspicion, not whether it is individualized.[94] Harcourt and Meares recognize the arbitrariness inherent in their randomization proposal, but contend that it is an improvement over the current means of making such decisions, which is biased and manipulable.[95]

The most common solution offered to the indeterminacy of individualized suspicion is the use of cost-benefit analysis. Professor Schauer argues, for example, that because all evidence sits somewhere on the spectrum of generalization, which ranges from extremely broad to sufficiently specific, he argues that the costs and benefits of the proposed action determine where on the spectrum qualifies as "individualized" in any given instance. Returning to his pit bull ordinance, for example, he argues that it is too costly to wait for a pit bull to bite someone before requiring the animal be euthanized. As a result, he is willing to accept a concededly over-inclusive rule.[96]

Professor Taslitz also views individualized suspicion as a spectrum but argues that at some point, a sufficiently specific generalization becomes more than a difference in degree; it becomes a difference in kind and therefore qualifies as individualized.[97] As with Schauer, Taslitz would locate that point by weighing the costs of the government action against its resulting benefits. Taslitz's sensitivity to the costs of individualized suspicion, given the problematic nature of errors in the criminal justice context, results in an effort to mitigate the damage of reliance on probabilities.[98] First, he argues that we should err on the side of under-inclusive rules, placing the costs of mistakes on society rather than on individual suspects. And second, he suggests greater willingness to expend resources necessary to avoid errors in the first place or to correct errors that are made.[99] Nevertheless, at bottom, his vision employs a cost-benefit analysis premised on the likelihood that a stop or search will yield investigative benefits.

Professor Jane Bambauer has suggested treating the individualized suspicion requirement as a form of cost-benefit analysis that embraces the use

---

94. Of course, another way to view Harcourt and Meares' proposal is to say not that it sets a threshold for the individualization determination but rather that, once such a determination has been made with respect to a number of individuals, the decision of who to actually inconvenience should be out of the hands of government officials.

95. Harcourt & Meares, *supra* note 19, at 815–16.

96. SCHAUER, *supra* note 2, at 69–72.

97. *See* Taslitz, *supra* note 19, at 155–61; Taslitz, *Police Are People Too*, *supra* note 82, at 51 ("Some use of generalizations is unavoidable in human reasoning, so no reasoning is purely individualized or purely generalized. Rather, there is a spectrum, and the reasonable suspicion and probable cause concepts should be understood as favoring the more individualized end of that spectrum." (footnote omitted)).

98. *See* Taslitz, *supra* note 19, at 210.

99. *Id.* at 162–64 (noting that error-correction efforts can be expensive and may itself produce errors).

of statistical evidence, one that is focused on minimizing a particular type of cost: false positives.[100] Thus she would consider a decision to conduct a search "individualized" so long as the hit rate for such searches were sufficiently high, and the false positives—what she terms the "hassle" rate—were sufficiently low. Individualization thus does not result in suspicion regarding one individual. Rather, it identifies a category of searches whose investigative benefits outweigh the costs they impose on innocents. Bambauer posits a hypothetical (reliable) study concluding that 60 percent of on-campus dorm rooms at Harvard contain illicit drugs.[101] Since there is general agreement that the threshold for probable cause sits below "more likely than not," a 60 percent likelihood of finding evidence of criminality would seem sufficient to justify a search.[102] She argues that, in such instances, the investigative benefit—the hit rate of 60 percent—should be balanced against what she labels the "hassle rate"—the cost that false positives impose on innocents.[103] Whenever the benefit exceeds the cost, she argues, that search should be considered "individualized."

With each of these cost-benefit approaches, the question of when a determination is sufficiently individualized is determined by asking whether the suspicion level for a certain subset of people is sufficiently high that the police are justified in searching any of them. The question is thus not whether evidence or a decision is individualized in isolation, but rather whether the costs of its over-inclusiveness are outweighed by its benefits.[104] This type of analysis generates a sliding scale—the more severe the costs of false positives—viewed as either the rate of false positives or the intrusiveness of the government action—the higher the investigative payoff must be.[105]

Whatever their appeal, none of these proposals would satisfy existing individualized suspicion requirements. Indeed, the Supreme Court consistently has rejected assigning either probable cause or reasonable suspicion a numerical value, insisting that they are not standards that can be determined with mathematical precision.[106] Any magistrate judge asked for a warrant to search any particular Harvard dorm room on the basis of

---

100. *See generally* Bambauer, *supra* note 4 (considering and rejecting four ways that courts and commentators have conceptualized the individualization requirement).

101. *Id.* at 462–63 (adapting a hypothetical posed in Orin Kerr, *Why Courts Should Not Quantify Probable Cause*, *in* THE POLITICAL HEART OF CRIMINAL PROCEDURE: ESSAYS ON THEMES OF WILLIAM J. STUNTZ 131, 135–37 (Michael Klarman et al. eds., 2012)).

102. *Id.*

103. *Id.* at 464–66 (arguing that hassle rates serve one role normally envisioned for individualization—limiting the number of innocents subjected to police intrusion).

104. SCHAUER, *supra* note 2, at ix–xi.

105. Bacigal, *supra* note 3, at 332–35 (arguing that probable cause should be a sliding scale depending on how serious the crime or risk of future harm is, and identifying five categories of required suspicion); Taslitz, *supra* note 19, at 155–68.

106. *E.g.*, Maryland v. Pringle, 540 U.S. 366, 371 (2003) (noting that probable cause "is incapable of precise definition or quantification into percentages").

Bambauer's hypothetical study would reject the application on the grounds that the evidence is not sufficiently individualized—that there is no fact specific to any particular room (other than the fact that it is on Harvard's campus) that makes it more likely to contain contraband than any other.[107] In other words, because the evidence is purely probabilistic, the magistrate judge would reject the warrant. While Bambauer concedes that the traditional application of individualized suspicion would prevent the imposition of search costs on the 40 percent of Harvard students on campus who do not harbor illegal drugs in their room, she argues that cost is exceeded by the benefits of being able to "create inroads for criminal enforcement within elite communities otherwise immune to the enforcement of minor criminal laws."[108] Similarly, Harcourt and Meares concede that their suspicion-sufficient randomization paradigm is, under existing law, impermissibly arbitrary, yet believe the benefits of its discretion-limiting effects outweigh the costs of arbitrariness.

The flaws these scholars have identified are very real and very troubling. Unfortunately, as the next Section will demonstrate, their chosen means of addressing the problem not only will fail to meaningfully constrain law enforcement discretion but also will undermine other, less concrete purposes underlying the individualized-suspicion requirement.

C. The Illusory Constraining Power of Numerical Thresholds

To the extent that confidence in statistical probabilities and numerical thresholds rests on their ability to further the aim of constraining government discretion, that confidence is largely misplaced. Justifications that courts have accepted as sufficiently particularized often look more like categorical judgments than individualized suspicion.[109] As Justice Marshall famously pointed out in the context of drug courier profiling, courts have held numerous and inconsistent factors relevant to the individualized suspicion determination: suspect was first to deplane, suspect was last to deplane, suspect deplaned from middle; suspect traveled on one-way tickets, suspect traveled on round-trip tickets; suspect took a nonstop flight, suspect changed planes; suspect traveled with no luggage, suspect carried a gym bag, suspect carried a new suitcase; suspect traveled alone, suspect traveled with a companion.[110] When just about any behavior can be characterized as suspicious, law enforcement discretion is essentially unfettered.

This phenomenon is particularly pronounced when it comes to the reasonable-suspicion standard needed to justify an investigative stop. Courts

107. Bambauer, *supra* note 4, at 462–63.

108. *Id.* at 465.

109. Harris, *supra* note 28, at 987–1012 (arguing that the bases for reasonable suspicion are so broad as to be categorical rather than individualized).

110. United States v. Sokolow, 490 U.S. 1, 13–14 (1989) (Marshall, J., dissenting).

have often deemed this standard met on the basis of broad categories of behavior, such as engaging in "furtive movements" or being located in a "high crime area."[111] The combination of this low doctrinal hurdle and the government's sophisticated investigative tools mean there will be a sufficient amount of suspicion about a sufficient number of people to almost entirely defang the reasonable-suspicion requirement. With respect to investigative stops, therefore, the rationale for many commentators' support for use of probabilistic evidence begins to crumble.

Modern technology has only exacerbated this concern. The time is approaching, if it has not arrived already, when government surveillance tools will allow law enforcement to record, store, and analyze nearly everything we do.[112] As the Supreme Court has recognized in cases like *United States v. Jones*,[113] *Riley v. California*,[114] and *Carpenter v. United States*,[115] the government can both cheaply and easily capture not only previously unimaginable volumes of information, but it can also aggregate that information in ways that disclose new forms of information.[116] And while these three cases demonstrate that the Court seems inclined to impose some limits on some types of data collection, there is a variety of technologies being employed (and more being developed all the time) that currently remain outside the scope of Fourth Amendment protection. From Stingrays[117] to drones to facial and license-plate recognition software to digital searches at the border to the output of commercial data aggregators,[118] vast amounts of data about every American

---

111. *See* Harris, *supra* note 28, at 976.

112. *See, e.g.*, Aliza Hochman Bloom, *When Too Many People Can Be Stopped: The Erosion of Reasonable Suspicion Required for a* Terry *Stop*, 9 ALA. CIV. RTS. & CIV. LIBERTIES L. REV. 257, 264–69 (2018); Joh, *supra* note 7, at 15–19; Maclin, *supra* note 27, at 395–98.

113. United States v. Jones, 565 U.S. 400, 412–13 (2012).

114. Riley v. California, 573 U.S. 373, 385–87, 393–95 (2014).

115. Carpenter v. United States, 138 S. Ct. 2206, 2211–13, 2216–18 (2018).

116. *See Jones*, 565 U.S. at 415–16 (Sotomayor, J., concurring) (recognizing that our digital data trails "generate[] a precise, comprehensive record of a person's public movements that reflects a wealth of detail about her familial, political, professional, religious, and sexual associations[,] . . . [that] [t]he Government can store such records and efficiently mine them for information years into the future," and that because digital surveillance is cheap and goes largely undetected, "it evades the ordinary checks that constrain abusive law enforcement practices: 'limited police resources and community hostility'" (quoting Illinois v. Lidster, 540 U.S. 419, 426 (2004))); *see also* Emily Berman, *When Database Queries Are Fourth Amendment Searches*, 102 MINN. L. REV. 577, 579–80 (2017) (discussing how the information the government can glean by aggregating data can reveal private information).

117. "Stingray" is a brand name for a device that collects the identity of cellular phones in a particular geographic location by impersonating a cell phone tower. *See, e.g.*, Jessica Glenza & Nicky Woolf, *Stingray Spying: FBI's Secret Deal with Police Hides Phone Dragnet from Courts*, GUARDIAN (Apr. 10, 2015, 10:49 AM), https://www.theguardian.com/us-news/2015/apr/10/stingray-spying-fbi-phone-dragnet-police [https://perma.cc/CGP6-QCQ4].

118. *See, e.g.*, Aaron Gregg, *For this Company, Online Surveillance Leads to Profit in Washington's Suburbs*, WASH. POST (Sept. 10, 2017), https://www.washingtonpost.com/business/economy/

with an internet connection are available to the government. This means that law enforcement likely can find individualized suspicion for *something* about each and every one of us. Who among us has not done something that might qualify as individualized suspicion for a stop or a search—rolled through a stop sign, jumped a subway turnstile, brought Cuban cigars back from a trip abroad, purchased a "tobacco" pipe at a head shop? So, no matter the numerical threshold one sets, it remains up to the government's discretion to determine where to impose the costs of law enforcement. Having articulated this skepticism regarding the idea that probabilistic evidence can meaningfully constrain law enforcement's investigative actions, this Part now turns to a discussion of how reliance on probabilistic accuracy also will undermine the other, less concrete goals of the individualized-suspicion requirement.

D. *The Crucial Role of Non-Probabilistic Evidence*

The Fourth and Fifth Amendment procedures for assessing the sufficiency of individualized suspicion are relatively straightforward. Government officials must articulate a justification explaining their decision. They must also provide sufficient evidence to meet the relevant standard, whether that is probable cause or reasonable suspicion. And to make these two requirements meaningful, error-correction methods must be available to provide an affected individual his day in court to challenge the government's explanation as erroneous or insufficient.

A closer look at the nature of these mechanisms, however, provides clues as to why using probabilistic evidence alone—so appealing to scholars seeking more objective legal standards—has failed to win over courts. The upshot is that it is not enough merely for these procedures to exist. Rather, the manner in which they are conducted—and the nature of the evidence they assess—is critical. Our procedures to ensure that an individual knows the nature of the allegations against her and receives a justification for the government's decision to impose burdens on her reflect the way in which humans make decisions, what humans find persuasive, and what humans deem necessary to legitimate government incursions into private spaces. Even the Supreme Court has recognized that a decision that relies solely on a high likelihood of success is simply inconsistent with these ideas. Despite the fact that, "[i]n dealing with probable cause . . . as the very name implies, we deal with probabilities,"[119] the Supreme Court has made plain that establishing probable cause (or reasonable suspicion) is not a mathematical exercise. Rather, "probable cause is a fluid concept—turning on the assessment of probabilities in particular factual contexts—not readily, or even usefully,

for-this-company-online-surveillance-leads-to-profit-in-washingtons-suburbs/2017/09/08/6067c924-9409-11e7-89fa-bb822a46da5b_story.html [https://perma.cc/9CUX-LCYG].

119. Brinegar v. United States, 338 U.S. 160, 175 (1949).

reduced to a neat set of legal rules."[120] Not only has the Court rejected calls to reduce probable cause or reasonable suspicion to a numerical formula, it also has emphasized that these standards require a non-technical inquiry into the totality of the circumstances. The question is not simply "how likely is it that criminal activity is afoot?," but rather whether the facts and circumstances within government officials' knowledge and of which they had reasonably trustworthy information were sufficient to warrant a person of reasonable caution in the belief that criminal activity is afoot.[121] As a result, valid determinations of individualized suspicion require something beyond probabilistic evidence.

A closer examination of three characteristics of the relevant procedures will illustrate this point. The first such feature is the means of evaluating how convincing any given piece of evidence may be. Second is the way the standard of review for individualized suspicion is articulated. And third is the nature of the "totality of the circumstances" analysis. It is thus not just the existence of a process for establishing individualized suspicion that matters. How it is carried out is a critical means of vindicating the values that individualized suspicion promotes.

Take first the nature of the evidence that suffices to establish reasonable suspicion. Proponents of statistical decision-making argue that the inevitable inferential leap from the accumulated evidence to a legal conclusion means that all evidence is probabilistic. All inferences, however, are not created equal.[122] And the types of inferences deemed acceptable by the courts have a different nature from those inferences articulated through statistical or probabilistic evidence. One way to think of the difference between probabilistic evidence, which the legal system rejects, and "individualized" evidence is to consider the difference between a *justification* for the government action and an *explanation* for government action.

Human decision-makers—whether they are judges at bail hearings, juries in criminal cases, or cops on the beat—infer from known facts to the explanation that best explains those facts, a process sometimes referred to as "inference to the best explanation."[123] This method requires a decision-maker

---

120. Illinois v. Gates, 462 U.S. 213, 232 (1983).

121. *See id.* at 231 ("Perhaps the central teaching of our decisions bearing on the probable cause standard is that it is a 'practical, nontechnical conception.'" (quoting *Brinegar*, 338 U.S. at 176)); Carroll v. United States, 267 U.S. 132, 162 (1925).

122. *See* Michael S. Pardo & Ronald J. Allen, *Juridical Proof and the Best Explanation*, 27 L. & PHIL. 223, 227 (2008) (articulating the distinctions between inductive and deductive inferences). The legal system almost always implicates induction—such as whether the available evidence adds up to probable cause or whether the evidence presented to a jury proves guilt beyond a reasonable doubt. *Id.* at 229–33.

123. *See id.* at 227–33; Brennan-Marquez, *supra* note 7, at 1277 ("Although it is true that all evidence relies, at some level, on generalization, a meaningful line can be drawn between inferences that merely draw predictions from observed facts and inferences that purport to explain those facts. Explanatory power, in other words, is not an epistemic illusion.").

to generate in her mind all of the possible explanations of the known evidence or circumstances and choose the best explanation from among those options.[124] This is the way the legal system presents questions—it asks decision-makers to assess available evidence and determine what story is most likely to explain that evidence. In the world of inductive inferences there will always be alternative possibilities, but some will be more plausible than others. Generally speaking, the more compelling the explanation, the more convincing the case or the more reasonable the suspicion.[125] Another name some scholars have assigned to the best or most plausible justification is "normic justification." A belief is normically justified if, in a normal world, the known facts would lead one to hold that belief.[126] Because the evidence would "normally" mean the belief is true, if the belief turns out to be *false*, then there must be some "mitigating circumstances [that explain the error] in terms of disobliging environmental conditions, or cognitive or perceptual malfunction or some such."[127] Statistical information, regardless of how accurate, cannot provide a normic justification.

Consider some examples. First, imagine that a defendant's DNA is discovered at a crime scene. This fact might lead us to infer that the defendant was at that location at some point. That DNA, however, might have been brought to the crime scene through some other mechanism—someone who was at the crime scene wore a coat with the defendant's hair stuck to it, or borrowed a coffee mug from the defendant's house and left it at the scene. The best explanation—the defendant had at one time been at the crime scene—is not guaranteed to be true, but if it is not, something unexpected intervened. In other words, the best inference seeks to provide an answer to the question, "Why Story A *rather than* Story B?"[128]

Or recall Professor Bambauer's hypothetical about marijuana in the Harvard dorms. Rather than explain, "why search a Harvard dorm room?," an inference to the best explanation would explain, "why search *this* Harvard dorm room rather than another?"[129] If smoke and the odor of marijuana are emanating from behind a dorm room door, for example, law enforcement

---

124. *See* Pardo & Allen, *supra* note 122, at 229–33 (describing the process of inference to the best explanation).

125. *See id.* at 229 ("[E]xplanatory considerations help to determine how likely one judges particular hypotheses or conclusions to be.").

126. Martin Smith, *What Else Justification Could Be*, 44 NOÛS 10, 13–19 (2010). *See generally* David Enoch et al., *Statistical Evidence, Sensitivity, and the Legal Value of Knowledge*, 40 PHIL. & PUB. AFF. 197 (2012) (identifying a dichotomy between evidence that is "counterfactually sensitive" —evidence that will be different if the truth is different—and non-sensitive evidence as relevant to the distinction between individualized and statistical evidence).

127. Smith, *supra* note 126, at 17 (emphasis omitted).

128. Pardo & Allen, *supra* note 122, at 232.

129. Ian Kerr & Jessica Earle, *Prediction, Preemption, Presumption: How Big Data Threatens Big Picture Privacy*, 66 STAN. L. REV. ONLINE 65, 68–69 (2013); *see* Bambauer, *supra* note 4, at 462; *see also supra* notes 101–03, 107–08 and accompanying text.

would be normically justified in determining that the resident possessed illicit drugs—in a normal world, the best explanation for smoke and the odor of marijuana is that there is, in fact, marijuana in the room. By contrast, if the resident did *not* possess such drugs, there would have to be some less likely reason why the odor and smoke were coming from behind the door nevertheless. If, by contrast, the only evidence of the presence of marijuana is a 60 percent statistical likelihood applicable to all Harvard dorm rooms, however, and the police select a room to search but do not find marijuana, there is no need to provide an explanation for that failure. That room simply happened to be one of the four in ten where no drugs are present.

Similarly, if the reason we believe that the Blue Bus Company caused the accident is the fact that it owns 80 percent of blue buses, but it turns out instead that the Red Bus Company is to blame, that is simply the result of, as it were, the luck of the draw. But if an eyewitness in circumstances where witnesses are correct 80 percent of the time identified the bus as a Blue-Bus-Company Bus, we would want to know how the witness got it wrong—it was dark, she was not wearing her glasses, etc. The result may be the same—erroneous results in two of ten cases. But at the time the decision is made, an eyewitness account justifies a particular conclusion (even if less likely to be accurate) in a way that pure statistical likelihood does not.[130] One type of evidence does not differentiate between cases, while the other requires law enforcement to have reason to believe that *this particular person*, or *this particular bus* is a justifiable target of state intervention. When it comes to insisting on individualized decision-making, it is therefore not sufficient that a legal conclusion is, to some extent, probable. Rather, the reasons that conclusion is probable—the ability of the story to explain that conclusion—matters.[131]

130. It is possible that some statistical evidence is so likely to be true that the likelihood it creates crosses the line into normic justification. Some forms of DNA evidence might be an example. *See* David Enoch & Talia Fisher, *Sense and "Sensitivity": Epistemic and Instrumental Approaches to Statistical Evidence*, 67 STAN. L. REV. 557, 587–92 (2015). If a defendant's DNA is found at a crime scene, law enforcement is normically justified in believing the defendant himself was there. In a normal world, someone's DNA is present in a specific location because that individual left it there. If, however, the defendant was *never* at the crime scene, there would have to be some additional explanation how his DNA got there. Perhaps, one day, computer models will reach the reliability of (properly collected and tested) DNA evidence, but we are not there yet.

131. *See* Brennan-Marquez, *supra* note 7, at 1255, 1265 ("It is possible for an inference to be likely, in a probabilistic sense, without being relatively plausible. The latter depends not only on the predictive power of an inference, but also on its 'quality.' Specifically, it depends on whether the factual inputs giving rise to the inference enable an observer to meaningfully compare different explanations before deciding which to entertain." (footnote omitted)); Taslitz, *Police Are People Too*, *supra* note 82, at 54 ("Experience and intuition aid the imagination, but the act of imaginative testing is also a decidedly conscious one, so its outcomes can be explained to others. If the outcomes are not adequate, the expert relies on intuition to craft an alternative hypothesis, imaginatively testing it as well. When a hypothesis is found that adequately survives testing, existing mental models generate action scripts—guides to choice of decision and resulting behavior." (footnotes omitted)).

In addition to explaining why probabilistic evidence is rejected in favor of evidence that might be less accurate, this view of the government's obligation to justify its decisions is reflected in Fourth Amendment procedural requirements. Applicable legal doctrines assume fact-finders will engage in inference to the best explanation to generate or select a convincing narrative. For example, when making a probable cause determination, a police officer or a judicial magistrate must generate a story of how and why they reached the conclusion that probable cause existed, why the choice to search this individual rather than another is justified.[132] If probabilities alone were dispositive, we would use statistical, rather than narrative justifications.[133]

Those who favor the increased reliance on probabilistic evidence, unsurprisingly, reject the idea that a narrative-focused approach is compatible with existing legal rules while probabilistic evidence is not. What is important, they argue, is not the number of factors taken into account or the plausibility of an explanation, but the accuracy of the predictions. They often advocate for a Bayesian approach, rather than a narrative approach.[134] A Bayesian

---

132. *See* Taslitz, *Police Are People Too*, *supra* note 82, at 32 ("Indeed, because the individualization mandate impliedly contains a mandate for officer explanation of his actions, the officer must anticipate objections to gaps in the evidence, thus seeking to fill them, perhaps even to explore alternative potential perpetrators." (emphasis omitted)). The extent to which this requirement constrains government action is, of course, a separate question. A long-standing critique of Fourth Amendment law is that police officers can act first and think of (or make up) sufficient justification for their actions after the fact. "'Articulable bases' [of action] become rote, readily repeatable phrases, rather than clear articulations of reasons for action in a particular case." *Id.* at 75. Moreover, some of the justifications that courts accept as sufficiently individualized are applicable to broad swaths of the population. *See* Maclin, *supra* note 27, at 415 (arguing that the individualized suspicion requirement of probable cause has been diluted). These critiques, while certainly justified, are properly targeted at law enforcement officials for being insufficiently forthcoming and at the courts for being too permissive, not at the value of requiring government agents to justify imposing burdens on individuals.

133. *See* Pardo & Allen, *supra* note 122, at 226 (discussing the phenomenon in the context of admissibility of evidence in civil cases).

134. Bayes Theorem is a means of determining the probability of a particular fact, given some piece of evidence—the likelihood that A is true, given that we know B to be true. For example, what is the likelihood that a patient has a disease, given that she tested positive for that disease? We might know that two percent of the population has the disease. In the absence of the positive test, there is therefore a two percent likelihood that any given individual has the disease. If, however, the test is 95 percent accurate (it correctly identifies which individuals *have* the disease 95 percent of the time, and it correctly identifies which individuals *do not have* the disease 95 percent of the time) and our patient tested positive, we can update our prior probability of two percent to take account of this new fact. Bayes Theorem indicates that the likelihood that someone has a disease present in two percent of the population, given that she tested positive for the disease using a 95 percent accurate test, is almost 28 percent. This is the "conditional probability" that the patient has the disease, given that she tested positive for it. *See, e.g.*, DAVID H. KAYE ET AL., THE NEW WIGMORE: A TREATISE ON EVIDENCE: EXPERT EVIDENCE, POSTERIOR PROBABILITIES AND BAYES' RULE § 14.2.1 (2019) (providing a very basic explanation of Bayes Theorem); *see also, e.g.*, JAMES V. STONE, BAYES' RULE, A TUTORIAL INTRODUCTION TO BAYESIAN ANALYSIS 3–11 (2013) (providing a more in-depth, though still relatively accessible, explanation of Bayes' Theorem).

approach seeks the absolute likelihood of any given outcome by aggregating the statistical probabilities of all the possible stories supporting that outcome.[135] As it turns out, however, this is not the way humans—and therefore police officers, judges, administrative officials, and juries—process information. As one commentator has pointed out, "[e]ven if all evidence is statistical," that is to say, "even if 'case-specific' evidence is an illusion—there can still be case-specific explanations of evidence."[136] And it is through these case-specific narratives, rather than statistical probabilities, that humans assess likelihoods.

Consider an experiment conducted by the famous behavioral economists Amos Tversky and Daniel Kahneman.[137] In the experiment, test subjects were given information that there was a hit-and-run accident involving a taxi cab in a city where 85 percent of cabs are green and 15 percent are blue. They were also told that a witness identified the cab as blue and that a test revealed that the witness could accurately identify the color of a cab under conditions similar to those of the accident 80 percent of the time. The test subjects were then asked what the probability was that the cab involved in the accident was blue rather than green. Bayes' Theorem indicates that the likelihood that the cab was blue rather than green is 41 percent.[138] Test subjects, however, placed the probability at 80 percent, ignoring the fact that many more cabs in the city are green and the impact of the differential base rates on the statistical probability of a blue cab being the culprit. When instead the test subjects were told that, although the number of blue and green cabs is roughly equal, 85 percent of cab accidents in the city involve green cabs and 15 percent involve

---

135. Pardo & Allen, *supra* note 122, at 249–50 (Explanatory and Bayesian approaches are not necessarily incompatible, "to the extent Bayesian perspectives can clarify and approve on those considerations").

136. Brennan-Marquez, *supra* note 7, at 1279–80 (emphasis omitted).

137. Pardo & Allen, *supra* note 122, at 258–61 (citing Amos Tversky & Daniel Kahneman, *Evidential Impact of Base Rates*, *in* JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES 153–60 (Daniel Kahneman et al. eds., 1982)).

138. Taking account of the 15 percent base rate of blue cabs and the 80 percent accuracy of the witness, Bayes' Theorem indicates a 41 percent conditional probability that the cab involved in the accident was blue. Note that determining the base rate of any given phenomenon requires identifying the reference class. Rather than all cabs, for example, we could ask for the base rate of blue cabs among cabs in a certain city, or cabs that were on the road at the time the accident occurred, or cabs that picked up or dropped of a passenger in the vicinity of the accident.

> We can continue to specify more and more detailed classes, but eventually we arrive at the event itself . . . . The incidental base rates, thus, are subject to a particular reference class and without some guarantee that there is some degree of homogeneity within the class (e.g., are [g]reen cabs 85% more prevalent everywhere in town?), the data may not be very useful in telling us about the particular event.

Pardo & Allen, *supra* note 122, at 260. By contrast, explanations establish necessary connections that explain both the class and the event (green-cab drivers are bad drivers). There are reference class problems here too—maybe the green drivers on the road that night had pristine records —but the stronger explanation leads people closer to the statistical answer: "explanatory considerations guide inference and likelihood assessments." *Id.*

blue cabs they put the likelihood that the cab in question was green at 60 percent. From a purely Bayesian perspective, this scenario results in the exact same answer: 41 percent. So why do people come closer to the accurate statistical answer in the second scenario? Perhaps because there is a story to tell that calls into question the reliability of the witness's testimony. The cab in the accident is more likely to be green because green cabs are more likely to be driven by accident-prone drivers. When the base rate is presented as pure market share rather than accident-proneness, there is no reason to discount the witness's testimony.[139] The upshot here is that human predictions tell a story, and whether that story is a sufficiently reasonable one to tell given the available evidence is matter of debate. This both legitimates the government intrusion and provides a basis by which the lawfulness of that intrusion can be challenged.

The way the probable cause and reasonable suspicion standards are articulated also assumes human decision-making. As an initial matter, while the standard is objective, the question in probable cause is whether "the facts available to the officer at the moment of the seizure or the search 'warrant a man of reasonable caution in the belief' that the action taken was appropriate."[140] This formulation indicates that what renders an intrusion into Fourth Amendment space reasonable is whether one could believe it to be sufficiently justified in that particular instance. The reasonable suspicion standard is similarly formulated. The Supreme Court has "invariably" insisted that before stopping and searching an individual for weapons, an officer must have "a reasonable belief" that the subject of the stop is armed and dangerous.[141] Again, what renders the law enforcement action permissible is that an officer's *belief* of dangerousness would be reasonable. So Fourth Amendment doctrine does not ask whether it is sufficiently *likely* that a stop, search, or arrest would lead to evidence of criminality. Rather, the question is whether law enforcement officials had sufficiently convincing *reason to believe* that a stop, search, or arrest *in this particular instance* would do so.[142]

139. Pardo & Allen, *supra* note 122, at 259 ("The difference, in other words, is explanatory. A causal explanation explains why more [g]reen cabs are involved in accidents and also why the one involved in this crash is more likely to be [g]reen—namely, that [g]reen cabs are driven by bad drivers. . . . By contrast, no such explanatory connection exists in the first scenario.").

140. Terry v. Ohio, 392 U.S. 1, 21–22 (1968) (quoting Carroll v. United States, 267 U.S. 132, 162 (1925)).

141. *See* Ybarra v. Illinois, 444 U.S. 85, 942–94 (1979) (holding that in order to perform a frisk, police must have a reasonable belief that the suspect is armed and dangerous); Sherry F. Colb, *Probabilities in Probable Cause and Beyond: Statistical Versus Concrete Harms*, 73 L. & CONTEMP. PROBS. 69, 100, 105 (2010) (discussing human tendencies to put more weight on actual, rather than statistical harm, in the context of legal decision-making).

142. To be sure, the Supreme Court has determined that government officials' motivations for conducting stops, searches, or seizures are irrelevant, so long as there is probable cause. Whren v. United States, 517 U.S. 806, 813–15 (1996). So the individual officer need not himself hold the belief that the individual is guilty, so long as a reasonable person could so believe.

Finally, there is the fact that individualized suspicion assessments are made under a totality-of-the-circumstances standard. The use of a totality-of-the-circumstances standard is inherently non-probabilistic for at least two reasons. First, it is not a rule but a standard, which means that the individualized suspicion inquiry is not a question of fact, but a mixed question of fact and law. The process of deciding mixed questions is at the heart of our common-law system, the exercise of judgment through which law is made and developed. A critical role of the legal system is to provide a forum in which to argue about whether a particular set of facts meets a particular legal standard, such as probable cause. Ambiguities are inherent in standards.[143] In this way, cases that lie in the interstices of settled law are categorized as sitting on one side of the line or another, and the law's content and meaning are clarified through this articulation.[144] This process is a collaborative one, in which different actors in the system—lawyers, judges, defendants, law enforcement officials—collaborate in articulating and refining legal norms.[145] In other words, "law is our collective creation," not a set of rules mechanically applied.[146] Each probable-cause or reasonable-suspicion analysis is therefore an exercise in developing legal rules through a contest over which factors should be considered, how heavily they should be weighed, and what each ultimately means. This process cannot be reduced to numerical values, which may be one reason the Supreme Court has refused to identify either probable cause or reasonable suspicion as a particular numerical likelihood.[147]

143. *See* Brennan-Marquez, *supra* note 7, at 1292; Rich, *supra* note 7, at 887 (noting that individualized suspicion analysis has two steps: "[(1)] the events which occurred leading up to the stop or search, and then [(2)] the decision whether these historical facts, viewed from the standpoint of an objectively reasonable police officer, amount to reasonable suspicion or to probable cause" (alteration in original) (quoting Ornelas v. United States, 517 U.S. 690, 696 (1996))); *see also* Citron, *supra* note 16, at 1256 (describing errors written into computer code intended to implement Colorado's public benefits program).

144. Tribe, *supra* note 2, at 1391 (pointing out that legal "rules and institutions are often significant, not only as means of achieving various ends external to themselves, but also as ends in their own right, or at least as symbolic expressions of certain ends and values").

145. *See* Richard H. Fallon, Jr., *"The Rule of Law" as a Concept in Constitutional Discourse*, 97 COLUM. L. REV. 1, 19 (1997) ("[T]heories approaching the Legal Process ideal type tend to conceive the subjects of legal justice as reasonable persons, open to argument and persuasion, and deserving of reasoned explanations that the law should aspire to provide."); Frank Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, 87 GEO. WASH. L. REV. 1, 45 (2019) (explaining that where legal technology "reduces a legal relationship to a 'clear prescription' . . . it is unlikely to meet the complex standards of review and appeal embodied in the Legal Process conception of the rule of law" (quoting Fallon, *supra*, at 14)); Jeremy Waldron, *Theoretical Foundations of Liberalism*, 37 PHIL. Q. 127, 134 (1987) (arguing that in liberal society, "principles must be amenable to explanation and understanding, and the rules and restraints that are necessary must be capable of being justified to people who are to live under them").

146. *See* Andrew C. Michaels, *Abstract Innovation, Virtual Ideas, and Artificial Legal Thought*, 14 J. BUS. & TECH. L. 1, 30 (2019).

147. Brennan-Marquez, *supra* note 7, at 1266 ("Plausibility is qualitative, not quantitative. So it only stands to reason that the Court's analysis would shy away from numbers, in favor of narrative explanations.").

Second, use of a totality-of-the-circumstances standard seeks to minimize over- and under-inclusiveness by tailoring each decision to the specific case at hand by assessing all relevant factors. This means that unanticipated, anomalous, or unique factors should form part of the analysis, as do intangibles such as a decision-maker's empathy or compassion. Only the totality of an individual's circumstances can provide a basis on which to assess what those circumstances indicate about that individual's intent and volition.[148] Mere statistical probability might take some—or even many—features into account, but not all relevant features can be reduced to numerical value.

As the foregoing makes plain, the use of statistical thresholds or cost-benefit analyses alone to define the level of individualized suspicion necessary to justify government action cannot be sustained. To be sure, predictive accuracy must form a part of the individualized-suspicion inquiry. If the standard fails to actually identify likely instances of criminal activity at some reasonable rate, then it is not serving its purpose of channeling government discretion in a way that respects privacy, autonomy, or procedural justice. But something more is also required. So while statistical likelihood might be a satisfactory answer to the question "why this individual?," any acceptable articulation of individualized suspicion must go beyond that to address the question "why this individual *rather than any other*?" Probabilities and statistics cannot do so. The next Part turns to the implications of this analysis for the use of predictive machine learning.

IV. INDIVIDUALIZED SUSPICION AND ALGORITHMIC DECISION-MAKING

Having laid out both the values and the challenges that the individualized-suspicion requirement embodies as well as how it operates from a procedural perspective, this Part turns to the use of machine learning and its implications for individualized suspicion determinations. Section IV.A will provide a brief primer on the phenomenon of predictive analytics, focusing on the characteristics of the process that implicate individualized suspicion analysis. Section IV.B will then argue that reliance on predictive accuracy through artificial intelligence—even when that accuracy is very high—will not only implicate the same concerns as other forms of probabilistic evidence but will actually intensify those concerns.

148. *See* Andrew E. Taslitz, *Myself Alone: Individualizing Justice Through Psychological Character Evidence*, 52 MD. L. REV. 1, 16 (1993) ("Observers are often unaware of 'the variety of historical and contextual forces that impinge on [an individual actor's] behavior' in ways that are not immediately obvious." (alteration in original) (quoting RICHARD LEMPERT & JOSEPH SANDERS, AN INVITATION TO LAW AND SOCIAL SCIENCE: DESERT, DISPUTES, AND DISTRIBUTION 51–52 (1986))).

A. ALGORITHMIC DECISION-MAKING

Machine learning, a powerful form of datamining, is an umbrella term that encompasses many different techniques.[149] As a general matter, however, machine learning entails the deployment of an "algorithm"—a sequence of instructions telling a computer what to do—that examines an enormous data set and generates a "model" from it.[150] A model is the mathematical depiction of the relevant relationships among the data that the computer extracts.[151] This model can then be used to analyze new or existing data, often revealing patterns or relationships that humans could not have discovered unaided.

Most predictive computer models are developed through what is known as supervised machine learning.[152] This is a process used to develop a model that best captures the relationship between a set of "features" of the relevant data—its descriptive characteristics—and the "target feature"—the information we want to the computer to figure out.[153] It begins with a "training set," or a data set of examples for which we know both the characteristics of the data and the target feature.[154] For example, if we want to train a computer to recognize whether a particular image is an image of a cat, we would feed it a large database of images, some of which depict cats, as well as the appropriate label for each image—cat or not-cat. The computer will apply an algorithm to the labeled data to develop a model that identifies which images are cats. The hope is that this model will be able to accurately categorize new images as either cat or not-cat when it encounters them. For a model to be useful as a predictive tool, it must therefore be able to analyze accurately data that was not in the training set.[155] To determine whether the model can do so, usually some of the available data will be withheld from the training set to be used as a "test set," which will be given to the computer

149. For a more detailed discussion of the machine-learning process, see Berman, *supra* note 6, at 1284–90, from which this brief summary is drawn.

150. PEDRO DOMINGOS, THE MASTER ALGORITHM: HOW THE QUEST FOR THE ULTIMATE LEARNING MACHINE WILL REMAKE OUR WORLD 23–24 (2015).

151. Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 91 (2014) ("The goal of [a machine learning] algorithm is to build an internal computer model of some complex phenomenon . . . that will ultimately allow the computer to make automated, accurate classification decisions."). In essence, algorithms instruct computers to "figure it out on their own, by making inferences from data. . . . Now we don't have to program computers; they program themselves." DOMINGOS, *supra* note 150, at xi. Those relationships can be descriptive, meaning that they simply seek to identify properties of the available data set, or predictive, where the knowledge is extracted from known data for the purposes of predicting properties of new data. *Id.* at vx ("[A]t its core, machine learning is about prediction: predicting what we want, the results of our actions, how to achieve our goals, how the world will change."); *see also* FLACH, *supra* note 23, at 18–19.

152. There is also unsupervised, semi-supervised machine learning. *See* FLACH, *supra* note 14, § 1.1.

153. KELLEHER ET AL., *supra* note 15, at 3.

154. *Id.* at 4.

155. *Id.* at 5–11.

without the target variables.[156] The goal is to find a model that captures meaningful relationships between the data's features and the target variable such that when confronted with data it has not seen before, it will nevertheless produce the correct target variable.[157]

So in order to make a prediction, an algorithm constructs a model. The prediction is the model's determination regarding what is probable. Whether a pre-trial detainee is likely to appear for her court date. Whether an individual is likely to be transporting contraband. Whether a beneficiary of a government entitlement program is engaged in fraudulent activity. How accurate those conclusions are will depend on the chosen algorithm's accuracy rate, but they remain probabilities. A computer model is therefore conceptually akin to probabilistic evidence. The model might qualify as a very detailed generalization—e.g., unemployed males under age 25 who have a criminal history, no family support system, and less than a high school education are unlikely to appear for their court date and should therefore be denied bail—but they are generalizations that yield accurate answers at a predictable rate.

Embedded in this brief, highly simplified description of predictive machine learning are several important features of the process. First and foremost is the idea of predictive or preemptive decision making generally. To be sure, this feature is not unique to machine-learning analytics. Decisions regarding pretrial release, parole, sentencing, presence on the No-Fly list, where to deploy law enforcement resources, and many more are implicitly or explicitly premised on predictions about human behavior.[158] Nevertheless, as some commentators have pointed out, "big data's predictive benefits belie an important insight historically represented in the presumption of innocence and associated privacy and due process values—namely, that there is wisdom in setting boundaries around the kinds of assumptions that can and cannot be made about people."[159] In other words, just because a prediction can be made does not mean that making that prediction is consistent with the underlying values that have driven the development of individual rights

156. DAVID SKILLICORN, KNOWLEDGE DISCOVERY FOR COUNTERTERRORISM AND LAW ENFORCEMENT 76 (2008). If there is insufficient data to divide it into a training set and a test set, programmer can use "cross validation" to check the model's accuracy, in which the training data is sliced up in various ways for the model to analyze. *Id.* The more consistent the outcomes, the more accurate the model. *Id.*

157. Note that this simplified description of the process fails to capture the complexity of machine learning. The process is not a linear one, moving methodically from training to verifying to implementation. Rather, machine learning—like any form of data-mining—is a dynamic, iterative "process of retrieving, excluding, comparing[,] reorganizing, digging and pulling, etc." Liane Colonna, *A Taxonomy and Classification of Data Mining*, 16 SMU SCI. & TECH. L. REV. 309, 337–51 (2013); *see also* KELLEHER ET AL., *supra* note 15, § 1.5 (describing the phases of building a predictive model and noting the iterative nature of the process).

158. *See* Kerr & Earle, *supra* note 129, at 68–69.

159. *Id.* at 66.

protections.[160] Machine learning's potential for generating reliable predictions may tempt us to use them when we should not.

Second is the oft-repeated refrain that correlation does not necessarily mean causation. Regardless of how strong a correlation may be, "this knowledge may only concern populations while actions are directed towards individuals."[161] As artificial intelligence and computer-modeling experts have explained, current modeling capacity produces algorithms that are "statistically impressive, but individually unreliable."[162] In other words, they might be highly accurate in the aggregate but can still err in any given case. In fact, there may be multiple models that all reflect meaningful—though different—relationships among the data. As I have pointed out elsewhere, the result is that algorithms that are equally accurate, as an aggregate matter, can reach different results in individual circumstances.[163] Moreover, this variation can be significant. In the context of algorithmically determined creditworthiness scores, one commentator pointed out that, out "of 500,000 files, 29% of consumers had credit scores that differed by at least 50 points between the three credit bureaus."[164] So in any individual case, the prediction could depend on the model that is chosen just as much as it depends on an individual's characteristics or actions.

Another risk is that the correlation that the model identifies is not actually meaningful. Even if a model is consistent with the data on which it was trained, that may mean nothing more than it memorized the data.[165] Or it may have identified a pattern among the data that is simply not generalizable—meaning it will not necessarily describe accurately the relationship among data not included in the training set. Professor Frank Pasquale identifies an example where an algorithm sought to learn to differentiate between pictures of dogs and pictures of wolves. The algorithm initially seemed remarkably accurate. Upon closer inspection, however, researchers determined that the feature that the algorithm relied upon most

160. *See id.* at 70–71 ("[T]he presumption of innocence and related private sector due process values can be seen as wider moral claims that overlap and interrelate with core privacy values. Taken together, privacy and due process values seek to limit what the government (and, to some extent, the private sector) is permitted to presume about individuals absent evidence that is tested in the individuals' presence, with their participation.").

161. Brent Daniel Mittelstadt et al., *The Ethics of Algorithms: Mapping the Debate*, BIG DATA & SOC'Y, July–Dec. 2016, at 1, 5 (citation omitted).

162. John Launchbury, Director, DARPA Information Innovation Office, PowerPoint Presentation, A DARPA Perspective on Artificial Intelligence, at slide 23 (Feb. 15, 2017), *available at* https://www.darpa.mil/about-us/darpa-perspective-on-ai [https://perma.cc/F7UD-CNMR]; *see also* Mittelstadt et al., *supra* note 161, at 5 (noting that algorithms infer correlations based on "populations while actions are directed toward individuals").

163. *See generally* Berman, *supra* note 6 (identifying multiple uses of predictive analytics in the governmental context and discussing the variety of results that could be produced).

164. Citron & Pasquale, *supra* note 6, at 12.

165. KELLEHER ET AL., *supra* note 15, at 7–8. Note that there may be several models that are consistent with the data set from which an analyst must choose one. *See id.* at 17.

heavily in distinguishing between the two different animals was the presence of snow in the photograph.[166] Given pictures of dogs in a snowy landscape or wolves in a green field, the model was a failure.[167]

Third, when decision-makers rely on predictive machine-learning, they accept the inevitability of a certain number of erroneous decisions.[168] Predictive analytics might yield a relatively accurate rule. As with all rules, however, it will be both over- and under-inclusive. And with a computer model, we will know from its performance what those error rates will be. Even if an algorithm is 90 percent accurate, a choice to implement that algorithm reflects an explicit acceptance that 10 percent of decisions will go awry. We thus know that use of an algorithmic risk assessment that returns one false positive for every nine accurate predictions will deny bail unnecessarily to ten people out of 100. In other words, use of predictive algorithms recognizes that errors will, in fact, happen at a predictable rate. As many scholars have pointed out, "[a]ll evidence is probabilistic, in the sense that there is a risk of error in relying on it to support a factual conclusion."[169] Indeed, there will always be some errors regardless of how a decision is made. The very existence of legal standards rather than rules—reasonable suspicion, probable cause, preponderance of the evidence—is a recognition of this inevitability.[170] Nevertheless, the use of predictive models accepts the inevitability of erroneous determinations in some cases in order to achieve aggregate accuracy.

Fourth, any information that contributes to the development of a computer model must come in a form that can be entered into a database. In the world of machine learning, "[i]f you can't count it, it doesn't exist."[171] Data-driven tools press policymakers to focus on data-driven factors. This means consideration of factors that are quantifiable, rather than factors arising from human judgment or policy decisions that cannot be expressed in ones and zeros. As a result, automated systems are best used for enforcing

166. *See* Marco Tulio Ribeiro et al., *"Why Should I Trust You?": Explaining the Predictions of Any Classifier*, ACM, Aug. 2016, at 8–9, *available at* https://arxiv.org/pdf/1602.04938.pdf [https://perma.cc/5ZH6-LAJP].

167. *Id.*

168. *See* Jonathan J. Koehler & Daniel N. Shaviro, *Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods*, 75 CORNELL L. REV. 246, 251–52 (1990) (recognizing that use of "[o]vertly probabilistic evidence . . . makes the risk of error explicit"); Roth, *supra* note 90, at 1165.

169. Koehler & Shaviro, *supra* note 168, at 252; Bacigal, *supra* note 3, at 295 ("All proof is ultimately 'probabilistic' in the sense that no conclusion can ever be drawn from empirical data without some step of inductive inference.").

170. Bambauer, *supra* note 4, at 482 (arguing that doubts and errors are part of criminal procedure).

171. *See* Tribe, *supra* note 2, at 1361–66. *See generally* BERNARD E. HARCOURT, AGAINST PREDICTION: PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE (2007) (arguing that actuarial methods used in the criminal legal system are increasingly erroneous, leading to increased profiling and punishment that ultimately harms people).

rules, not for enforcing standards, which "explicitly or implicitly require the exercise of human discretion."[172]

Finally, many computer models that emerge from the machine-learning process cannot be explained in terms intelligible to humans—they cannot provide an explanation for their outputs and predictions that humans can understand.[173] They often do not identify the data on which they rely in generating their model, nor how certain features are weighed relative to others. They represent a black-box phenomenon where the inputs go in and the outputs emerge, but there is no means of tracking or describing what happens in between. Moreover, there are reasons to doubt that machine-learning algorithms can ever be rendered intelligible without sacrificing the analytical value for which they are prized.[174]

B. Algorithms & Individualized Suspicion

Taking into account these inevitable features of machine-learning predictions, it becomes clear that there is no way to analyze individualized suspicion requirements through machine learning and simultaneously preserve individualized suspicion's non-accuracy-focused purposes. First consider the idea of articulating a justification for a decision. Algorithms' use of correlation to predict outcomes means that there is no way to determine whether a prediction is inferred through seeking the "best" explanation because machine-learning predictions do not seek to provide narrative explanations. Like probabilistic evidence from other sources, they represent associative hypotheses—the proposition that an association exists between two or more factors—as opposed to telling a story. Computer models do not purport to identify combinations of factors that *explain* the resulting prediction. They simply identify patterns within data and infer correlations from those patterns. In generating those correlations, they might rely on categories of information that humans might not consider relevant—or at the very least not the best explanation—even if it does turn out to be an accurate predictor. A machine-learning algorithm might, for example, prove highly accurate in distinguishing between dogs and wolves, but whether that is due to some intrinsic characteristics of wolves qua wolves or due to the prevalence of snow in pictures of wolves is not something the algorithm can explain to

172. Citron, *supra* note 16, at 1304.

173. *See* FLACH, *supra* note 14, at 27–29, 32; *see also* RUSSELL & NORVIG, *supra* note 17, at 707.

174. *See* Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 FORDHAM L. REV. 1085, 1126–29 (2018) (arguing that human means of normatively evaluating explanations can never effectively be applied to algorithms). Nevertheless, computer scientists continue their efforts to improve interpretability of algorithmic models. *See, e.g.*, Alfredo Vellido et al., *Making Machine Learning Models Interpretable*, *in* EUROPEAN SYMPOSIUM ON ARTIFICIAL NEURAL NETWORKS, COMPUTATIONAL INTELLIGENCE AND MACHINE LEARNING 163–65 (2012). These efforts may one day minimize or eliminate the opacity of machine learning predictions, but, as Selbst and Barocas argue, *supra*, they may not.

us.[175] This means that not only do we lack a narrative justification for government action, but it is also difficult to assess the role that features we might consider invalid—race or sex, for example—are playing.

Further complicating the need for explanation is the fact that the most powerful (i.e., useful) algorithms are unintelligible. When algorithms are unintelligible, no explanation can be forthcoming. An individual must forego an explanation of why *she*, rather than someone else, was searched. How then, can an individual seek to establish that an algorithm is in error, or appeal to a higher authority, or contest what the probable cause standard actually means. Other black-box tools that the legal system relies upon, such as DNA tests, radar guns, and drug-sniffing dogs, seek to answer questions of fact —whether a DNA sample from the crime scene matches that of the defendant, whether a car was moving faster than 65 miles per hour, whether drugs are present. Predictive analytics, by contrast, go beyond yes or no factual questions to predictions regarding whether a whole host of factors adds up to suspicion justifying government action.[176] Even if an algorithm is, in fact, intelligible to humans, in some instances the explanation for the decision would be "because we chose computer model A instead of computer model B, and while we know that each of these models makes errors, there is no means available to identify which individual case represents one of those errors." Efforts to find fault with or challenge such an explanation would face obstacles similar to those encountered when the government offers no justification at all.

Even if computer models could provide detailed justifications for their decisions, those decisions still might not truly be made on the basis of the totality of the circumstances. Machines can only take account of information that is provided to them, which may or may not include all factors necessary to reach an accurate result. Of course, there is no guarantee that a human decision-maker will be apprised of all possible relevant facts either. A human decision-maker, however, "is not committed in advance of decision to the factors that will be considered and the rule for combining them. He is free to respond to individual differences whose relevance was not anticipated."[177] Moreover, there is some information that goes into legal decision-making that lends itself to empirical treatment and may be reduced to ones and zeroes relatively easily (e.g., a criminal record); other information, however, does not

175. *See* Ribeiro et al., *supra* note 166, at 8–9.

176. Rich, *supra* note 7, at 892.

177. Barbara D. Underwood, *Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgment*, 88 YALE L.J. 1408, 1423 (1979); *see also* Brennan-Marquez, *supra* note 7, at 1298 ("To say that a problem is best resolved by prudence rather than principle is to express doubt about the possibility of fashioning second-order rules for navigating the collision between first-order values. Prudence becomes important, in other words, to the extent that conflict between competing goods is hard to reduce to fixed equations. When that happens, case-specific judgments—as opposed to generalized principles—must carry the day.").

(e.g., intent).[178] As a result, decision-making methods, like computer models, that rely on objectively measurable factors will inevitably "shift the focus away from such elements as volition, knowledge, and intent, and toward such elements as identity and occurrence."[179] There are also factors relevant to legal decision-making that machines are simply incapable of taking into account. Any emotional component that we might want decision-makers to employ—compassion, empathy, hope, anger, fear—is an intangible not susceptible to modeling.

Use of algorithmic predictions also poses challenges for providing a meaningful opportunity to contest or refute the government's account. To mount a coherent challenge to a particular decision, we must know how that decision was made. With machine learning, this will often be difficult to discern because scrutiny of the outcome of a computer model cannot necessarily reveal what exactly is included in the model, how each factor is weighed, or whether there are factors included in the model that perhaps should *not* be taken into account.[180] If a computer model identifies a dorm room as likely to contain marijuana, what factors did it use to generate that result? Was it simply that the dorm is on a campus known for pervasive marijuana use? That it is occupied by a male student? That the occupant has long hair and wears Grateful Dead T-shirts?[181] How do we know that the factors relied upon are sufficiently rare among the general population to be considered "individualized"?[182] Individualized suspicion doctrine requires that—at the time a government official imposes burdens on an individual's privacy or autonomy, she has reason to think that this specific stop, search, arrest, or other action is justified. Thus, the human-driven regime preserves the opportunity to strive for accurate results and bars action in the absence of a plausible basis for decision-makers to believe that they have reached the correct result. If, on the other hand, such decisions are instead based on computer models that classify certain individuals as permissible subjects of, say, a stop or a search, the absence of a convincing narrative justification makes the right to challenge government decision-making all the more difficult.

But it goes beyond that. The values underlying individualized suspicion demand that burdens are imposed on citizens because of what they themselves have done, not because they happen to be unlucky enough to

178. *See* HARCOURT, *supra* note 171, at 31–34, 188–89. *See generally* Taslitz, *supra* note 19, at 165–68 (asserting that individualized suspicion is the heart of probable cause doctrine).

179. Tribe, *supra* note 2, at 1366.

180. Protected classes such as race and gender might qualify.

181. *See* SCHAUER, *supra* note 2, at 9–19 (discussing the importance of making informed generalizations). *See generally* Taslitz, *supra* note 19 (describing a wide variety of scenarios and factors that constitute reasonable suspicion on a case-by-case basis).

182. *See* Andrew D. Selbst, Response, *A Mild Defense of Our New Machine Overlords*, 70 VAND. L. REV. EN BANC 87, 99–101 (2017).

share the features that an algorithm determines to be relevant. When the decision-maker is a computer model, an individual has no means of arguing that the correlations identified in the algorithm are inaccurate as applied to her; no way to show that unquantifiable factors that distinguish her as an autonomous, unique individual, dictate a different outcome; no way of ensuring that all relevant circumstances are taken into account.[183] We are reduced to a list of attributes. As a result, predictive analytics disregard one of the crucial elements of autonomy—free will. Decisions based on predictive tools operate as if the future is set in stone. Their predictions, based on correlations of quantifiable information, will be based on factors such as identity (e.g., age, race, gender) or past actions (e.g., criminal history, education, drug abuse). An individual need not necessarily be connected to any inherently suspicious activity—someone is suspicious only because the model said they were. Yet at any point, any individual can make choices that defy predictions about them.[184] Individuals grow and develop. They have the capacity to change the path they are walking down. To base decisions on the assumption that human behavior can be predicted thus fails to recognize that nothing is preordained and places an obstacle to consideration of a counterargument in a defendant's way.

Concern regarding the inconsistency between predictive analytics and free will is compounded by the fact that government use of predictive analytics can operate as self-fulfilling prophecies. Some predictive modeling operates independently of the outcomes it predicts. Take life insurance, for example. If an actuarial table predicts a certain life expectancy for a non-smoking woman with a family history of breast cancer, and an insurance company bases her life insurance premiums on that data, that prediction is unlikely to impact the woman's life span.[185] The same cannot be said for many of today's uses of machine learning by the government. Instead, there are two ways in which computer models can affect the phenomena they purport to predict. First, they can focus government attention on certain populations. Consider, for example, an algorithm that predicts high likelihood of drug-related activity in a particular neighborhood.[186] This prediction then will prompt law enforcement to devote significant resources to investigating drug crime in that neighborhood, with the result that police will find evidence of such crimes. It is impossible to know in such instances whether a different

183. *See* Taslitz, *supra* note 19, at 165–68.

184. *See* Underwood, *supra* note 177, at 1414 ("[R]espect for individual autonomy requires recognition of the possibility that an individual can choose to refute any prediction about himself.").

185. *See* HARCOURT, *supra* note 171, at 185–86.

186. It is well-documented that crime-related data will disproportionately identify poor and minority communities as loci of crime because those are the neighborhoods that historically have been heavily policed, thereby generating a disproportionate amount of data indicating criminal activity there. *See* HARCOURT, *supra* note 171, at 112–19; Ferguson, *supra* note 13, at 401–03; Selbst, *supra* note 182, at 99–102.

allocation of resources would have resulted in different outcomes. Perhaps devoting similar law enforcement resources to investigating drug-related activity on the local college campus rather than the low-income, minority-dominated neighborhood would yield similar numbers of drug crimes. But if law enforcement acts according to the model's predictions, it will always focus on the same neighborhoods, thereby making it likely that they will find crime there, which will, in turn, strengthen "the seeming validity of the profile even if it does not match the" full picture of who is engaged in drug crime.[187]

In other words, computer models can only measure one form of error —false positives. False negatives will, by definition, go undetected. Thus, if a predictive tool grows more accurate over time—as machine-leaning algorithms are meant to do—we cannot know whether that is because the model is better reflecting the state of the world, or whether the model is actually generating the world it purports to describe. It is hard to imagine government action more inconsistent with respect for individual autonomy than using the power of the state to determine someone's life path. In addition to undermining autonomy, the self-fulfilling prophecy of predictions will reinforce existing disparities based on race and poverty and exacerbate procedural justice concerns.[188]

A similar concern arises when computer models are used in contexts where their determinations are non-falsifiable. With respect to most individualized suspicion determinations, this will not be the case. If a criminal defendant has evidence seized from his home in a search not supported by probable cause, he will both know that the search took place and have an opportunity to challenge its constitutionality in court. Other errors, however, will be insulated from challenge. Consider a determination, based on reasonable suspicion, that a non-citizen poses security concerns and should therefore be denied entry into the country.

Effectively implementing the error-correction methods that individualized decision-making demands is also made more difficult when the government accepts that a predictable number of decisions will be incorrect. When we rely on an algorithm, we know its rate of accuracy. Therefore, we might know that ten out of every 100 decisions will be wrong, but we lack means to determine on a case-by-case basis which ten of the 100 outcomes are erroneous. Such decision-making, in a sense, "gambles" with a citizen's liberty or privacy, placing the state's stamp of approval on burdening some number of innocent individuals.[189] Nine out of ten individuals identified by an algorithm might possess evidence of criminal activity, but if there is no story

187. Ferguson, *supra* note 7, at 297.

188. *See* Ferguson, *supra* note 13, at 401 (noting that the burden of false positives will likely fall most heavily on individuals with prior interaction with the criminal justice system, deriving suspicion from that correlation and that, as a result, "[t]hose with lengthy criminal records or gang associations may be stopped because of who they are and not what they are doing").

189. Bacigal, *supra* note 3, at 295–97.

that points to any one of those ten above the others, there is no normic justification for the intervention. Our innate, "perhaps culturally instilled, American sense of fairness"[190] rebels against the idea of passing judgment on individuals purely based on the statistical likelihood that they are "one among many, or even a few, who could have" acted in the relevant fashion.[191] Of course all forms of human decision-making will produce errors. And many, if not most, decision-making methods will be less accurate than a computer algorithm. But the requirement that an officer have reason to believe that *this person* is engaged in criminal activity seems to reject the use of a system where the best we can say is that there is reason to entertain that belief with respect to nine out of these ten people. And any system that insists upon individuals' right to challenge the government determination surely contemplates something other than the response that "you must just be one of the 10% of cases where we will be wrong."

Some scholars suggest that use of machine learning in the Fourth Amendment context is not problematic so long as a human retains the final decision-making authority.[192] But the "human-in-the-loop" solution requirement does not alleviate the concerns identified above. A computer model's prediction subsequently confirmed by a judge or police officer is no more intelligible to the subject of a search or arrest than the computer model itself. Moreover, absent conclusive evidence of error, human decision-makers are likely to defer to an algorithm's output, regardless of how they would have made the decision acting alone. Predictive analytics, when viewed from the outside, seems like a precise scientific calculation, rather than the probabilistic judgment that it actually represents. Government officials' ability to effectively assess such a judgment, especially in the absence of information regarding how the judgment was reached, is too thin a reed on which to place our Fourth Amendment rights.

V. CONCLUSION

The awesome power of the state and law enforcement's broad discretion combine to raise real and significant concerns regarding how to best constrain government action. In such a world, it is tempting to seek to rely on hard and fast data, information that is not subject to interpretation or manipulation. Translating the idea of individualized suspicion into a mathematical (or algorithmic) inquiry, however, is not the answer. Even if such measures alleviate concerns about the use of state discretion—a dubious claim at best—they do so at the expense of other, equally valuable goals that individualized

190. Taslitz, *Police Are People Too*, *supra* note 82, at 67.

191. Bacigal, *supra* note 3, at 298; *see also* Taslitz, *Police Are People Too*, *supra* note 82, at 15 (arguing that moral concepts of just deserts are guided by a wrongdoer's intentions and her ability to do otherwise).

192. *See* Ferguson, *supra* note 7, at 311; Simmons, *supra* note 7, at 1017.

suspicion is designed to promote. Any viable definition of individualized suspicion needs to ensure that government decisions are both sufficiently likely to yield criminal evidence and true to our commitment to human dignity, autonomy, and procedural justice.