

Michigan Law Review

Volume 113 | Issue 4

2015

Hassle

Jane Bambauer

University of Arizona James E Rogers College of Law

Follow this and additional works at: <http://repository.law.umich.edu/mlr>



Part of the [Fourth Amendment Commons](#), and the [Law Enforcement and Corrections Commons](#)

Recommended Citation

Jane Bambauer, *Hassle*, 113 MICH. L. REV. 461 (2015).

Available at: <http://repository.law.umich.edu/mlr/vol113/iss4/1>

This Article is brought to you for free and open access by the Michigan Law Review at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Michigan Law Review by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

HASSLE

Jane Bambauer*

Before police perform a search or seizure, they typically must meet the probable cause or reasonable suspicion standard. Moreover, even if they meet the appropriate standard, their evidence must be individualized to the suspect and cannot rely on purely probabilistic inferences. Scholars and courts have long defended the distinction between individualized and purely probabilistic evidence, but existing theories of individualization fail to articulate principles that are descriptively accurate or normatively desirable. They overlook the only benefit that the individualization requirement can offer: reducing hassle.

Hassle measures the chance that an innocent person will experience a search or seizure. Because some investigation methods meet the relevant suspicion standards but nevertheless impose too many stops and searches on the innocent, courts must have a lever independent from the suspicion standard to constrain the scope of criminal investigations. The individualization requirement has unwittingly performed this function, but not in an optimal way.

Individualization has kept hassle low by entrenching old methods of investigation. Because courts designate practices as individualized when they are costly (for example, gumshoe methods) or lucky (for example, tips), the requirement has confined law enforcement to practices that cannot scale. New investigation methods such as facial-recognition software and pattern-based data mining, by contrast, can scale up law-enforcement activities very quickly. Although these innovations have the potential to increase the accuracy of stops and searches, they may also increase the total number of innocent individuals searched because of the innovations' speed and cost-effectiveness. By reforming individualization to minimize hassle, courts can enable law-enforcement innovations that are fairer and more accurate than traditional police investigations without increasing burdens on the innocent.

TABLE OF CONTENTS

INTRODUCTION	462
I. INDIVIDUALIZED SUSPICION AMISS	468
A. <i>The Snowflake Theory</i>	469
B. <i>The Felt Belief Theory</i>	473
C. <i>Suspicious Conduct Theories</i>	478
D. <i>Crime-Out Investigations</i>	479

* Associate Professor of Law, University of Arizona James E. Rogers College of Law. Deep gratitude is owed to Derek Bambauer, Barbara Atwood, Josh Blackman, Shun-Ling Chen, Ed Cheng, Mike Chiorazzi, James Cooper, Deven Desai, Jeffrey Fagan, Kristine Huskey, Shasta Inman, Orin Kerr, Jason Kreag, Dave Marcus, Toni Massaro, Jeff Rachlinski, Carol Rose, Roy Spece, Rodney States, Alan Trammell, and Brent White for excellent feedback, and to Melissa Ahlers and Nathan Berry for research assistance.

E. <i>Ending the Status Quo</i>	481
II. HASSLE	482
A. <i>A Compact Model for Criminal Investigations</i>	483
B. <i>The Importance of Hassle</i>	487
C. <i>Individualization Reduces Hassle</i>	490
D. <i>Other Instincts Explained</i>	494
1. Odd but Legal Behavior	495
2. A Preference for More Factors	496
3. Undefined Suspicion Standards	498
4. The Search for Distributional Justice	498
III. THE FUTURE OF SUSPICION	501
A. <i>More Open and Honest Experimentation</i>	501
B. <i>Fewer Vague Narratives</i>	504
C. <i>More Randomness</i>	506
D. <i>Objections</i>	508
CONCLUSION	510

INTRODUCTION

A police officer has submitted an application to a magistrate judge requesting warrants to search every dorm room in the Harvard College residence halls for illegal drugs. To establish probable cause, the officer furnished a copy of a recently published study about student life at Harvard. The study is methodologically sound and concludes that 60% of the on-campus dorm rooms contain illicit drugs.

Naturally, the magistrate judge will deny the application, but why? While probable cause is not defined as a precise probability, we are told that the standard sits below “more likely than not.”¹ A 60% likelihood of finding drugs in any given dorm room should easily clear the bar.

The textbook response to this Harvard dorm hypothetical, a variant of a puzzle posed by Orin Kerr,² is that the suspicion stemming from a statistical study is not *individualized*. The evidence is not tailored to each Harvard student whose home is to be searched. The study offers only probabilistic

1. *Illinois v. Gates*, 462 U.S. 213, 235 (1983); *Texas v. Brown*, 460 U.S. 730, 742 (1983) (plurality opinion).

2. Orin Kerr, *Why Courts Should Not Quantify Probable Cause*, in *THE POLITICAL HEART OF CRIMINAL PROCEDURE* 131, 135–37 (Michael Klarman et al. eds., 2012). This is actually a variant on Kerr’s original hypothetical. Kerr asked what should happen if the police request a warrant to search one specified Harvard student’s dorm in order to explore the importance of back stories and the fruitless efforts of the police that may not be reported to judges. *Id.* Kerr was chiefly concerned about the potential that the officer selected the dorm room for an improper reason. *Id.* I changed the hypothetical to cover all Harvard dorm rooms in order to reduce the potential of abuse and to illustrate how hassle animates intuitive responses to this hypothetical.

evidence. No matter how great the chances of finding drugs may be, lack of individualization presents a disqualifying flaw.³

Individualized suspicion therefore consists of two distinct prongs. The police must have suspicion—a fairly good chance of finding evidence of a crime—and they also must have individualization. This Article explores the second prong, the individualization requirement, in search of a principled distinction between particularized evidence and evidence that is probabilistically sufficient but constitutionally flawed.

The predominant justification for the individualization requirement appeals to a rejection of cold calculations and group-based generalizations, but this appeal falls flat with a little scrutiny. Police and judges must always resort to rough estimates of the conditional probability that a suspect has engaged in crime. Whether police use a collection of details (a partially corroborated tip and unusual travel habits) or just one important detail (a weaving car or a tattoo matching a victim's description), they ultimately must ask whether the innocent explanations for the observations are much more probable than the illicit ones.⁴ Indeed, even vocal opponents of actuarial policing have acknowledged that *all* suspicion is built from probabilistic inferences.⁵

Moreover, the distinctions that we intuitively draw between individualized and merely mathematical suspicion get the public policy backwards. Because the investigation methods approved by courts usually rely on the observations and perceptions of police, the “particularized” evidence is likely to be biased, error prone, and disproportionately aimed at poor and minority residents living in higher-crime areas.⁶ Subjective factors like a suspect's “nervousness” or “furtive movements” can be imagined or, worse still, manufactured through deceit.⁷ And the long, detailed narratives that courts have

3. See *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) (separately listing “reliability” and “particularity” requirements for evidence supporting probable cause); Sherry F. Colb, *Probabilities in Probable Cause and Beyond: Statistical Versus Concrete Harms*, LAW & CONTEMP. PROBS., Summer 2010, at 69, 70.

4. See Colb, *supra* note 3, at 78.

5. Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1330 n.2 (1971); see also BERNARD E. HARCOURT, *AGAINST PREDICTION: PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE* 17 (2007).

6. This occurs not only or primarily due to implicit bias or prejudice among police. It is also the natural result of deploying police in greater numbers to high-crime neighborhoods. See, e.g., William J. Stuntz, *Unequal Justice*, 121 HARV. L. REV. 1969, 1997 (2008).

7. See Report of Jeffrey Fagan at 51 tbl.11, *Floyd v. City of New York*, 959 F. Supp. 2d 540 (S.D.N.Y. 2013) (No. 08 Civ. 01034) [hereinafter Report of Jeffrey Fagan], available at http://ccrjustice.org/files/Expert_Report_JeffreyFagan.pdf. The Fagan Report found that, among the more than 2 million stops performed by the New York Police Department between 2004 and 2009 that had sufficient documentation for analysis, 42% listed “furtive movements” among the reasons justifying the stop, and 10.4% listed “suspicious bulge.” *Id.* When the police suspected a weapons violation, “furtive movements” was reported 60% of the time, and “suspicious bulge” over 34% of the time. *Id.* Police frisks uncovered weapons only 15% of the time. *Id.* at 64 tbl.15.

come to expect from police in order to satisfy particularization requirements are so inconsistent that they risk diluting the suspicion requirement.⁸ Individualization seems to be intellectually bankrupt and morally hazardous.

Nevertheless, our fumbling with the notion of individualization performs latent but valuable work. Given that all evidence is probabilistic, the virtue of individualization has little to do with the nature of probabilistic calculations. It doesn't even have much to do with the particular suspect. Instead, individualization protects everybody *else* from the potential costs of law-enforcement investigations. The techniques courts and scholars accept as individualized exclude most of the population from the practical likelihood of police intrusions. Because they rely on perceptions of police officers or on happenstance like tips, the traditional methods cannot scale. Actuarial methods can.

This Article argues that the purpose of individualization is to minimize hassle. *Hassle* is the chance that the police will stop or search an innocent person against his will.⁹ After all, while we may not all be Harvard students, we do inevitably engage in activities that predict a high chance of crime. We pace. We circle the block. We travel with bulky or light luggage. And we attend Phish concerts. If the police were able to act on all reasonably predictive statistical models en masse, we would experience an inappropriate and dramatic increase in suspicion-based searches and seizures. The individualization requirement constrains hassle by ensuring that an innocent person is unlikely to be stopped or searched even if he seems suspicious from time to time.

The twin prongs of individualized suspicion—suspicion and individualization—ought to be loosely guided by hit and hassle rates. A hit rate is the probability that a stop or search will uncover evidence of a crime. Hit rates measure suspicion and must meet the relevant standard (reasonable suspicion for *Terry* stops and probable cause for full-blown searches and

8. Justice Marshall described the reasonable suspicion standard as having a “chameleon-like way of adapting to any particular set of observations.” *United States v. Sokolow*, 490 U.S. 1, 13 (1989) (Marshall, J., dissenting) (quoting *United States v. Sokolow*, 831 F.2d 1413, 1418 (9th Cir. 1987)) (internal quotation marks omitted). In his dissent in *Sokolow*, Justice Marshall included a string citation with pieces of evidence that had been used to build suspicion in previous cases: suspect was first to deplane; last to deplane; deplaned from the middle; purchased one-way tickets; purchased round-trip tickets; traveled on a nonstop flight; changed planes; traveled with no luggage; traveled with a gym bag; traveled with new suitcases; traveled alone; traveled with a companion; acted nervous; acted too calm. *Id.* at 13–14 (citing ten cases).

9. I use the word “hassle” to capture a broad range of governmental impositions, from inconveniences to serious affronts to dignity. The word will fail to reflect the depth of intrusion in some cases (invasive searches, wrongful arrests, and uses of force, for example). But “hassle” does reflect the disruption of liberty that carries Fourth Amendment significance in even the most minor cases of involuntary searches or seizures. Part II defines the term more precisely.

seizures).¹⁰ Hassle rates, by contrast, measure the probability that an innocent person within the relevant population will be stopped or searched under the program. Hassle rates speak to individualization. If a program is likely to cause too much hassle, the police have not sufficiently narrowed the scope of the investigation, no matter how high the hit rate may be. Hassle rates keep track of the societal costs of criminal investigations, and hit rates ensure that the costs are justified.

Courts and scholars have already grown accustomed to examining hit rates when data are available.¹¹ When they are not, courts and scholars have used common sense and experience to estimate the same thing: whether the police had a decent shot of uncovering incriminating evidence during a stop or search. But hit rates alone cannot keep the government in check. Hassle rates are also crucial to the Fourth Amendment's protection. For rare crimes, like murder, a high hit rate can guarantee a low hassle rate. But for more common crimes, such as drug possession,¹² an additional constraint must curb governmental intrusions. At a higher level of abstraction, hit and hassle perform the delicate balancing of interests that the Fourth Amendment demands.

Hassle explains many of the instincts already embedded in the individualization precedent. For example, it explains the courts' consistent preference for police narratives chock-full of detail, even when each additional detail does not contribute much to the amount of suspicion.¹³

Hassle can also explain the Harvard dorm room hypothetical. If an officer requests warrants to search the dorms of all 6,000 students living in Harvard residential halls, we know *ex ante* that approximately 2,400 of them will not have illegal drugs. In one fell swoop, the police will have imposed significant costs on the innocent population in the Harvard and Cambridge communities.

On the other hand, the Harvard dorm room hypothetical also commands attention to a lost opportunity. If the warrant were issued, the Cambridge police could have greater success searching a Harvard student and could create inroads for criminal enforcement within elite communities otherwise immune to the enforcement of minor criminal laws. But the dominant understanding of individualization will push the police out of the Harvard dorms and back into the homes and pockets of the poor, the uneducated, and the traditionally suspect.

10. *Terry v. Ohio*, 392 U.S. 1, 21 (1968) (requiring articulable suspicion for a stop); *Carroll v. United States*, 267 U.S. 132, 149 (1925) (requiring probable cause for a search).

11. See, e.g., *Floyd v. City of New York*, 959 F. Supp. 2d 540, 558 (S.D.N.Y. 2013); DAVID A. HARRIS, *PROFILES IN INJUSTICE: WHY RACIAL PROFILING CANNOT WORK* 13–14, 79–84 (2002).

12. See NAT'L CTR. FOR HEALTH STATISTICS, U.S. DEP'T OF HEALTH & HUMAN SERVS., HEALTH, UNITED STATES, 2012, at 193–94 tbl.58 (2013) [hereinafter CDC DRUG USAGE TABLE], available at <http://www.cdc.gov/nchs/data/hus/hus12.pdf> (finding that, in 2012, 9.2% of Americans aged 12 and older had used illicit drugs in the past month and that 21.3% of Americans between the ages of 18 and 25 had done so).

13. See *infra* Section II.D.2.

Courts and advocates should reform the concept of individualization to focus on minimizing hassle.¹⁴ And they can do so without making significant changes to existing doctrine. This Article explains why and how in three parts.

Part I collects the definitions and justifications for individualization that have floated around the legal scholarship over the last five decades. Individualization has attracted the attention of a long list of distinguished scholars, and with the exception of Fred Schauer, all have vigorously defended the concept.¹⁵ Charles Nesson champions the notion of case-by-case assessment.¹⁶ Laurence Tribe highlights the importance of human intuition.¹⁷ Andrew Taslitz argues that suspicion should be based on conduct under the control of the suspect,¹⁸ and Bernard Harcourt suggests that police should trace suspicion from a crime to a suspect rather than attempt to predict which individuals are criminals.¹⁹ Each of these theories fails to describe actual or desirable outcomes. Nesson's and Tribe's theories entrench the discretion of police officers in the teeth of ample evidence of bias and error, while Taslitz's and Harcourt's theories put impracticable limits on criminal investigation.

There is a way out. Part II introduces hassle, a concept that operates behind the scenes of the individualization doctrine and deserves attention. Hassle measures how much pain an investigatory program will impose on the innocent even when the program is moderately successful at detecting crime. Even if a new police tool does a very good job of detecting suspicious conduct, if the tool is inexpensive and used with abandon, the hassle it brings to the wrongly suspected should justify Fourth Amendment scrutiny on its own. The concept of hassle will become increasingly valuable in an era of technological change. Indeed, the Foreign Intelligence Surveillance Court

14. I do not mean to imply that data science is immune from human error and influence. See, e.g., Kate Crawford, *The Hidden Biases in Big Data*, HARV. BUS. REV. BLOG NETWORK (Apr. 1, 2013, 2:00 PM), <http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/>. Data-driven policing is bound to have its problems, but it beats the alternatives.

15. FREDERICK SCHAUER, *PROFILES, PROBABILITIES, AND STEREOTYPES* (2003); see also Colb, *supra* note 3, at 71, 102 (neither rejecting nor endorsing the distinction between probabilistic and concrete harms but recognizing that such a distinction is "irrational" and "at times arbitrary").

16. Charles R. Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187, 1191–92 (1979).

17. See Tribe, *supra* note 5, at 1332–38.

18. Andrew E. Taslitz, *What Is Probable Cause, and Why Should We Care?: The Costs, Benefits, and Meaning of Individualized Suspicion*, LAW & CONTEMP. PROBS., Summer 2010, at 145, 146.

19. See HARCOURT, *supra* note 5, at 18. Harcourt also distinguishes the actuarial policing that he discourages from "clinical methods" that "rel[y] primarily on the subjective judgment of experienced decision makers." *Id.* at 16 & 269 n.48. But Harcourt has also criticized the case law on individualized suspicion, calling the concept "a rhetorical placeholder used to bless police practices without providing policing agencies with any guidance or requirements." Bernard E. Harcourt & Tracey L. Meares, *Randomization and the Fourth Amendment*, 78 U. CHI. L. REV. 809, 851 (2011).

reached for the concept of hassle without having a name for it when the court invalidated aspects of the National Security Agency (“NSA”)’s Upstream program.²⁰ And hassle matches and explains judicial instincts about individualization better than other theories. Although courts are likely to resist defining hassle thresholds in precise statistical terms, hassle provides a badly needed benchmark. It fills the vacuum in individualization’s organizing principles.

Part III explores the implications of hit- and hassle-based individualized suspicion. Not only is a hassle-based individualization requirement more descriptively accurate but it also has desirable normative implications for police practices. It tolerates innovation in policing. It also permits law-enforcement agencies to limit their operations using carefully designed random selection instead of relying on the luck and resource constraints that currently control their scope.²¹ As long as a program has a high enough hit rate to satisfy suspicion requirements, the police should be permitted to keep hassle rates low by selecting among suspicious individuals in a mechanical but evenhanded way. For example, suppose new software can analyze video footage from security cameras to detect hand-to-hand drug sales with a high hit rate (that is, most of the time the software alerts, it is correct). Because of the high frequency of this sort of crime, incessant monitoring throughout a city is likely to produce too much hassle. Rather than abandoning the software altogether, the police could reduce hassle by responding only to a randomly selected portion of the alerts. Granted, this means that police will choose arbitrarily among individuals who are equally likely to have committed a crime, but arbitrary selection is more legitimate, less biased, and less prone to manipulation than the organic selection that police use today.

These changes offer some hope of redistributing the costs of fruitless searches from the poor and minority communities that frequently come into contact with police to the wealthier, whiter communities that traditionally have lived above suspicion.²² This would be a welcome improvement.

20. See *In re* Government’s Ex Parte Submission of Reauthorization Certification and Related Procedures, at 75–78 (Foreign Intelligence Surveillance Ct. 2011), available at https://www.eff.org/files/filenode/fisc_opinion_-_unconstitutional_surveillance_0.pdf (declassified in 2013).

21. This lines up with the recommendations of Bernard Harcourt and Tracey Meares in their persuasive article on randomization. See Harcourt & Meares, *supra* note 19. This Article will promote randomization for many of the same reasons that they have, with an emphasis on its ability to constrain the hassle of potentially expansive law-enforcement programs.

22. Our cognitive dissonance at the thought of attributing suspicion to attending Harvard reveals the limits of traditional methods.

I. INDIVIDUALIZED SUSPICION AMISS

Before conducting a stop, search, or another invasion of a Fourth Amendment interest, the government usually must have sufficient suspicion.²³ This suspicion is measured by the chance that law enforcement will uncover evidence of a crime.

Sometimes the government will have abundant suspicion, as when a police officer observes an illegal assault rifle, or finds a DNA match, or follows an impaired driver as he pulls off the road, climbs up a tree, and yells “I’m an owl!”²⁴ But the police do not need certainty or anything close to it. They need only to satisfy the probable cause standard in the case of searches, arrests, and exigencies and to satisfy the reasonable suspicion standard in the case of brief stops and limited frisk searches.²⁵

The Supreme Court has declined to define these standards with statistical precision, but it has stated that probable cause requires a “fair probability” that evidence of a past or future crime will be uncovered.²⁶ This standard is lower than the preponderance standard, and it is roughly estimated as a 33¹/₃% chance of recovering evidence of a crime.²⁷ The reasonable suspicion standard is lower still.²⁸ Thus, police must have some chance of success before engaging in a search or a stop, but that chance need not be large.

23. *City of Indianapolis v. Edmond*, 531 U.S. 32, 37 (2000) (“The Fourth Amendment requires that searches and seizures be reasonable. A search and seizure is ordinarily unreasonable in the absence of individualized suspicion of wrongdoing.”). The courts have recognized an exception to the suspicion requirement in a line of “special needs” and “administrative search” cases. *See, e.g.*, *Mich. Dep’t of State Police v. Sitz*, 496 U.S. 444 (1990) (permitting DUI checkpoints); *Nat’l Treasury Emps. Union v. Von Raab*, 489 U.S. 656 (1989) (allowing suspicionless drug testing of U.S. Customs employees); *Colonnade Catering Corp. v. United States*, 397 U.S. 72 (1970) (permitting inspections of liquor stores); *United States v. Edwards*, 498 F.2d 496 (2d Cir. 1974) (permitting airport security screens); *United States v. Davis*, 482 F.2d 893 (9th Cir. 1973) (same), *overruled on other grounds by United States v. Aukai*, 497 F.3d 955 (9th Cir. 2007).

24. Sebastian Murdock, *‘I’m an Owl’ Says Drunk Driver Troy Prockett Hiding in Tree: Cops*, HUFFINGTON POST, http://www.huffingtonpost.com/2014/01/10/troy-prockett-drunk-owl_n_4577317.html (last updated Jan. 23, 2014, 6:56 PM).

25. *Terry v. Ohio*, 392 U.S. 1, 10 (1968); *Carroll v. United States*, 267 U.S. 132, 149 (1925).

26. *Illinois v. Gates*, 462 U.S. 213, 238 (1983).

27. *Gerstein v. Pugh*, 420 U.S. 103, 121 (1975) (explaining that probable cause does not require a preponderance of evidence); Edward K. Cheng, *Comment on Dawid, Faigman, and Fienberg* (2014), 43 SOC. METHODS & RES. 396, 397 (2014); Christopher Slobogin, *The World Without a Fourth Amendment*, 39 UCLA L. REV. 1, 40 (1991); Andrew E. Taslitz, *Foreword, LAW & CONTEMP. PROBS.*, Summer 2010, at i–ii. *But see* Jeffrey J. Rachlinski et al., *Probable Cause, Probability, and Hindsight*, 8 J. EMPIRICAL LEGAL STUD. (SPECIAL ISSUE) 72, 86–88 (2011) (showing that judges seem to work with a probability closer to 50% to determine probable cause).

28. *Terry*, 392 U.S. at 20–22; Craig S. Lerner, *Reasonable Suspicion and Mere Hunches*, 59 VAND. L. REV. 407, 460 (2006).

But courts frequently make a further, independent inquiry. In addition to meeting the threshold probability for suspicion, the “belief of guilt must be *particularized* with respect to the person to be searched or seized.”²⁹ Thus, it is not enough for law enforcement to have adequate suspicion. Police officers must have *particularized*, or, as it is often called, *individualized* suspicion. While courts have applied the individualization rhetoric inconsistently,³⁰ the requirement is sacrosanct for most criminal-procedure scholars.³¹ And yet its purpose remains elusive.

This Part will identify and assess four theories that have attempted to describe and justify the individualization requirement. In order, they are the snowflake theory (each case is unique and cannot be based on generalizations); the felt belief theory (to promote institutional legitimacy, law enforcement and judges actually need to harbor a belief of guilt); the suspicious conduct theory (suspicion should be based on the suspect’s actual behavior); and the crime-out theory (suspicion should build from the crime toward a suspect rather than the other way around).

All of these explanations for the meaning and purpose of individualization make wishful distinctions. They fail to describe actual jurisprudential outcomes and, in any event, are normatively undesirable.

A. *The Snowflake Theory*

The most natural reading of “individualization” draws a distinction between generalizations and case-specific facts.³² It demands holistic reasoning—rather than the application of probabilities—so that individuals are judged for who they really are.³³

29. *Maryland v. Pringle*, 540 U.S. 366, 371 (2003) (emphasis added); *see also* *United States v. Cortez*, 449 U.S. 411, 418 (1981) (describing two separate elements for “articulable suspicion,” the first measuring suspicion and the second assessing particularization).

30. *See* *United States v. Montoya de Hernandez*, 473 U.S. 531, 541 (1985) (conflating particularization and suspicion); *Harris v. State*, 806 A.2d 119, 127–28 (Del. 2002) (same).

31. Scholars tend to jockey for position as the underdog. I acknowledge that is precisely what I am doing here. Harcourt, a strong critic of actuarial policing, has insisted that *my* position (in favor of actuarial methods) dominates the reality, if not the theory, of criminal procedure. HARCOURT, *supra* note 5, at 19–21. But Schauer is the lone scholar whom Harcourt identifies as a champion of actuarial justice (aside from economic defenses of racial profiling—references that are hardly likely to curry favor among legal academics). *Id.* As this Part will show, the long list of scholars writing in favor of particularization includes Laurence Tribe, Charles Nesson, Edward Cheng, Andrew Taslitz, and, of course, Harcourt himself. *But see* Harcourt & Meares, *supra* note 19, at 848 (embracing probabilistic *ex ante* estimates of suspicion).

32. Taslitz, *supra* note 18 (discussing how probable cause must be measured by the facts of a particular case); Barbara D. Underwood, *Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgment*, 88 YALE L.J. 1408, 1425–27 (1979).

33. Thanks to Derek Bambauer for summarizing uniqueness arguments as expressing the unfounded belief that each case is “its own special snowflake.”

Holistic reasoning finds support in a range of contexts from affirmative action to consumer protection.³⁴ Nesson famously argued that the reasonable doubt standard of proof required for criminal prosecutions “does not lend itself to being expressed in correlative probabilistic terms and, indeed, operates in an environment judicially structured to submerge probabilistic quantification in the factual complexity and uniqueness of specific cases.”³⁵ Holistic reasoning also incorporates the preference for human intuition celebrated in the *Trial by Mathematics* literature, in which Tribe and others argue that human intuition has its own irreplaceable wisdom that scientific predictions cannot match.³⁶

In the context of criminal investigation, Taslitz has been the most vocal advocate for case-by-case reasoning. He argues that individualization has an important function when criminal suspects present “extraordinary combinations of behaviors and traits” that generalities cannot adequately capture.³⁷

Neither Taslitz nor Nesson harbors any illusions that holistic determination is error free. Rather, they argue that justice requires considering combinations of facts unique to the individual beyond what systems of statistical inference can offer. Each case deserves its own evaluation.

The Supreme Court shares some responsibility for this romance with unique cases, as its criminal-procedure jurisprudence emphasizes the importance of case-by-case, fact-specific analysis. Although the Court has approved the use of some formulaic profiles,³⁸ more often it has insisted that police couch their justifications in ornate descriptions to show that each suspect was chosen for unique and special reasons.

The reasoning of *Richards v. Wisconsin*³⁹ is typical. In that case, the Court had to decide whether a police officer could execute an arrest warrant without first knocking on the door and announcing himself if the arrest concerned a drug-related offense. The Fourth Amendment usually requires

34. Ronald J. Allen, *Factual Ambiguity and a Theory of Evidence*, 88 Nw. U. L. REV. 604 (1994). For example, Edith Ramirez, chairwoman of the Federal Trade Commission, has warned that increased use of consumer data to make lending or other decisions is unfortunate because companies will “make determinations about individuals, not based on concrete facts, but on inferences or correlations that may be unwarranted.” Edith Ramirez, Chairwoman, Fed. Trade Comm’n, Keynote Address at the Technology Policy Institute Aspen Forum: The Privacy Challenges of Big Data: A View from the Lifeguard’s Chair (Aug. 19, 2013). For a critique of the holistic reasoning used in affirmative action cases, see Ian Ayres & Sydney Foster, *Don’t Tell, Don’t Ask: Narrow Tailoring After Grutter and Gratz*, 85 TEX. L. REV. 517 (2007).

35. Nesson, *supra* note 16, at 1191.

36. Tribe, *supra* note 5; Ronald J. Allen, *A Reconceptualization of Civil Trials*, 66 B.U. L. REV. 401 (1986); Kerr, *supra* note 2, at 139; Richard O. Lempert, *Modeling Relevance*, 75 MICH. L. REV. 1021 (1977); Peter Tillers, *Trial by Mathematics—Reconsidered*, 10 L. PROBABILITY & RISK 167 (2011).

37. Taslitz, *supra* note 18, at 158.

38. *United States v. Sokolow*, 490 U.S. 1, 10 (1989); see also Tung Yin, *The Probative Values and Pitfalls of Drug Courier Profiles as Probabilistic Evidence*, 5 TEX. F. ON C.L. & C.R. 141 (2000).

39. 520 U.S. 385 (1987).

that police announce themselves when executing an arrest,⁴⁰ but if the police have reasonable suspicion that the defendant will destroy evidence or attack the arresting officers, they may forgo the knock-and-announce procedure.⁴¹ The government's theory was that the disposable nature of drugs made all drug-related arrests good candidates for a per-se knock-and-announce exception.⁴² Justice Stevens's unanimous opinion explained that a categorical rule allowing drug arrests to proceed without a knock and announce would overgeneralize since it would encompass many situations that pose no risk of evidence destruction or violence.⁴³ By contrast, if the police know that the arrestee is aware of their presence—as was the case for Richards—the suspicion that evidence may be destroyed *would be* tailored to the situation, and police would be free to overtake the premises without knocking.⁴⁴

Justice Stevens started his opinion by rejecting generalizations, but by the end, he did nothing of the sort. As he applied the law to the facts, he substituted one rule (drug crimes = no-knock exigency) for another (drug crimes + awareness = no-knock exigency).

Put another way, suppose that a drug offender is very likely to dispose of his drugs when the police knock and announce their presence. A rule allowing the police to enter without first knocking is no more mechanical than Justice Stevens's rule permitting police to enter when one more fact (awareness) is added. Justice Stevens's rule likely increases the probability that a defendant will attempt to destroy drugs, but this advances the *suspicion* prong of individualized suspicion, not the *individualization* prong. As a matter of individualization, Justice Stevens's rule relies just as heavily on generalizations as the rule he rejects.

Case-by-case reasoning cannot and does not eliminate generalizations. Instead, it forces the generalizations to operate in informal and ad hoc ways. An inevitable result of case-by-case thinking is the creation of squishy factors like “furtive movements” that give the government's cases a gloss of particularity. These factors are based on in-the-moment, holistic impressions that can absorb any number of details that the police observe.⁴⁵ The obligation to generate a narrative distracts from much more important questions, such as whether the government is doing a good job choosing its targets.

Cases can be unique in the sense that they involve one-of-a-kind combinations of factors, but the *reasoning* of a case cannot be unique. Prediction

40. *Wilson v. Arkansas*, 514 U.S. 927 (1995).

41. *Richards*, 520 U.S. at 394.

42. *See id.* at 390.

43. *Id.* at 393.

44. *Id.* at 396. Violation of the knock-and-announce rule no longer comes with a suppression remedy. *Hudson v. Michigan*, 547 U.S. 586 (2006). But the change in remedy does not affect this Article's exploration of the Fourth Amendment's bounds.

45. *E.g.*, *Michigan v. Long*, 463 U.S. 1032, 1050 (1983); *United States v. Glover*, 851 A.2d 473, 477 (D.C. 2004); *People v. Woods*, 475 N.E.2d 442, 442 (N.Y. 1984).

requires generalization. Whatever is truly unique about a case cannot support an educated guess about its outcome unless it is analogized to some other generalization. The generalizations can be more finely grained by adding variables,⁴⁶ but the nature of the prediction does not change. Schauer has it right: “[O]nce we understand that most of the ordinary differences between general and particular decisionmaking are differences of degree and not differences in kind, we become properly skeptical of a widespread but mistaken view that the particular has some sort of natural epistemological or moral primacy over the general.”⁴⁷

It might seem quite unnatural and even immoral to give up on the promise of unique treatment, but one need only think of *McCleskey v. Kemp*⁴⁸ to see the tragic legacy of “case-by-case” thinking.⁴⁹ In *McCleskey*, an African-American man convicted of murder challenged his death sentence using statistical evidence of racial bias.⁵⁰ *McCleskey*’s challenge relied on the Baldus study, an analysis of over two thousand death-eligible murder cases tried in Georgia during the 1970s.⁵¹ The study showed that the races of the victims and murderers were strong determinants of capital sentences even after controlling for dozens of other explanatory variables (such as the method of killing, the victim’s experience before death, the defendant’s prior record, and the number of victims).⁵² The Supreme Court accepted the validity of the statistical study, but it nevertheless rejected *McCleskey*’s challenge because each individual death-row defendant had his own chance, in front of his own jury, to show why his case was unsuitable for capital punishment based “on the particularized nature of the crime and the particularized characteristics of the individual defendant.”⁵³ This move allowed the Court to sweep glaring patterns of bias under the rug in order to preserve the illusion of case-by-case determination.

46. Predictive models can also add the interaction between two variables. For example, the combination of traveling for a very short amount of time *and* traveling without a suit might increase the chance of crime more than either of the two variables alone since business travelers frequently travel for short amounts of time but will also usually have a suit.

47. SCHAUER, *supra* note 15, at 69, 106 (“[I]ndividualized analysis is simply an aggregate of stereotypes”); accord DAVID L. FAIGMAN, CONSTITUTIONAL FICTIONS: A UNIFIED THEORY OF CONSTITUTIONAL FACTS 69 (2008) (“[W]here error may be recognizable in a class of cases, it may be practically unknowable in particular cases.”).

48. 481 U.S. 279 (1987).

49. Sherry Colb links the moral intuitions that dominate the “statistical-proof” debate to the reasoning in *McCleskey*, too, and she also connects these intuitions to the reasoning that led to the market share liability doctrine in the civil DES cases. But Colb is more cautious than I am about insisting that we change our moral barometers. Colb, *supra* note 3, at 79–82.

50. *McCleskey*’s writ challenged his sentence on the grounds of the Eighth and Fourteenth Amendments. *McCleskey*, 481 U.S. at 286.

51. The study was conducted by David C. Baldus, Charles Pulaski, and George Woodworth. *Id.*

52. See David C. Baldus et al., *Comparative Review of Death Sentences: An Empirical Study of the Georgia Experience*, 74 J. CRIM. L. & CRIMINOLOGY 661, 684–85 (1983).

53. *McCleskey*, 481 U.S. at 308 (quoting *Gregg v. Georgia*, 428 U.S. 153, 206 (1976)) (internal quotation marks omitted).

McCleskey embedded a misunderstanding of epic proportions into equal-protection law. The Court failed to appreciate the import of the Baldus study. Despite the trappings of a “holistic” process, *McCleskey* was judged using generalizations. But instead of being judged by the right generalizations (those related to the heinousness of his crime), he was judged by the wrong ones (his race and the race of his victim). By denying that generalizations were used at all in his conviction, the Court was able to avoid accountability. The individualization requirement has allowed the myth of unique cases to pollute the law of criminal investigations as well.

B. *The Felt Belief Theory*

Nesson’s work rejects purely probabilistic proof, at least in the case of criminal convictions, by arguing that the jury should have an “abiding conviction” that the defendant did it.⁵⁴ Jurors should harbor a complete and confident belief, even if past experience and common sense tell us that there is a nonnegligible chance they are wrong. Nesson argues that the mistakes of wholly convinced jurors, even if they are wrong more often than statistically derived determinations of guilt, serve an objective that is different from and sometimes superior to accuracy: final resolution.⁵⁵

Under this theory, judgments coming from wholly convinced judges and juries, flawed as they may be, serve the institutional interests of the courts by presenting the party—particularly the losing party—and the public with a monolithic outcome that leaves little room to doubt the factual findings.⁵⁶ This helps preserve the authoritative reputation and popular legitimacy of the judiciary⁵⁷ and potentially helps jurors and litigants feel more psychological closure with the matter.⁵⁸

Although Nesson focuses on the reasonable doubt standard required for criminal convictions, the same institutional interests are implicated with legal determinations using lower standards. Indeed, Nesson cites the famous blue bus hypothetical, based on *Smith v. Rapid Transit, Inc.*,⁵⁹ a civil action decided under the much lower preponderance standard. In *Smith*, the plaintiff suffered injuries when she was forced off the road by a bus barreling

54. Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357, 1378 (1985); Nesson, *supra* note 16, at 1192–93.

55. Nesson, *supra* note 16, at 1194.

56. The right to appeal, and the possibility of reversal even on determinations of fact, shows that there are some limits to how well Nesson’s theory describes the judiciary’s views of its own role.

57. Nesson, *supra* note 16, at 1194.

58. A related theory concerns the dehumanization of the legal process. Mechanical determinations may be an affront to the dignity each individual is owed when the government contemplates taking punitive action. Tribe, *supra* note 5, at 1375–77. For example, Tribe worries about the symbolic expressions of values that are broadcast by judicial systems that put too much emphasis on statistical models. *Id.* at 1392. See also Laurence H. Tribe, *An Ounce of Detention: Preventive Justice in the World of John Mitchell*, 56 VA. L. REV. 371, 386 (1970).

59. 58 N.E.2d 754 (Mass. 1945).

toward her at forty miles per hour.⁶⁰ To prove her case, the plaintiff offered her own testimony (which, for the purposes of this discussion, we can assume was credible) that the bus was blue—the color of the defendant’s bus fleet. Although the original case opinion did not offer a precise estimate of the chance that somebody else’s blue bus may have been operating on the street at the time of the plaintiff’s accident, the case has bred a sort of legal urban myth that the plaintiff submitted evidence showing that 80% of the blue buses running down that particular street were the defendant’s and 20% were not. Whatever the true probabilities, the action was thrown out for failure to prove the case using more than mere “mathematical chances.”⁶¹ A long line of distinguished scholars insist that this was the correct outcome.⁶² Indeed, they still defend the outcome of *Smith* even if the defendant had been responsible for 99% of the blue buses running down the street.⁶³ After all, a jury would not be able to quiet the nagging thought that it may have been that other infrequent bus.⁶⁴

The individualization requirement for criminal investigations might promote the same interests in finality and responsibility. Even if human decisionmaking is flawed, perhaps a court should still require that an officer feel certain that a suspect is guilty before stopping and searching him, or at least be able to gut check his decision rather than abdicate his judgment to a statistical process.

Sure enough, *Carroll v. United States*,⁶⁵ an old case frequently quoted for its definition of probable cause, could support a distinction between belief and unenthusiastic assent to mathematical inference. In that case, the Supreme Court explained that an officer must have sufficient information “to warrant a man of prudence and caution in *believing* that the offense has been committed.”⁶⁶ Many cases have made clear that an officer’s subjective belief is not enough without additional objective evidence that would lead a reasonable officer to conclude that crime might be afoot.⁶⁷ But these cases do not necessarily deny that subjective belief is a necessary condition for individualized suspicion.

While the felt belief theory and the snowflake theory both lead to rejections of purely statistical evidence, they do so on different grounds. The

60. *Smith*, 58 N.E.2d at 754–55.

61. *Id.* at 755.

62. Edward K. Cheng, *Reconceptualizing the Burden of Proof*, 122 YALE L.J. 1254, 1269–71 (2013); see also Nesson, *supra* note 16, at 1192–93; Tribe, *supra* note 5, at 1376.

63. Cheng, *supra* note 62, at 1270 (using a gatecrashers hypothetical).

64. The blue bus paradox and others like it have also inspired debate about whether arresting two or three people for a crime that the police know only one has committed is constitutionally permissible. The definiteness of an innocent person’s arrest triggers the “identifiable victim effect,” although, as Colb points out, the same doubts should trouble any system engaged in repeated actions. Colb, *supra* note 3, at 76–78, 101.

65. 267 U.S. 132 (1925).

66. *Carroll*, 267 U.S. at 161 (emphasis added) (quoting *Stacey v. Emery*, 97 U.S. 642, 645 (1878)) (internal quotation marks omitted).

67. See, e.g., *United States v. Foreman*, 369 F.3d 776, 781 (4th Cir. 2004).

snowflake theory is built on a fallacy; it assumes, wrongly, that distinctions can be drawn between cases without resorting to a combination of generalizations. Thus, the snowflake theory's flaws are insurmountable. Case-by-case reasoning simply doesn't exist. The felt belief theory, on the other hand, does articulate a standard that can be implemented. Individualization would be satisfied if a human can trust the evidence using his own senses and experiences.

Nevertheless, the felt belief theory has three significant and intertwined flaws. First, it is not possible to separate evidence that can lead to a felt belief from evidence that cannot. For example, there is no reason to assume that witness testimony about the operator of the blue bus can move a juror to complete belief but methodologically sound calculations of the chance that the defendant operated the blue bus cannot. The distinction imports somebody's strong preferences for certain types of evidence, allowing some types of evidence to persuade the juror, judge, or officer to a felt belief while denying other types of evidence the chance to do so.

Second, the theory gets the moral imperative backwards. Nobody—whether juror, magistrate, or police investigator—should maintain his beliefs with full conviction. Indeed, procedural rights and judicial appeals are designed to avoid any delusions of perfection.⁶⁸ Since most evidence should leave some doubt in the minds of the decisionmakers, evidence cannot be categorized along the lines that Nesson and others suggest.

Finally, if the felt belief theory rejects all mechanical forms of evidence, the theory runs into paradoxes when such evidence is extremely accurate. Consider the cold hit DNA case, which meets its burden of proof entirely through references to probability. Suppose the government presents evidence that a defendant's DNA matches a sample from a rape kit across all thirteen tested genetic biomarkers and that this unique combination of markers is likely to occur only one out of 56 billion times. The chance that somebody else in the world (let alone in the country) would share this much of the defendant's genetic code is negligibly low.⁶⁹

Yet the cold hit DNA case differs from the blue bus hypothetical only in degree. If the plaintiff in *Smith* could prove that the only blue bus, aside from the defendant's, drove down Main Street only once every thousand years, her case would rely just as much on probabilities. If pure mathematical proof—divorced from the sorts of evidence that juries can see, hear, and perceive themselves—cannot inspire a felt belief, then cold hit DNA cases

68. New Jersey's experience curbing the influence of eyewitness testimony provides a good example of the courts' using both appeals and procedural rights in order to correct errors at trial. After digesting the wide-ranging experimental and observational evidence on the inaccuracy of eyewitness testimony, the New Jersey Supreme Court introduced new limitations on the admission of such testimony at trial. *State v. Henderson*, 27 A.3d 872 (N.J. 2011).

69. But not as low as one might think. For a discussion of proper pigeonholing methods, see Philippe Golle, *Revisiting the Uniqueness of Simple Demographics in the US Population*, PROC. FIFTH ASS'N COMPUTING MACHINERY WORKSHOP ON PRIVACY ELECTRONIC SOC'Y 77 (2006), available at <http://dl.acm.org/citation.cfm?id=1179615>.

would have to fail. But this outcome is intolerable as a practical matter and perverse as a normative matter.

If DNA cases are legitimate, the blue bus hypothetical must be reconsidered, too. There is no reason to allow a plaintiff to win a suit against a bus company that runs only 49% of the city's blue buses when she supplements her evidence with a barely credible eyewitness while rejecting the suit of a plaintiff who sues a bus company that runs 99% of Main Street's blue buses.⁷⁰ Instead of fretting over the probabilistic nature of evidence, courts should strive to make peace with the doubt that does, and should, accompany all factual determinations in law. After all, the study of quantum physics is rapidly moving toward an understanding of the universe that is driven entirely by probability theory.⁷¹ If the laws of nature are probabilistic, surely the laws of law will have to be as well.

Courts have proven to be much less perplexed by probabilistic evidence than the academic debate would suggest. The Arizona Supreme Court had to decide a case with all the zany qualities of a law school hypothetical: a student at a boarding school contracted a severe case of salmonella from one of between 100 and 120 meals consumed on campus.⁷² Every meal but one was prepared by an independent contractor, but that one meal was prepared by the school, and none of the evidence presented could help establish or rule out the school-made meal as the cause of the plaintiff's illness.⁷³ Despite the similarities to the famous blue bus hypothetical, the Arizona Supreme Court had no difficulty granting the school's motion for summary judgment and allowing the case against the contractor to proceed.⁷⁴

In the context of criminal procedure, courts have already accepted investigatory tools that substitute for human judgment. Cold hit DNA matching is one unusually accurate statistical tool, but there are other error-prone tools that have been deemed sufficient for Fourth Amendment standards. For example, the Supreme Court has approved the use of alerts from narcotics-sniffing dogs to establish probable cause to search cars and luggage.⁷⁵ Justice Souter strongly disagreed with these decisions partly on the grounds

70. This is the perspective that Schauer ascribes to statisticians in his assessment of the gatecrashers and blue bus problems. See SCHAUER, *supra* note 15, at 88–89; see also F.E. Guerra-Pujol, *Visualizing Probabilistic Proof: The Case for Bayes*, 7 WASH. U. JURISPRUDENCE REV. (forthcoming 2015) (manuscript at 7), available at <http://ssrn.com/abstract=2271870> (pointing out that the *Smith* plaintiff could have testified that she had seen the defendant's logo on the bus but that her identification testimony might well have been wrong).

71. See *Does Probability Come from Quantum Physics?*, U.C. DAVIS NEWS & INFO. (Feb. 5, 2013), http://news.ucdavis.edu/search/printable_news.lasso?id=10491&table=news.

72. *Orme Sch. v. Reeves*, 802 P.2d 1000, 1002 (Ariz. 1990).

73. *Id.* at 1002–03.

74. *Id.* at 1009–10.

75. *Illinois v. Caballes*, 543 U.S. 405, 409 (2005); *United States v. Place*, 462 U.S. 696, 707 (1983). The Supreme Court recently decided that bringing a dog up to the front door of a home is unconstitutional, but its decision involved a different issue (whether the physical trespass of the police dog was an impermissible search). *Florida v. Jardines*, 133 S. Ct. 1409 (2013).

that narcotics dogs make errors.⁷⁶ He rightly criticized the Court's early opinions on the topic for either failing to address canine error or, worse yet, implying that the dogs are infallible.⁷⁷ Subsequent cases show that the Court now acknowledges that the dogs make mistakes, but it remains willing to permit their use anyway.⁷⁸

This is sensible. Although narcotics dogs do create false alerts,⁷⁹ their hit rates far outperform those of highway patrolmen in predicting which cars contain drugs.⁸⁰ They even outperform the success rates of warrant-based home searches—the gold standard for criminal investigation.⁸¹

For some scholars, however, the dog-sniff cases showcase everything that is wrong with criminal procedure today.⁸² Most scholars object for the reasons just discussed—that drug-detecting dogs are prone to error.⁸³ Others object that dogs can be used in discriminatory ways, which is true enough depending on how or, more importantly, where they are deployed.⁸⁴ But this remains a concern for every type of investigation, including those based on human observations and judgments. Moreover, the random error of the dog nose, when it does misfire, is usually free from the prejudices that can influence the (nonrandom) error of human police.

76. *Caballes*, 543 U.S. at 411–12 (Souter, J., dissenting). Justice Ginsburg also noted, in a remark unrelated to the discussion here, that the presence of a police dog causes intimidation and embarrassment, regardless of whether the dog alerts. *Id.* at 421 (Ginsburg, J., dissenting).

77. *See Caballes*, 543 U.S. at 409; *Place*, 462 U.S. at 707.

78. *See Florida v. Harris*, 133 S. Ct. 1050, 1056–58 (2013).

79. *United States v. Donnelly*, 475 F.3d 946, 954 (8th Cir. 2007) (54% hit rate); *United States v. Koon Chung Wu*, 217 F. App'x 240, 246 (4th Cir. 2007) (60% hit rate); *United States v. Scarborough*, 128 F.3d 1373, 1378 (10th Cir. 1997) (92% hit rate).

80. *See Harcourt & Meares*, *supra* note 19, at 849 (Maryland's state patrol hit rate ranged between 32% and 34%); *see also* Max Minzner, *Putting Probability Back into Probable Cause*, 87 TEX. L. REV. 913, 925 (2009) (citing, based on highway-patrol data, hit rates ranging from 35.1% to 52.5% for probable-cause car searches).

81. Jane Bambauer, *Defending the Dog*, 91 OR. L. REV. 1203, 1206 (2013) (comparing dog hit rates to home warrant hit rates). *But see* Minzner, *supra* note 80, at 923 n.38 (citing success rates as high as 93% for warrant-based searches).

82. *See, e.g.*, Lewis R. Katz & Aaron P. Golembiewski, *Curbing the Dog: Extending the Protection of the Fourth Amendment to Police Drug Dogs*, 85 NEB. L. REV. 735 (2007); Richard E. Myers II, *Detector Dogs and Probable Cause*, 14 GEO. MASON L. REV. 1, 22 (2006) (discussing false-positive errors); Andrew E. Taslitz, *Does the Cold Nose Know? The Unscientific Myth of the Dog Scent Lineup*, 42 HASTINGS L.J. 15 (1990) (criticizing dog-sniff tracking as evidence at trial). *But see* James B. Johnston, *Drugs, Dogs, and the Fourth Amendment: An Analysis of Justice Stevens' Opinion in Illinois v. Caballes*, 24 QUINNIPIAC L. REV. 659 (2006); Kit Kinports, *The Dog Days of Fourth Amendment Jurisprudence*, 108 NW. U. L. REV. COLLOQUY 64 (2013), <http://www.law.northwestern.edu/lawreview/colloquy/2013/7/LRColl2013n7Kinports.pdf>.

83. *See supra* note 82.

84. *See* Dan Hinkel & Joe Mahr, *Drug Dogs Often Wrong*, CHI. TRIB., Jan. 6, 2011, at 1, available at http://articles.chicagotribune.com/2011-01-06/news/ct-met-canine-officers-20110105_1_drug-sniffing-dogs-alex-rothacker-drug-dog (uncovering evidence of biased handling—the rate for Hispanic drivers was much lower).

Still, some may be troubled that searches based on dog alerts are not grounded in any evidence of suspicious behavior. This objection will be taken up next.

C. Suspicious Conduct Theories

Taslitz proposes an appealing theory that individualization requires an assessment based on the suspect's own conduct or behavior.⁸⁵ This theory eliminates the injustice of treating individuals differently due to attributes that they cannot manage or alter themselves.

Racial profiling is the motivating example for Taslitz. Although crime rates do vary by race and gender,⁸⁶ the variance by race is greatly diminished when socioeconomic conditions are controlled.⁸⁷ And since law enforcement has historically given too much weight to race and gender,⁸⁸ a prophylactic prohibition on using those factors in criminality profiles helpfully pushes law enforcement to use other, more predictive factors to determine reasonable suspicion and probable cause.

But the conduct theory of individualization falls apart outside the context of racial profiling. While the conduct theory matches some of the classic means of building suspicion—casing a store⁸⁹ or looking around nervously⁹⁰—it cannot explain large swaths of criminal procedure.⁹¹ Police are allowed to stop and search some people who have not exhibited any odd behaviors. Take, for example, the unlucky lot whose cars are searched in response to a false alert from a narcotics-sniffing dog. Unless they were carrying a trunkful of Snausages, these individuals' conduct played no role in the police's decision to search. Similarly, individuals matching a description from a victim or witness cannot alter their behavior to avoid suspicion. And

85. Taslitz, *supra* note 18, at 146 (“Roughly defined, individualized suspicion is the idea that the state should judge each citizen based upon his own unique actions, character, thoughts, and situation.”); *id.* at 160 (“Yet this judgment still retains a *relatively* individualized quality because it turns on observations of *this individual's* prior behavior”); see also Andrew Guthrie Ferguson, *Big Data and Predictive Reasonable Suspicion*, 163 U. PA. L. REV. (forthcoming 2015); David A. Harris, *Factors for Reasonable Suspicion: When Black and Poor Means Stopped and Frisked*, 69 IND. L.J. 659, 685–86 (1994); Underwood, *supra* note 32, at 1447. The Supreme Court has suggested that individualization depends on the suspect's conduct as well. See *Reid v. Georgia*, 448 U.S. 438, 441 (1980) (per curiam).

86. See, e.g., ALEXIA COOPER & ERICA L. SMITH, U.S. DEP'T OF JUSTICE, NCJ 236018, HOMICIDE TRENDS IN THE UNITED STATES, 1980–2008, at 14 (2011), available at <http://www.bjs.gov/content/pub/pdf/htus8008.pdf>.

87. See Julie A. Phillips, *White, Black, and Latino Homicide Rates: Why the Difference?*, 49 SOC. PROBS. 349 (2002); Robert J. Sampson et al., *Social Anatomy of Racial and Ethnic Disparities in Violence*, 95 AM. J. PUB. HEALTH 224 (2005).

88. See Section II.D.4 for a discussion of racial bias in policing.

89. *Terry v. Ohio*, 392 U.S. 1, 6 (1968).

90. *United States v. Ocampo*, 650 F.2d 421, 424 (2d Cir. 1981).

91. Colb points out that suspicion is controlled by what the government knows and that it remains extrinsic to an individual's conduct and culpability. Sherry F. Colb, *Innocence, Privacy, and Targeting in Fourth Amendment Jurisprudence*, 96 COLUM. L. REV. 1456, 1459 (1996).

the subjects of tips form another class of searched individuals whose suspicion was not necessarily based on conduct.⁹² Like the targets of a mistaken dog nose, these subjects will be unable to alter their behavior to avoid the stop.

Even if it were possible to reconfigure individualized suspicion so that it conformed to a conduct-driven rule, it would be unwise to do so. First, the line between conduct and attribute is difficult to manage. For example, is a person's location a form of conduct? If so, this detracts from the rule's appeal. Surely many people who live in or travel through "high-crime" neighborhoods would avoid those areas if they could. And what about the Harvard students living in a dormitory? Attending college is a choice. Would the conduct rule allow the officer in the Article's introductory hypothetical to obtain a warrant to search all Harvard dormitories? The line between conduct and nonconduct is inadministrable.

Moreover, even if the rule were administrable, requiring suspicion to arise from observed conduct alone would displace some of the more reliable gauges for suspicion (DNA matching, dogs, tips from reliable informants) and would consequently put more pressure on some of the less reliable profiles for crime (nervousness, flight, style of dress). This shift in focus could exacerbate the already lopsided distribution of law-enforcement costs on the poor.

D. *Crime-Out Investigations*

A final theory for individualization rejects law-enforcement investigations that go out in search of suspicious people in favor of investigations stemming from an already committed crime. Harcourt's book *Against Prediction* warns against actuarial law-enforcement methods that attempt to sort out suspicious people.⁹³ He objects in particular to predictions of criminal personality.⁹⁴ Profiles developed to identify a segment of the population that is more likely to commit crime bear an unnerving resemblance to past eugenics movements.⁹⁵ Probabilistic inferences that stem from the scene of a

92. See, e.g., *People v. Navarette*, No. A132353, 2012 WL 4842651 (Cal. Ct. App. Oct. 12, 2012), *aff'd*, *Navarette v. California*, 134 S. Ct. 1683 (2014); *People v. Johnson*, 235 Cal. Rptr. 62 (Ct. App. 1987).

93. HARCOURT, *supra* note 5. Despite his thoroughness in every other aspect, Harcourt spends little time defining these actuarial methods and explaining how they differ from other methods. First, Harcourt defines by example. He differentiates actuarial profiles from the process of identifying a suspect based on an eyewitness description. *Id.* at 18. Similarly, he notes that DNA evidence used to match a suspect to a crime does not "concern itself with the offending rates of any particular group." *Id.* The examples that Harcourt later uses to distinguish actuarial policing (a homicide investigation and an eyewitness account) are also crime out. *Id.* at 104.

94. See, e.g., *id.* at 174–80 (discussing Burgess and other pioneers of predictive modeling for penal decisions).

95. Richard Hil & Barbara Hocking, *Genes for Crime: Do They Fit?*, in *RECODING NATURE* 163, 164–66 (Richard Hindmarsh & Geoffrey Lawrence eds., 2004) (describing the work of Charles Davenport, Francis Galton, and Karl Pearson in attempting to define and weed out a

particular crime, by contrast, can avoid these uncomfortable similarities. As long as police departments work outward from a particular crime rather than surveying the public and looking for suspicious types, we can be confident that suspicion is growing from the facts of a particular, individual case.

In other words, we might distinguish “crime-out” investigations from “suspect-in” investigations. But while this distinction is useful,⁹⁶ it cannot do all the constitutional work here. The scope of suspect-in investigations includes many sound and sanctioned practices. To give just one example, the conduct that caught Officer McFadden’s eye in *Terry v. Ohio* led to a suspect-in investigation. Terry was casing a men’s clothing store, pacing in front of it and conferring with his conspirator. McFadden’s observations of this behavior aroused his suspicion even though he was not following the leads from an already committed crime.⁹⁷

McFadden’s decision to stop and question Terry was sufficiently particularized for the Court.⁹⁸ What McFadden observed is a useful and often deployed actuarial profile: people who pace in front of a building, paying attention to its details, might be plotting a robbery. Of course, the profile is not perfect—chronic pacers, window shoppers, and architectural enthusiasts may be lumped in with the robbers. But courts have found that this type of evidence is sufficiently individualized.

Moreover, the *Terry* case points to a larger problem with Harcourt’s definition: adhering to the crime-out mode of individualization would wipe out most of the opportunities for police to thwart attempted criminal activity.⁹⁹ It is also of limited value for underreported crimes. Crime-out investigations only begin once a crime has been committed *and* the authorities know that it has been committed.¹⁰⁰ The scope is much too narrow for a constitutional limit on investigation.

genetic component of criminal behavior and tracing its relationship to the eugenics movements).

96. Christopher Slobogin uses a similar “event-driven” versus “suspect-driven” dichotomy in his work. CHRISTOPHER SLOBOGIN, *PRIVACY AT RISK: THE NEW GOVERNMENT SURVEILLANCE AND THE FOURTH AMENDMENT* 183–86 (2007).

97. 392 U.S. 1, 5 (1968).

98. *Terry*, 392 U.S. at 28.

99. One could imagine some plotted and premeditated crimes that could trigger investigations before the crime is completed—for example, if a potential victim reports that somebody is following him in a suspicious way.

100. Harcourt goes on to explain that actuarial methods are “criminal justice determinations that do not rest simply on probabilities but on statistical correlations between group traits and group criminal offending rates.” HARCOURT, *supra* note 5, at 18. The scope of this definition depends on how “group traits” are defined, however. If groups can be determined by any shared quality, then the subjects of tips or chronic pacers (who appear to be casing) would constitute an actuarial profile. Even the subset of people who happen to match eyewitness descriptions of perpetrators shares a trait that could be used to define a group, although members of that subset look very little like each other. It’s true that these are not demographic or immutable characteristics that rattle scholars the most, but neither are the shared traits in drug-courier profiles (such as use of air fresheners or the presence of fast-food wrappers). If *any* set of generalizations is actuarial, this leaves no room for the criminal justice determinations that *aren’t* actuarial. See *supra* Section I.A.

E. *Ending the Status Quo*

Existing theories of individualization fail to provide a satisfying account of its purpose. In practice, the doctrine has served only to ossify the familiar methods and tools, thereby preserving the status quo for no principled reason. Supreme Court precedents place great emphasis on the narrative of the experienced policeman,¹⁰¹ and as a result the most respected sources of *individualized* suspicion are the least reliable and the least fair.

An officer's testimony about what he or she observed is prone to misjudgment or even outright deceit (a practice common enough to have inspired a cute nickname—"testilying"¹⁰²). And apart from perjury, judges expect officers to use squishy, subjective factors like furtive movements,¹⁰³ suspicious bulges,¹⁰⁴ the officer's training and experience,¹⁰⁵ "surveillance-conscious behavior," and "high-crime areas" to build up suspicion in a particularized way, despite ample evidence that these factors are poor predictors of criminal activity.¹⁰⁶ While such evidence seems to satisfy courts' desire for individualization, the evidence does not reliably increase the chance that the target is engaged in crime.¹⁰⁷ Thus, current attempts at individualization do worse than nothing. They corrode the suspicion standards by allowing an officer's unscientific opinions to guide predictions of crime.

101. *Illinois v. Wardlow*, 528 U.S. 119, 124–25 (2000) (relying on officer testimony about a suspect's nervous, evasive behavior to establish reasonable suspicion); *Terry*, 392 U.S. at 22–23 (analyzing Officer McFadden's observations and experience).

102. See, e.g., ALAN M. DERSHOWITZ, *THE ABUSE EXCUSE* 233 (1994); David N. Dorfman, *Proving the Lie: Litigating Police Credibility*, 26 AM. J. CRIM. L. 455, 480 (1999); Christopher Slobogin, *Testilying: Police Perjury and What to Do About It*, 67 U. COLO. L. REV. 1037 (1996).

103. *People v. Woods*, 475 N.E.2d 442, 442 (N.Y. 1984).

104. *People v. De Bour*, 352 N.E.2d 562, 570 (N.Y. 1976); *People v. Hudson*, 527 N.Y.S.2d 919, 919 (App. Div. 1988).

105. *Terry*, 392 U.S. at 27; *United States v. Brown*, 159 F.3d 147, 149–50 (3d Cir. 1998); *Harris v. State*, 806 A.2d 119, 121 (Del. 2002); *State v. Lafferty*, 967 P.2d 363, 366 (Mont. 1998), *abrogated on other grounds by State v. Flynn*, 251 P.3d 143 (Mont. 2011).

106. "High-crime area" was used as a justification in over 55% of the stops performed in New York between 2004 and 2009. Jeffrey Fagan compared the use of "high-crime area" as a justification across precincts to see if that justification correlated with actual crime data. It did not. Even in the precincts with the lowest crime rates, "high-crime area" was used as a justification nearly 55% of the time. Report of Jeffrey Fagan, *supra* note 7, at 53–54; Andrew Guthrie Ferguson & Damien Bernache, *The "High-Crime Area" Question: Requiring Verifiable and Quantifiable Evidence for Fourth Amendment Reasonable Suspicion Analysis*, 57 AM. U. L. REV. 1587, 1594, 1605–22 (2008) ("[T]here is no definitional clarity to the 'high-crime area' term now regularly used by the courts post-*Wardlow*.").

107. See Ferguson & Bernache, *supra* note 106, at 1605–22. Judge Kozinski on the Ninth Circuit Court of Appeals criticized his colleagues' acceptance of testimony about a "high-crime area" without any hard evidence. *United States v. Montero-Camargo*, 208 F.3d 1122, 1143 (9th Cir. 2000) (Kozinski, J., concurring) ("Are such estimates sufficiently precise to tell us anything useful about the area? I wouldn't have thought so, although I could be persuaded otherwise. But my colleagues don't even pause to ask the questions. To them, it's a high crime area, because the officers say it's a high crime area.").

It gets worse. Traditional routes to individualization distribute their intrusions in severely regressive ways. It's no secret that discretion- and observation-driven policing lead to more searches of poor and minority subjects.¹⁰⁸ This is at least partially a result of where police spend their time. But the accumulation of recent Fourth Amendment rules has not helped. The upper classes can afford copious curtilage,¹⁰⁹ usually hang out in "low-crime areas,"¹¹⁰ and may wear form-fitting bulgeless clothing more often.¹¹¹ Thus, poor and minority communities serve a disproportionate share of the prison time for minor drug convictions, despite having drug-usage rates similar to those of upper-class and white communities.¹¹² In contrast, algorithms are more likely to cast their cold accusations on everybody.

The unavoidable conclusion is that individualization, when courts have insisted on it, has permitted a high degree of error to infect the criminal-investigation process. Justifications for the individualization requirement have spared the courts the embarrassment of rejecting long-trusted sources of suspicion, but so far existing theories of individualization have offered very little to justify its exalted place in constitutional law.¹¹³

II. HASSLE

The individualization half of "individualized suspicion" is in trouble. Scholars who defend it cannot articulate which types of criminal investigations should be considered unconstitutional despite their high likelihood of success. And yet the concept has irrefutable magnetism.

108. See DAVID K. SHIPLER, *THE RIGHTS OF THE PEOPLE: HOW OUR SEARCH FOR SAFETY INVADES OUR LIBERTIES* 55 (2011).

109. For example, in *Florida v. Jardines*, 133 S. Ct. 1409 (2013), the Supreme Court found that bringing a drug-sniffing dog to the door of a house constituted a search. But because the opinion dealt with physical trespass onto the curtilage, lower courts have permitted the same technique on the front doors of apartments. See, e.g., *State v. Nguyen*, 841 N.W.2d 676 (N.D. 2013).

110. See Julie Berry Cullen & Steven D. Levitt, *Crime, Urban Flight, and the Consequences for Cities*, 81 REV. ECON. & STAT. 159 (1999) (demonstrating that high-income households leave urban centers when the risk of crime is high); see also Ferguson & Bernache, *supra* note 106, at 1591–93 (criticizing the term "high-crime area" for lacking objectivity and cataloguing the literature that asserts that the factor has been used in racially and class-biased ways).

111. See, e.g., Dean A. Dabney et al., *The Impact of Implicit Stereotyping on Offender Profiling: Unexpected Results from an Observational Study of Shoplifting*, 33 CRIM. JUST. & BEHAV. 646, 669 (2006) (raising the possibility that peoples' perceptions about how "baggy" clothing is, and thus how likely it might be to conceal stolen merchandise, could vary by race).

112. CDC DRUG USAGE TABLE, *supra* note 12. But the government may use drug-offense pleas to bargain away the prosecution of more serious crimes. See K. JACK RILEY ET AL., *JUST CAUSE OR JUST BECAUSE?: PROSECUTION AND PLEA-BARGAINING RESULTING IN PRISON SENTENCES ON LOW-LEVEL DRUG CHARGES IN CALIFORNIA AND ARIZONA* (2005).

113. I am by no means the first to push back against the status quo in favor of some more overtly statistical decisionmaking. See Jonathan J. Koehler & Daniel N. Shavero, *Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods*, 75 CORNELL L. REV. 247, 250 (1990); Andrea Roth, *Safety in Numbers? Deciding when DNA Alone Is Enough to Convict*, 85 N.Y.U. L. REV. 1130 (2010).

This Part proposes a new understanding of the individualization requirement. Courts and scholars have focused exclusively on the particulars of the stopped or searched suspect. They ask, “Why *her*?” This is the wrong question. More precisely, the suspicion layer of Fourth Amendment protection already addresses this question. Why *her*? Because she lives in a Harvard dorm room, so there is a 60% chance that she has narcotics.

Instead, the individualization inquiry should ask, “Why not *everybody else*?” After all, while we do not all live in Harvard dorm rooms, we do all wind up, at some point, in circumstances in which our benign behavior may legitimately pique law-enforcement suspicion. We stand on corners, and look over our shoulders, and purchase Bob Marley CDs. The Fourth Amendment must offer some assurance that most of the time, most of us will be excluded from stops and searches even though we pass through these temporary states of heightened suspicion.

This Part identifies the unsung virtues of individualization. If “suspicion” can be summarized mathematically as a hit rate (the chance that evidence will be discovered in the course of a stop or search), individualization can be captured as a hassle rate: the chance that an innocent person will have to undergo a stop or search.

Section II.A provides a descriptive model that captures in a nutshell how well a given criminal-investigation program is working. Section II.B then establishes the connection between hassle rates and individualization. The individualization practices routinely endorsed in case law and scholarship accomplish the goal of limiting how many other people the police can practicably search. This reduces hassle on the innocent people. Moreover, although discussions of individualization have historically focused on the searched or stopped suspect, the interest in excluding others has quietly guided the reasoning. Some courts have explicitly considered hassle without quite knowing how it fits into the Fourth Amendment jurisprudence. Finally, Section II.C concludes by showing that understanding individualization in terms of hassle explains many other instincts that have circulated in the individualized-suspicion case law.

By renovating individualization consciously to minimize hassle, courts can transparently address problems that have been latent motivators in probable-cause and reasonable-suspicion cases. At the very least, renovating the concept of individualization gives it something useful to do while satisfying a demand for a type of justice that has not yet found a home in the Fourth Amendment’s protection.

A. *A Compact Model for Criminal Investigations*

A criminal-investigation program can be summarized using four statistics: the base rate, the hit rate, the miss rate, and the hassle rate.

The base rate is the rate at which crime is committed.

$$\text{Base Rate} = \frac{N}{X}$$

where:

X = total population for the relevant jurisdiction

N = number of people within the jurisdiction committing the particular crime or crimes.

The hit rate is the proportion of searched individuals who are caught with contraband or evidence of a crime.¹¹⁴

$$\text{Hit Rate} = \frac{C}{S}$$

where:

S = subset of the population searched

C = subset of the population searched *and* caught.

The miss rate reflects the proportion of criminals who are not caught.

$$\text{Miss Rate} = \frac{N-C}{N}$$

where:

N = number of people within the jurisdiction committing the particular crime or crimes

C = subset of the population searched *and* caught.

And finally, the hassle rate shows the proportion of the innocent population who are searched fruitlessly.

$$\text{Hassle Rate} = \frac{S-C}{X-N}$$

where:

S = subset of the population searched

C = subset of the population searched *and* caught

X = total population for the relevant jurisdiction

N = number of people within the jurisdiction committing the particular crime or crimes.

114. For now, to keep the discussion simple, I am ignoring the additional harm from searching the same person more than once. A better measure of hassle would find the proportion of the innocent population stopped once and then add to it a weighted proportion of the innocent population stopped twice, another weighted proportion of the innocent population stopped three times, and so forth. The weights should be nonlinear since studies have found that the frustration and distrust caused by stops and searches increase exponentially by the number of compelled stops and searches.

Each of these figures will vary by crime, program, jurisdiction, time frame, and population. And we will often lack the data needed to calculate each of the statistics with precision. This is especially true for miss rates and base rates since, by definition, criminals who are not detected by law enforcement cannot appear in police data. But taken together, rough estimates of these statistics form a thumbnail sketch of the costs of criminal-investigation processes. In fact, estimates of any three will suffice since the fourth can be derived from the others.¹¹⁵

An optimal criminal-investigation system will have low base rates, miss rates, and hassle rates, and high hit rates. Ideally these rates will not vary dramatically across race and class. But the incentives to optimize the scale and distribution of these rates come from varying sources of law.

Base rates are not in the government's direct control because they measure the criminal behavior of the population. We depend on political processes to create policies that will deter the commission of future crime. Since crime deterrence is a majoritarian interest, the incentives do not have or need constitutional reinforcement. The Fourth Amendment is also agnostic about miss rates, again leaving it to political processes to create the right incentives for detection.¹¹⁶ It is worth noting, though, that disparities in miss rates across race (and other constitutionally protected classes) could implicate the Fourth Amendment or the Equal Protection Clause if race is used to select a population for further investigation and screening.¹¹⁷

Hit rates are a different matter. The Fourth Amendment keeps a watchful eye on them through the suspicion standards (either probable cause or reasonable suspicion). Because many jurisdictions have monitoring and record-keeping requirements as a result of prior consent decrees,¹¹⁸ there are some fairly good data on the hit rates for a smattering of programs. In the late 1990s, Maryland's highway patrol successfully recovered contraband in over half of the warrantless car searches it conducted based on probable cause (as opposed to less fruitful searches conducted with the consent of the drivers).¹¹⁹ Other jurisdictions didn't perform quite as well but still cleared

115. The four statistics are composed of four variables: N, X, S, and C. Thus, three independent equations allow us to solve for the fourth.

116. Police departments are zealously focused on base rates and miss rates, at least by assumption. The Supreme Court has called detecting and preventing crime the "most basic function of any government." *Illinois v. Gates*, 462 U.S. 213, 237 (1983) (quoting *Miranda v. Arizona*, 384 U.S. 436, 539 (1966) (White, J., dissenting)). *But see* Joseph Goldstein, *Police Discretion Not to Invoke the Criminal Process: Low-Visibility Decisions in the Administration of Justice*, 69 *YALE L.J.* 543, 552–54 (1960) (arguing that, because police decisions not to enforce are not visible to the community, those decisions are not subject to the same political pressures that would result from community and administrative review of affirmative police conduct).

117. Samuel R. Gross & Katherine Y. Barnes, *Road Work: Racial Profiling and Drug Interdiction on the Highway*, 101 *MICH. L. REV.* 651, 737–38, 740 (2002).

118. *See, e.g., Consent Decree Overview*, L.A. POLICE DEPARTMENT, http://www.lapdonline.org/search_results/content_basic_view/928 (last visited Aug. 27, 2014).

119. Gross & Barnes, *supra* note 117, at 700 tbl.14.

the 33¹/₃% hurdle.¹²⁰ In contrast, a district court found that New York's stop-and-frisk program was unconstitutional partly because of its low hit rate.¹²¹ Hit rates are useful signals for the adequacy (or not) of police methods that attempt to single out suspicious behavior. But hit rates cannot alone capture the costs to society and the people in it.

The hassle rate is critical.¹²² The word "hassle" will at times understate the intrusion and disrespect that can characterize stops and searches (particularly when the use of force is involved¹²³), but the virtue of the term is that "hassle" does not exaggerate the problem. Sometimes, hassle is all that it is. A police officer may briefly detain a person to ask a few questions and, when the misunderstanding is resolved, all parties may go on their way having suffered only mild irritation. When a search is unusually intrusive, courts will require additional suspicion.¹²⁴ When a search is *really* intrusive, courts will invalidate it on substantive due process grounds.¹²⁵ But courts have not been in the habit of paying close attention to the typical stops and searches.

These ordinary stops and searches have significant costs. When the experience of a typical member of a community involves involuntary stops or searches, it is natural for the community to question whether the government has overreached its authority. And when an innocent person is stopped more than once in a short time, the effects are much more severe.

120. See, e.g., Minzner, *supra* note 80, at 931–33 (surveying Florida state police officers).

121. *Floyd v. City of New York*, 959 F. Supp. 2d 540, 586 (S.D.N.Y. 2013).

122. Colb, *supra* note 91, at 1472, 1505 (describing the Fourth Amendment as balancing the intrusions on the innocent with the need for law enforcement).

123. See generally *CTR. FOR CONSTITUTIONAL RIGHTS, STOP AND FRISK: THE HUMAN IMPACT* (2012), available at <http://stopandfrisk.org/the-human-impact-report.pdf>.

124. *Safford Unified Sch. Dist. No. 1 v. Redding*, 129 S. Ct. 2633, 2641–42 (2009); *United States v. Afanador*, 567 F.2d 1325, 1328 (5th Cir. 1978) (“[W]hat constitutes ‘reasonable suspicion’ to justify a particular search may not suffice to justify a more intrusive or demeaning search.”); *United States v. Love*, 413 F. Supp. 1122, 1127 (S.D. Tex.) (“[T]he greater the intrusion, the greater must be the reason for conducting a search that results in such invasion.”), *aff’d*, 538 F.2d 898 (5th Cir. 1976).

125. *Rochin v. California*, 342 U.S. 165, 174 (1952) (finding that using a stomach pump to retrieve swallowed drugs was unconstitutional on due process grounds); *cf. Winston v. Lee*, 470 U.S. 753, 755 (1985) (finding that surgically extracting a bullet would violate the Fourth Amendment even if the government had probable cause and sought a warrant in advance); Jacob Sullum, *Victim of Dog-Authorized Anal Assault Receives \$1.6 Million Settlement*, REASON.COM (Jan. 14, 2014, 5:58 PM), <http://reason.com/blog/2014/01/14/victim-of-dog-authorized-anal-assault-re>; Alex Greig, *Boy, 16, Suffers Ruptured Testicle After Rough ‘Pat Down’ by Police Woman and Now Faces Infertility*, DAILY MAIL (Jan. 18, 2014, 12:32 PM), <http://www.dailymail.co.uk/news/article-2541780/Boy-16-suffers-ruptured-testicle-police-pat-infertile.html>. But *Rochin*-style challenges to intrusive searches on the basis of substantive due process may run into trouble in light of *Graham v. Connor*, 490 U.S. 386 (1989), which held that the Fourth Amendment preempts substantive due process challenges for factual circumstances that are governed by it. See also Toni M. Massaro, *Reviving Hugo Black? The Court’s “Jot for Jot” Account of Substantive Due Process*, 73 N.Y.U. L. REV. 1086, 1087 (1998).

Repeated police stops are likely to whittle away a person's sense of democratic belonging and trust. At best, the community will come to regard police presence as a mixed blessing.¹²⁶ In essence, most people are willing to endure the inconvenience of a stop or search despite their innocence, but only if the hassle is infrequent.

B. *The Importance of Hassle*

Hassle is a problem of constitutional importance. If suspicion requirements ensure that the hit rate stays high enough, the individualization requirements should ensure that the hassle rates stay low enough. This simple insight unlocks the motivation behind the requirement of individualized suspicion. Individualization measures the societal costs of a law-enforcement program, and suspicion measures its justification. Both are crucial, and they depend on one another to cabin law enforcement appropriately. This Section explores the theoretical and practical importance of hassle to Fourth Amendment interests.

A constitutionally sound hit rate reveals nothing about a program's hassle rate until we know the base rate for the crime and the miss rate for the program.¹²⁷ An example will illustrate the point. Suppose the Boston Police Department develops a profile to detect a particular crime, and the profile has a $33\frac{1}{3}\%$ hit rate. When the profile identifies a suspect, two out of three times it is wrong, and the search is fruitless.¹²⁸ If the crime is quite rare (for example, murder), then the hassle rate is guaranteed to be low as long as the hit rate is respectable (as it is here). In 2011, only 403 murders occurred in the entire city of Boston.¹²⁹ Even if, by some miracle, the profile managed to detect *every* murder (a miss rate of 0), the false positives would have affected only 806 people. In a city of nearly 640,000 people,¹³⁰ this works out to a hassle rate of 0.13%: that is, a 0.13% chance that an innocent person would be questioned or searched in connection with a murder. Put another way, out of 100,000 innocent people, a maximum of 126 innocent individuals

126. CTR. FOR CONSTITUTIONAL RIGHTS, *supra* note 123.

127. Christopher Slobogin and Craig Lerner have discussed the relationship between false positives and crime base rates. Lerner, *supra* note 28, at 444–46; Christopher Slobogin, *Government Data Mining and the Fourth Amendment*, 75 U. CHI. L. REV. 317, 325 (2008). Slobogin is troubled when false positives outnumber hits—in other words, any time the hit rate falls below 50%. But since probable cause does not usually require a preponderance of evidence, a hit rate below 50% should not automatically make us wary of algorithms.

128. For an example of algorithmic detection, consider the algorithms identifying credit-card theft and fraud. The hit rate for these sorts of algorithms was 1 hit for every 2.6 suspects in 2003 (which amounts to 27%). AMY BELASCO, CONG. RESEARCH SERV., RL31786, TOTAL INFORMATION AWARENESS PROGRAMS: FUNDING, COMPOSITION, AND OVERSIGHT ISSUES 15–16 (2003).

129. *Crime Rate in Boston, Massachusetts (MA)*, CITY-DATA.COM, <http://www.city-data.com/crime/crime-Boston-Massachusetts.html> (last visited Aug. 27, 2014).

130. *Boston Population 2013*, WORLD POPULATION STAT. (Sept. 3, 2013), <http://www.worldpopulationstatistics.com/boston-population-2013/>.

would be stopped or searched. Again, this assumes that *every* murder is detected—it represents an upper bound for hassle. I suspect most people would be willing to take these odds of having to undergo a stop or search if it meant that the police department were able to detect every last murderer.

By contrast, if the crime is quite common, such as theft, then even a respectable hit rate cannot guarantee a low hassle rate unless it also happens to have a high miss rate. In 2011, there were just under 88,000 thefts in Boston.¹³¹ Let's assume again that the profile has a 33 $\frac{1}{3}$ % hit rate—one-third of the suspects identified turn out to be thieves. If the algorithm were deployed over the entire city and managed to detect all of the 90,000 thieves, it would have caused an additional 180,000 false-positive stops or searches in the process. That is a lot of hassle for a city of Boston's size. If the algorithm avoided searching the same innocent person more than once, the average Bostonian would face a 28% chance of being stopped or searched in the course of the year.¹³²

Still, even for theft, a profile with a 33 $\frac{1}{3}$ % hit rate could be used without reaching these astronomical hassle rates. The police department could keep the hassle rate low by using the profile sparingly—that is, by using it to identify suspects less often. Or the profile itself may keep the hassle rate low if it regularly fails to alert, even in the presence of thieves. But these are dynamics that often go unnoticed by courts, which have so far focused on ensuring only that hit rates are high enough. By considering hit rates alone, courts risk accepting investigation methods with high hassle rates and rejecting methods with low ones.

The hassle is potentially much worse under the more lenient *Terry* standard for stops and frisks, which requires only a reasonable suspicion of criminal conduct and officer danger rather than probable cause.¹³³ Since *Terry* stops are used to detect a wide range of offenses, some of which are quite common (high base rates), a program that satisfies the reasonable suspicion standard could cause a good deal of pain and grief, as measured by hassle.

Two snapshots of New York City demonstrate the problem. During a two-year period from 1997 to 1998, the Street Crime Unit of the New York Police Department (“NYPD”) stopped 45,000 people based on reasonable suspicion.¹³⁴ These stops resulted in 9,500 arrests—a hit rate of 21%.¹³⁵ The rest—the other 35,500—were false alerts. In absolute terms, this looks like a

131. *Crime Rate in Boston, Massachusetts (MA)*, *supra* note 129.

132. Of course, the rate would be lower if some innocent people were searched more than once, as would almost certainly be the case without additional intervention by the police. But this would cause another type of hassle—multiple searches for those innocents who manage to trigger the algorithm more than once.

133. *Terry v. Ohio*, 392 U.S. 1, 27 (1968).

134. Jeffrey Goldberg, *The Color of Suspicion*, N.Y. TIMES, June 20, 1999, § 6 (Magazine), at 51, available at <http://www.nytimes.com/1999/06/20/magazine/the-color-of-suspicion.html>.

135. *Id.*

lot of stops, but in a city of over 7 million people (at the time),¹³⁶ the hassle rate was actually quite low. Assuming that each stopped suspect was unique,¹³⁷ the average New Yorker had only a 0.5% chance of being stopped during the two years. Of course, the chance of being stopped was not actually distributed evenly across society, so the hassle rate even in the late 1990s may have been disproportionate for some precincts and for some groups defined by race and gender. For now, let's put aside these equitable distribution problems. We will return to them shortly.

Contrast the 1997–1998 hassle rates with the rates that developed in 2010–2011, at the height of NYPD's controversial stop-and-frisk program.¹³⁸ During those two years, New York police conducted nearly 1.3 million stops.¹³⁹ The city's population had grown to nearly 8.3 million by that time,¹⁴⁰ so if each stopped suspect was unique (and of course the suspects were not—some were stopped more than once), the chance that an average New Yorker would be stopped during the two-year period was over 16%.¹⁴¹

The magnitude of the stop-and-frisk program, and the fact that the vast majority of stops were fruitless intrusions on the innocent, convinced Judge Scheindlin to find the program unconstitutional in *Floyd v. City of New York*.¹⁴² The program had several Fourth and Fourteenth Amendment infirmities. The absence of sufficient suspicion was one of them. (Only 6% of the stops resulted in an arrest, a rate that was too low even under the more

136. *Estimated Population by Sex and Age, July 1, 1998*, N.Y. STATE DEPARTMENT OF HEALTH, http://www.health.ny.gov/statistics/vital_statistics/1998/table01.htm (last visited Aug. 27, 2014).

137. This is not very likely, but to simplify the discussion, I will assume that the costs of stopping 35,500 unique individuals are at least as harmful (if not more so) as the costs of exposing some of the population to more than one stop. That is, I will assume that stopping one person twice is at least as bad as stopping two different people.

138. Between 2002 and 2011, NYPD under the leadership of Commissioner Kelly made a concerted effort to increase the use of *Terry* stops in order to reduce violent crime. *Floyd v. City of New York*, 959 F. Supp. 2d 540, 591–602 (S.D.N.Y. 2013).

139. See Second Supplemental Report of Jeffrey Fagan at 10, *Floyd v. City of New York*, 959 F. Supp. 2d 540 (S.D.N.Y. 2013) (No. 08 Civ. 01034).

140. *Population, Land Area, and Population Density by County, New York State – 2011*, N.Y. STATE DEPARTMENT OF HEALTH, https://www.health.ny.gov/statistics/vital_statistics/2011/table02.htm (last visited Aug. 27, 2014).

141. The Newark Police Department (“NPD”)’s stop-and-frisk practices apparently were also quite aggressive. A recently released U.S. Department of Justice (“DOJ”) study of NPD analyzed 39,308 stops made over the course of just two years. Because Newark had a population of only 277,140, a typical Newark resident had a 14% chance of being stopped during that period. CIVIL RIGHTS DIV., U.S. DEP’T OF JUSTICE, INVESTIGATION OF THE NEWARK POLICE DEPARTMENT 5, 8 (2014), available at <http://www.justice.gov/usao/nj/Press/files/pdf/2014/NPD%20Findings%20Report.pdf>. Unfortunately, DOJ was not as careful about reporting hit rates as Judge Scheindlin. Thus, although the report thoroughly describes NPD’s practice of failing to document its rationale for conducting stops, frisks, and searches, it is difficult from the outside to be sure that NPD was as astoundingly bad at selecting individuals to stop as NYPD had been.

142. *Floyd*, 959 F. Supp. 2d at 562.

permissive *Terry* standard.¹⁴³) But one of the most serious constitutional flaws, and the first detail mentioned in Judge Scheindlin's decision, was the astronomical number of stops conducted under the program. The great amount of hassle that the searches created was as concerning to Judge Scheindlin as the low level of suspicion supporting them.

C. *Individualization Reduces Hassle*

The last Section explained why courts should monitor hassle as a Fourth Amendment interest separate from suspicion. This Section demonstrates that they already do. Although courts and scholars rarely make explicit reference to the concept of hassle, the practices that pass the mysterious mandate for individualization nonetheless have the effect of reducing hassle rates.¹⁴⁴ They do so by using natural limits on how many individuals a law-enforcement agency can investigate at one time. Informant and witness information is difficult to investigate. Suspicious bulges usually go unnoticed unless a police officer happens to be nearby. And police dogs are not so numerous that they can be used everywhere at once. A police unit cannot practicably expand these old practices to investigate large swaths of the population, and these resource limitations keep the hassle rates down. Although the connection to hassle is subconscious, it is not coincidental.

In the course of considering whether the government had sufficient suspicion to stop or search a suspect, courts often incorporate an analysis of hassle. Consider the case *Reid v. Georgia*,¹⁴⁵ in which the Supreme Court decided that a Drug Enforcement Administration ("DEA") officer could not stop a drug-courier suspect in an airport based on the facts that the suspect (1) arrived from Fort Lauderdale, a city known to be a source for cocaine distribution; (2) arrived early in the morning; (3) appeared to be avoiding the perception that he was traveling with his companion; and (4) arrived with no luggage other than a shoulder bag.¹⁴⁶ The Court concluded that this combination of factors could not support reasonable suspicion because the circumstances "describe a very large category of presumably innocent travelers, who would be subject to virtually random seizures were the Court to conclude that as little foundation as there was in this case could justify a seizure."¹⁴⁷

This analysis touches on both hit and hassle rates. The conclusion that the combined factors would amount to "virtually random" selection shows that the Court questioned whether this profile had a better-than-random hit rate. This inquiry goes to the suspicion requirement. Reasonable minds

143. *Id.* at 558. Another 6% resulted in tickets or summons. *Id.*

144. Hassle maps fairly well onto Colb's Innocence Model. See Colb, *supra* note 91, at 1476-77.

145. 448 U.S. 438 (1980) (per curiam).

146. *Reid*, 448 U.S. at 439-41.

147. *Id.* at 441.

could differ on whether the Court got this right—the third and fourth factors may increase a hit rate more than the Court cared to acknowledge. But the Court was quite clearly animated by hassle as well. The justices were concerned that the profile failed to exclude enough “presumably innocent travelers” who could have been swept up by the profile that the DEA agent employed.¹⁴⁸ This concern would persist even if the profile *did* have a decent enough hit rate to satisfy the suspicion requirements.

Reid is not unusual. Courts frequently blend and blur the Fourth Amendment interests in suspicion and hassle, treating them as a single goal. One court explained the reasonable suspicion requirement by rhetorically asking, “[I]s it not better to frustrate the prosecution of an individual who may be guilty so that innocent citizens need not be fearful of a police stop and frisk under the circumstances here?”¹⁴⁹ And the Fourth Circuit Court of Appeals at one point had worked its way precisely to the notion of hassle rates. Describing what is required for using drug-courier profiles, the court said that “[t]he articulated factors together must serve to eliminate a substantial portion of innocent travelers before the requirement of reasonable suspicion will be satisfied.”¹⁵⁰ The Fourth Circuit did not recognize the radical nature of its approach. It was breaking the tradition of focusing on the defendant’s facts and instead looking at an investigation’s effects on everybody else.¹⁵¹

Jurists and scholars have not needed to disaggregate the suspicion and hassle issues because, until recently, most programs with decent hit rates were labor intensive and therefore had low hassle rates. But as policing tools become more sophisticated, enabling law enforcement to use computing power to sift through large amounts of digital information, these concepts will begin to diverge.

The separate and independent importance of hassle rates is on spectacular display in the context of the NSA’s surveillance programs. In the wake of the Edward Snowden leaks, the White House unsealed a Foreign Intelligence Surveillance Court opinion from a 2011 case decided by Judge Bates.¹⁵² This opinion is the first decision known to have invalidated aspects of the

148. *Id.*

149. *United States v. Carter*, No. Crim.A.99-50, 1999 WL 1007044, at *8 (D. Del. Oct. 22, 1999).

150. *United States v. Foreman*, 369 F.3d 776, 781 (2004).

151. Even Tribe and Taslitz hint at hassle rates while articulating their own suspect-centered theories for individualization. See Taslitz, *supra* note 18, at 176 (“[T]he citizen still retains significant control over the size of the *risk that his liberties will be infringed* by the police. Probable cause and reasonable suspicion thus help to protect citizen autonomy.” (emphasis added)); Tribe, *supra* note 5, at 1385 (asking that a system of justice take into account how many innocent men are likely to be erroneously convicted, which is a version of the hassle rate, albeit in absolute terms).

152. *In re Government’s Ex Parte Submission of Reauthorization Certification and Related Procedures*, at 78 (Foreign Intelligence Surveillance Ct. 2011), available at <http://goodtimesweb.org/covert-operations/2013/odni-fisc-memo-oct-3-2011.pdf> (declassified in 2013).

PRISM/Upstream program on constitutional as well as statutory grounds. The reasons were compatible with hassle.

Judge Bates wrote the opinion in response to new information about the scale of the NSA's Upstream program, under which the federal government directly collects information on Internet transactions (as opposed to obtaining transaction information indirectly from third parties, as the government does under the PRISM program).¹⁵³ The NSA was required under the Foreign Intelligence Surveillance Act of 1978 ("FISA") to use responsible means to "target" its information gathering at foreign communications and to minimize the collection or retention of any purely domestic communications unintentionally obtained in the process.¹⁵⁴

When the FISA court learned that the NSA was not able to eliminate the chance of collecting purely domestic communications, the court undertook its own study of a random sample of collected conversations in order to assess independently the NSA's filtering error. The court discovered that the filter's error rate was very low—only 0.197% of the collected communications were purely domestic. But because the NSA collected so many conversations (nearly 12 million in a six-month period), the error was great in absolute terms. The court inferred from its independent assessment that the NSA collected between 48,000 and 56,000 purely domestic communications each year.¹⁵⁵

The government predictably argued that the FISA court should be satisfied with the filter's hit rate since well over 99% of the collected communications involved a foreign party. The court responded, "That is true enough, given the enormous volume of Internet transactions acquired by NSA through its upstream collection But the number is small only in that relative sense."¹⁵⁶ Because the government overcollected domestic communications in an absolute sense and because it failed to implement reasonable minimization procedures, the FISA court held that the Upstream program violated the Fourth Amendment.¹⁵⁷

Although the national-security context removes this case from the street-level Fourth Amendment setting, the reasoning in the case foreshadows future litigation over law-enforcement technology. The mechanical nature of the NSA's filter did not put off the court. But the filter's hit rate proved insufficient on its own to justify the government's action. That a large number of individuals' conversations were accidentally swept into the

153. *Id.* at 29–30.

154. 50 U.S.C. §§ 1881a(d)(1), (e)(1) (2012); *In re* Government's Ex Parte Submission of Reauthorization Certification and Related Procedures, at 14.

155. *In re* Government's Ex Parte Submission of Reauthorization Certification and Related Procedures, at 34 n.32 ("Thus, NSA may be acquiring as many as 46,000 wholly domestic 'about' SCTs each year, in addition to the 2,000–10,000 MCTs referenced above.").

156. *Id.* at 73.

157. *Id.* at 78.

search was also an important factor. This measured the hassle of Upstream.¹⁵⁸ And it was enough, despite the very high hit rate, to raise constitutional concerns.

Courts have revealed a need and desire to consider hassle rates as a variable distinct from hit rates (which are already promoted by the Fourth Amendment's suspicion requirement). Since existing determinations of individualization correspond closely with practices that reduce the hassle rate, all that remains is tying the strands together. Individualization attempts to reduce hassle.

When a court considers the constitutionality of law enforcement's individualized suspicion, it should ensure that the government's investigative technique has a high enough hit rate to meet the suspicion standard and a low enough hassle rate to meet the individualization standard. Depending on its facts, a program would fall in one of the four quadrants in the following table:

TABLE 1.
HOW SUSPICION AND INDIVIDUALIZATION INTERSECT

SUSPICION	High Hit Rate	Constitutional	Too much hassle
	Low Hit Rate	Not enough suspicion	Not enough suspicion, too much hassle
		Low Hassle Rate	High Hassle Rate
		INDIVIDUALIZATION	

The table can come to life with a few illustrations. NYPD's stop-and-frisk program was unconstitutional because it operated without sufficient suspicion *and* with too much hassle. It would fall in the lower right quadrant.

158. While the two-part inquiry on hit rates and hassle rates seems wise, the hassle rate does not appear to raise the serious doubts that it did for the court. Fifty-six thousand conversations collected from a population of 300 million Americans amounts to a hassle rate of a 0.02% chance of inadvertent collection—or 19 out of 100,000 people. This rate seems especially low in light of the Supreme Court's one and only pronouncement on the Fourth Amendment's application to national-security investigations, in which the Court required *some* limits on governmental investigation but permitted the limits to be less restrictive than those applying to ordinary law enforcement. *United States v. U.S. Dist. Court (Keith)*, 407 U.S. 297 (1972).

An anonymous and uncorroborated tip, by contrast, would fall in the lower left quadrant. It would not cause a lot of hassle, but it fails the suspicion element since case law requires some corroboration or another objective reason to credit a tip's accuracy.¹⁵⁹

The upper left quadrant marks the constitutional ideal. These police investigations operate with enough suspicion to instill trust in the program and enough constraint to promise a low impact on the liberty of the average person. Corroborated tips and prudent use of *Terry* stops belong here.

The last quadrant, the upper right, will gain importance as some policing practices become automated. Big-data analytics may be able to meet the suspicion requirements, but if they operate on a large scale, they can quickly increase the number of fruitless searches and seizures.¹⁶⁰ Without constraints, data analytics can violate hassle thresholds even if their hit rates satisfy the suspicion element.

For example, if a pattern-based data-mining program detects copyright infringement with a 90% hit rate, the suspicion requirement would be satisfied. But its use may nevertheless fail the individualization requirement. Copyright law is violated so frequently that the relatively rare misfire could result in hundreds of thousands of futile investigations.¹⁶¹ If the analysis of individualization tunes into hassle, courts will be well equipped to identify constitutional problems with large-scale data-analytics programs without gutting the innovations entirely. The implications of hassle-driven individualized suspicion are explored in Part III.

Before we turn to implications, however, we will first explore how hassle explains some otherwise inexplicable patterns in Fourth Amendment case law. Even though courts have not referred often to hassle-style problems, there is ample evidence that they have been striving to reduce hassle all along. The indirect influence of hassle is considered next.

D. *Other Instincts Explained*

Courts sense the need to keep track of hassle, although they have not made a direct connection between hassle and individualization. This Section shows that linking individualization with hassle goes a long way toward explaining some otherwise curious jurisprudence. Indeed, the concept of hassle is legitimated by its subconscious application in the case law. First, hassle explains why courts sometimes react unfavorably to searches based on odd, but technically legal, behavior, while at other times they do not. Second, hassle explains why courts prefer profiles with many factors. Third, hassle

159. *Alabama v. White*, 496 U.S. 325, 330 (1990) (requiring the quantity or quality of an informant's tip to justify reasonable suspicion and giving strong weight to information that is independently verified by the police).

160. *Ferguson*, *supra* note 85.

161. Approximately 46% of American adults have pirated music, TV shows, or movies. THE AM. ASSEMBLY, COPYRIGHT INFRINGEMENT AND ENFORCEMENT IN THE US 2 (2011), available at <http://piracy.americanassembly.org/wp-content/uploads/2011/11/AA-Research-Note-Infringement-and-Enforcement-November-2011.pdf>.

offers another reason courts are unwilling to define probable cause and reasonable suspicion using precise probabilities. And finally, hassle establishes a concrete Fourth Amendment interest through which courts can demand distributive justice. Empirical analyses of racial bias in law enforcement have already exposed the disparate rate of stops and searches across races. Hassle gives these findings a home in Fourth Amendment doctrine.

1. Odd but Legal Behavior

In *People v. Parker*, a police officer suspected that a student who walked into a public school building, looked at the metal detector, and turned around to leave the building had a weapon.¹⁶² The Illinois Appellate Court ruled that the officer lacked particularized suspicion because the student “could have just turned around and gone home for any number of reasons, being sick, forgot something, forgot his lunch, forgot his books, forgot his homework or what have you.”¹⁶³ But just as important as what the *Parker* court said is what it did not say. The court did *not* suggest that the police officer’s stop and search were unlikely to produce a gun. It didn’t even say that these innocent explanations were more likely than the guilty explanations to justify the student’s behaviors.¹⁶⁴ For its part, the Pennsylvania Supreme Court used similar reasoning when it decided that a hand-to-hand trade of money for a “small object” occurring late at night in a tough part of North Philadelphia could not support probable cause.¹⁶⁵

These cases seem to conflict with Supreme Court precedents like *Illinois v. Gates*¹⁶⁶ and *Illinois v. Wardlow*,¹⁶⁷ both of which involved odd but legal behavior. *Gates* involved odd travel behavior that corroborated an anonymous tip.¹⁶⁸ And in *Wardlow*, a young man’s flight in reaction to the arrival of police was enough to justify an individualized suspicion even though some members of the Court acknowledged that there are innocent explanations for flight.¹⁶⁹

What explains courts’ willingness to accept innocent but odd behaviors as a basis for suspicion in some cases but reject them in others? It cannot be the mere possibility of innocent explanations. Those existed for both sets of cases.

Hassle can explain these cases’ seemingly schizophrenic outcomes. When police must rely on a tip or when they must be physically present to

162. 672 N.E.2d 813 (Ill. App. Ct. 1996).

163. *Parker*, 672 N.E.2d at 817.

164. It is possible that the chance of an innocent explanation *does* outweigh the chance that the student brought a weapon. My objection is that the court does not clarify that it harbors such a belief.

165. *Commonwealth v. Dunlap*, 941 A.2d 671, 679 (Pa. 2007), *disapproved of on other grounds* by *Commonwealth v. Thompson*, 985 A.2d 928 (Pa. 2009).

166. 462 U.S. 213 (1983).

167. 528 U.S. 119 (2000).

168. *Gates*, 462 U.S. at 226.

169. *Wardlow*, 528 U.S. at 124–25.

provoke flight, the investigation methods simply cannot scale up. Physical presence is costly, and tips are lucky. Even if a significant proportion of a jurisdiction's population would flee at the sight of policemen or engage in bizarre travel behavior, in order to take advantage of the case law, the police actually have to receive a tip or be physically present. But they cannot possibly show up to every public gathering of all youths who would decide to flee, and they will not receive tips about most of the people who engage in strange travel behavior.¹⁷⁰

Hassle rates may be driving these decisions where the government does not rely on some external limiting factor like a tip or a person's reaction to the physical presence of police. Circling a block and emerging with a bag is unusual behavior, to be sure. But we would not want *everybody* who circles a block and emerges with a bag of souvenirs, or who turns around at a metal detector because he forgot his homework, to become subject to a stop or search. Without something fortuitous like a tip, there is no external constraint on a police department that uses new technology to increase the number of stops or searches based on similar facts.¹⁷¹

2. A Preference for More Factors

When assessing an officer's decision to stop or search somebody, courts prefer to receive a long list of reasons justifying the decision. The more reasons the agent can recount, the better.¹⁷² This preference, too, can be explained by the invisible influence of hassle.

Consider drug-courier profiles. The Supreme Court condoned a DEA agent's decision to stop a man named Andrew Sokolow in the Honolulu airport as he returned from a trip to Miami.¹⁷³ The agent suspected that Sokolow was carrying drugs because he (1) paid \$2,100 for airplane tickets in cash; (2) traveled under a name that did not match the records for the telephone number he had provided; (3) came from Miami, "a source city for illicit drugs"; (4) stayed in Miami only forty-eight hours, even though he

170. I assume, for the purposes of this discussion, that the differences in hit rates between the first set of cases and the second set are negligible. In fact, courts rarely look for empirical evidence that searches tend to produce evidence of a crime.

171. This may be especially problematic if the suspect engages in the legal but suspicious behavior as part of what Elizabeth Joh calls a "privacy protest." These are practices of avoiding consented police encounters, buying burner phones with cash, using Tor to browse the Internet, and engaging in other activities intentionally designed to avoid police interaction. See Elizabeth E. Joh, *Privacy Protests: Surveillance Evasion and Fourth Amendment Suspicion*, 55 ARIZ. L. REV. 997 (2013).

172. The knock-and-announce case discussed earlier in the Article provides a good example. See *supra* notes 39–44 and accompanying text. Knowing that the defendant was going to be arrested for a drug-related crime did not provide probable cause to believe that evidence would be destroyed with advanced notice of the arrest, but adding one more factor (a recognition that the police were present) was sufficient. See also Taslitz, *supra* note 18, at 158 ("If hundreds of traits were added into the mix, it would be hard to see an individual [suspect] as just a collection of stereotypes.")

173. *United States v. Sokolow*, 490 U.S. 1 (1989).

spent over twenty hours in an airplane to get between Miami and Honolulu; (5) appeared nervous; and (6) checked no luggage.¹⁷⁴ The Court recognized that each one of these observations could have an innocent explanation but noted that, when taken together, these noncriminal acts could contribute to a “degree of suspicion” that surpasses the reasonable suspicion standard.¹⁷⁵

The Court’s reasoning is straightforward enough. The chance that a person flying out of Miami has narcotics on him is not very high. And the chance that a person flying with no checked luggage has narcotics on him is also not very high. But the odds increase when the person flies out of Miami *and* has no luggage. As more and more factors are added to the Venn diagram, the hit rate may continue to increase, and at some point the probability in the intersection will surpass the reasonable suspicion threshold.

This all seems perfectly reasonable, but given the low bar for reasonable suspicion, it is not clear that the DEA agent needed more than one or two of the factors to meet the threshold. The fact that Sokolow flew twenty hours and stayed in Miami for only two days may have on its own increased the chance of his carrying drugs to a figure that clears the suspicion threshold. And if that fact alone didn’t do it, perhaps because business travelers often fly long distances for short stays, adding just one more fact—that he paid for his ticket in cash—would have helped confirm that Sokolow was not a business traveler.¹⁷⁶ These two factors would have easily cleared the hit-rate hurdle, leaving us to wonder what work the others are doing.¹⁷⁷

Hassle helps explain why courts would prefer to see more factors than are necessary in order to clear the suspicion hurdle, even if the marginal returns on accuracy are negligible. Adding factors to the Venn diagram has an exclusionary effect. Each factor has the potential to exclude a swath of the population from the possibility of a search or seizure. Courts are reassured by longer lists of justifications because these lists roughly signal that the agent’s model cannot scale to a large number of people, many of whom may be innocent.

At their best, long lists of factors improve hit rates and decrease hassle rates at the same time.¹⁷⁸ As each factor contributes to the program’s accuracy, it also constrains the program’s scope. But some of the most frequently used suspicion factors appear to increase accuracy and constrain scope without actually doing so.¹⁷⁹

174. *Id.* at 3–4. The suspect also wore flashy clothing (a jumpsuit and gold jewelry). *Id.* at 4.

175. *Id.* at 9–10.

176. Even at the time, business travelers paid their airfare by check or credit card. *Id.* at 8.

177. Indeed, one of the factors (traveling with no luggage) flows from another (short trip).

178. Sharon Davies captures this idea with the phrase “circle of suspicion.” A good profile will narrow the circle’s scope in order to achieve a decent level of prediction. Sharon L. Davies, *Profiling Terror*, 1 OHIO ST. J. CRIM. L. 45, 52 (2003).

179. See Lerner, *supra* note 28, at 434–43, for a description of commonly applied subjective factors.

3. Undefined Suspicion Standards

Hassle can also explain why the Supreme Court has avoided giving mathematical precision to the probable cause and reasonable suspicion standards. The Court has declared that “[t]he probable-cause standard is incapable of precise definition or quantification into percentages because it deals with probabilities and depends on the totality of the circumstances.”¹⁸⁰ This particular explanation is immensely unsatisfying. The Court has attracted a lot of criticism for refusing to give a probability threshold for the probable cause and reasonable suspicion standards.¹⁸¹

But some scholars have come to the Supreme Court’s defense. Lerner argues that the standards do and should have differing tolerances based on the heinousness of the investigated crime.¹⁸² And Kerr argues that the probable cause standard must have some flexibility to accommodate factors relating to an investigation—factors that fall outside the four corners of a warrant application but that the judge is likely to know and use.¹⁸³

Hassle offers another reason to avoid fixing the probability thresholds for probable cause and reasonable suspicion. When the police investigate a crime with a low base rate, the suspicion rate can be lower than usual without significantly changing the hassle rates. For example, if police are investigating bribery, the crime is so infrequently committed that a few additional searches per bribe spread over an entire city’s population will not significantly affect the odds that a person will be searched.¹⁸⁴

Hassle may also explain why courts might prove more lenient with suspicion requirements in crime-out investigations.¹⁸⁵ When police are pursuing suspects of a single crime, searching four or five suspects will not affect a town’s hassle rate even though the hit rate may fall below the usual 33¹/₃% standard.¹⁸⁶

4. The Search for Distributional Justice

Finally, hit and hassle rates are attractive to courts and civil rights lawyers searching for a measure of law-enforcement bias. Large investigation

180. *Maryland v. Pringle*, 540 U.S. 366, 371 (2003).

181. *E.g.*, Harcourt & Meares, *supra* note 19, at 850. This is also consistent with what David Faigman calls the Court’s “nonrealist approach to the subject of constitutional application.” FAIGMAN, *supra* note 47, at 172.

182. Lerner, *supra* note 28, at 460; Craig S. Lerner, *The Reasonableness of Probable Cause*, 81 TEX. L. REV. 951 (2003).

183. Kerr, *supra* note 2, at 132–33.

184. Even if such additional searches did affect hassle rates, some crimes may be heinous enough to allow for more hassle, and lower hit rates, because our preference to root out a particular crime is so great that it changes the contours of our tolerances. Tribe maps out some examples of hit-rate and hassle-rate indifference curves in the context of criminal convictions. *See* Tribe, *supra* note 5, at 1387–88.

185. For a definition of crime-out investigations, see *supra* Section I.D.

186. A crime-out investigation can be thought of as a very low base rate crime—a crime that has been committed only once. Thus, this is merely a corollary of the low base rate point.

programs cause a lot of hassle even when their hit rates are high (and especially if they aren't). If the hassle disproportionately affects a suspect class, this could be useful evidence of bias.

Racial disparities in hit rates can be a telltale sign of police bias: if searches conducted on cars driven by minorities result in the recovery of contraband less often than searches conducted on cars driven by whites, the disparity suggests that the department might suffer from explicit or implicit bias that leads officers to believe that minorities are more likely to be engaged in criminal conduct than they actually are.¹⁸⁷ But the comprehensive record of stops can also reveal disparate hassle rates, *even if* hit rates are the same across all races. Sure enough, hassle rates (stops or searches per one hundred thousand or sometimes per ten thousand inhabitants) are routinely used in disparate impact studies of police programs.¹⁸⁸

For the last ten years, New York City's police department has operated under record-keeping obligations,¹⁸⁹ and the resulting data were critical to the constitutional challenge to NYPD's stop-and-frisk practices.¹⁹⁰ Judge Scheindlin was disturbed by the disparity between the rates at which black and Hispanic New Yorkers were stopped by NYPD and the rates at which whites were stopped.¹⁹¹

The government's expert suggested that the demographics of New York City's criminals would be unlikely to mirror the demographics of the city's entire population.¹⁹² If NYPD's stop-and-frisk program operated on a smaller scale—or operated for the purpose of rooting out a narrowly defined criminal act—the government's objections to Judge Scheindlin's reasoning would have some weight. The demographics of suspects selected from nuanced profiles with a good hit rate are often likely to diverge from the general demographics. For example, algorithms designed to detect white-collar crime are more likely to direct investigations disproportionately

187. Judge Scheindlin noted the disparity in hit rates in her *Floyd* opinion. Contraband and weapons were recovered from white suspects more often than they were from black suspects. *Floyd v. City of New York*, 959 F. Supp. 2d 540, 559 (S.D.N.Y. 2013).

188. See, e.g., IAN AYRES & JONATHAN BOROWSKY, A STUDY OF RACIALLY DISPARATE OUTCOMES IN THE LOS ANGELES POLICE DEPARTMENT (2008), available at <http://islandia.law.yale.edu/ayres/Ayres%20LAPD%20Report.pdf> (finding that, out of 10,000 white residents, the Los Angeles Police Department stopped only 1,750, whereas it stopped 4,500 out of 10,000 black residents; that minorities were also more likely to be searched than whites; and that the searches of minorities yielded lower hit rates for contraband); STEPHEN M. HAAS ET AL., WEST VIRGINIA TRAFFIC STOP STUDY: FINAL REPORT (2009), available at http://www.djcs.wv.gov/SAC/Documents/WVSAC_Traffic_NEWOverviewofStatewideFindings2009.pdf (using search rates and hit rates to assess racial disparities); GREG RIDGEWAY, ANALYSIS OF RACIAL DISPARITIES IN THE NEW YORK POLICE DEPARTMENT'S STOP, QUESTION, AND FRISK PRACTICES (2007), available at http://www.rand.org/content/dam/rand/pubs/technical_reports/2007/RAND_TR_534.pdf.

189. *Floyd v. City of New York*, No. 08 Civ. 1034(AT), 2014 WL 3765729, at *3 (S.D.N.Y. July 30, 2014).

190. *Floyd*, 959 F. Supp. 2d at 579, 588.

191. See *id.* at 560.

192. See *id.*

on educated white men.¹⁹³ Likewise, if NYPD's stop-and-frisk program succeeded in detecting crime, it would be appropriate for courts to permit some divergence between the demographics of the stopped and those of the whole population.¹⁹⁴ Establishing an appropriate baseline would have been difficult without good data. Judge Scheindlin was able to avoid the difficult task of establishing such a baseline by pointing to the program's ineffectiveness. Because the vast majority of stops (94%) did not lead to arrest, bona fide predictors of crime could not explain the disparity in stops.¹⁹⁵ NYPD's stop-and-frisk program had such a low hit rate, and was so active, that the enterprise consisted almost entirely of hassle. And that hassle had an outsize effect on minority communities.¹⁹⁶

Hassle rates are also very useful in combination with hit rates in order to show whether a police department or prosecutor declines to enforce the criminal laws more frequently for white criminals than for minority criminals, even when an officer or prosecutor has been presented with identical behaviors.¹⁹⁷ For example, Samuel Gross and Katherine Barnes took advantage of known hassle rates and base rates to determine whether Maryland's highway patrol was stopping and searching speeding minority drivers at higher rates than speeding white drivers (it was).¹⁹⁸

Given the prominence of hit rates as a measure of discrimination, it is surprising that the harm from hassle lacks a well-established home in Fourth Amendment doctrine.

If hassle and hit rates were recognized as the twin guiding lights for individualized suspicion, Fourth Amendment law would develop in a much more rational way. The next Part explores this potential.

193. See Laurie L. Ragatz et al., *The Psychological Profile of White-Collar Offenders: Demographics, Criminal Thinking, Psychopathic Traits, and Psychopathology*, 39 CRIM. JUST. & BEHAV. 978, 979 (2012).

194. In New York, at least in the 1990s, victim surveys suggest that about half of robberies were reportedly committed, according to their victims, by black perpetrators. And since more than half of the victims were also black, racial bias cannot explain all of the variance between the suspect reports and the racial composition of the city. See Goldberg, *supra* note 134.

195. Only 6% led to arrest. *Floyd*, 959 F. Supp. 2d at 558. But another 6% resulted in the issuance of a summons. *Id.*

196. Indeed, the hit rate was lower for stops of minorities than it was for stops of whites. And subjective factors like "furtive movements" and "bulge" were used more often for black suspects than for white suspects. Report of Jeffrey Fagan, *supra* note 7, app. at tbl.D1.

197. See HARCOURT, *supra* note 5, at 28; Katherine Y. Barnes, *Assessing the Counterfactual: The Efficacy of Drug Interdiction Absent Racial Profiling*, 54 DUKE L.J. 1089, 1099 (2005). Prosecutors might correct some of the racial disparities that come earlier in the criminal justice process. Christopher L. Griffin, Jr. et al., *Corrections for Racial Disparities in Law Enforcement*, 55 WM. & MARY L. REV. 1365, 1368–76 (2014).

198. See Gross & Barnes, *supra* note 117, at 665–66. Gross and Barnes had to use an independent measure of the base rate for speeding in combination with the hassle and hit rates in order to get at the key measure: the miss rate. The disparity in speeding enforcement was important because enforcing the speed limits was a feeder to drug interdiction. The Supreme Court has approved the use of stops for minor violations to further an investigation for more serious offenses. *Whren v. United States*, 517 U.S. 806, 819 (1996).

III. THE FUTURE OF SUSPICION

This Part considers the ramifications if hassle becomes a lodestar for individualized suspicion. Change will be slow at first. Hassle can provide only loose guidance for law enforcement and courts because they often lack the information needed to calculate hassle rates precisely. (For example, how will we know how many people *actually* pace in front of stores?) Just as with suspicion, hassle will be determined by experience and common sense for the foreseeable future. And much like for suspicion, the tolerance for hassle may vary based on the severity of the crime and on the intensity of governmental intrusion.

This Part shows that focusing individualized suspicion on hit and hassle rates would make its application more sound and consistent, would help eradicate some of the flawed traditional police practices, and would open up opportunities for new methods. The first and most important consequence would be an appetite for better information and a tolerance for accountable experimentation.¹⁹⁹ The second would be the development of a healthy judicial skepticism of police narratives. And third, focusing on hassle would allow law-enforcement agencies to use randomness to constrain the scope of their investigation programs. This Part closes with a review of some common objections to data-driven investigation methods.

A. *More Open and Honest Experimentation*

An explicit focus on hassle would very likely drive a judicial demand for data. This demand represents an indirect but important consequence. A generation ago, the Supreme Court, using a derisive tone, refused to set out quantitative definitions of probable cause.²⁰⁰ In an increasingly tech-savvy world, though, the Court's approach is aging poorly.²⁰¹ Contrary to the Court's predictions, police departments today are quite comfortable using statistics to develop law-enforcement policies.²⁰² And yet despite the growing facility with data, probable cause determinations are not "evidence based" in the sense that the phrase is used in every other serious discipline. Courts are

199. Fourth Amendment law may actually proscribe collecting data for these purposes, especially in light of the concurring opinions in *United States v. Jones*, 132 S. Ct. 945 (2012). *See id.* at 957 (Sotomayor, J., concurring) (positing that there should be some limit to the waiver of Fourth Amendment protection when data are disclosed for one reason but not another).

200. *Illinois v. Gates*, 462 U.S. 213, 231–32 (1983).

201. Faigman criticizes the Supreme Court for its "empirical cowardice" and suggests that constitutional law should begin to establish a "scientifically realistic jurisprudence." FAIGMAN, *supra* note 47, at 8, 26.

202. *See, e.g.*, Erica Goode, *Sending the Police Before There's a Crime*, N.Y. TIMES, Aug. 16, 2011, at A11, available at <http://www.nytimes.com/2011/08/16/us/16police.html>; Rocco Parasandola, *NYPD Commissioner Bill Bratton Tweets About First Compstat Meeting*, N.Y. DAILY NEWS, Jan. 24, 2014, available at <http://www.nydailynews.com/new-york/nyc-crime/nypd-commissioner-bill-bratton-tweets-compstat-meeting-article-1.1589955>.

permitted to define probable cause using their own untested and idiosyncratic assumptions about hit rates. Police methods that are acceptable to the judiciary are constitutional no matter how poorly they actually perform, and methods that are deemed unacceptable are proscribed without any consideration of whether they perform as poorly as the judiciary thinks they do.²⁰³ The judges sit with presumed and constitutionally insulated factual expertise.²⁰⁴

Jeffrey Rachlinski, Chris Guthrie, and Judge Andrew Wistrich collaborated on an impressive research project in order to better understand how state and federal judges respond to the same set of facts.²⁰⁵ Rachlinski and his coauthors were interested in testing whether judges suffer from hindsight bias—that is, whether judges are more likely to find probable cause when they know that the search ultimately uncovered evidence of a crime than when they reviewed the evidence prospectively. To everybody’s surprise, they don’t. Judges seem to be remarkably skilled at putting themselves behind the veil of ignorance for probable cause determinations.²⁰⁶

But the good news was mooted by the judges’ inconsistency with one another. The judges’ estimates for the likelihood of uncovering evidence of a crime, and their ultimate determinations of probable cause, were extraordinarily varied. The 224 judges responding to a hypothetical factual scenario were evenly split over whether probable cause existed to justify a search. About 56% thought it did. We might expect to see this sort of split if the estimates of the probability for a successful search hovered around 30% or 40%. (This would suggest that the facts make for a tough case.) But that is *not* what Rachlinski and his coauthors saw. Instead, the probability estimates were all over the map. The mean probable cause estimate was about 55%, and the standard deviation was a whopping 24 percentage points.²⁰⁷

With estimates this varied, it is clear that our criminal-procedure rules leave the fate of a police investigation and a defendant to the luck of the magistrate draw. To illustrate the arbitrary and subjective nature of the

203. In fairness to judges, the police who are making determinations on the ground may be no better at estimating the chance of success. William Stuntz called this overwhelming challenge “*Terry’s Impossibility*.” William J. Stuntz, *Terry’s Impossibility*, 72 ST. JOHN’S L. REV. 1213 (1998). *But see* Minzner, *supra* note 80, at 931–33 (showing that some patrol officers are very good at selecting vehicles for a car search).

204. Faigman suggests that something like a *Daubert* standard should apply to judicial recognition of facts that determine the constitutionality of an action. FAIGMAN, *supra* note 47, at 100–01.

205. Rachlinski et al., *supra* note 27.

206. *Id.* at 73.

207. Rachlinski provided the standard deviations to me at my request. E-mail from Jeffrey J. Rachlinski, Professor of Law, Cornell Law Sch., to Jane Bambauer, Assoc. Professor of Law, Univ. of Ariz. James E. Rogers Coll. of Law (Feb. 18, 2014, 5:57 AM MST) (on file with author). Adding to the confusion is that judges seem to have different standards for how much probability the probable cause standard requires. *See* C.M.A. McCauliff, *Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?*, 35 VAND. L. REV. 1293, 1327 (1982) (finding that judges’ responses ranged from 10% to 90% when asked how much chance of recovering evidence is required under probable cause).

judges' estimates of probable success, consider the estimates of ten hypothetical judges randomly drawn from the distribution, any one of whom could receive the officer's warrant request²⁰⁸:

TABLE 2.
HYPOTHETICAL SELECTION OF JUDGES'
ESTIMATES OF SEARCH SUCCESS

Judge 1	58%
Judge 2	67%
Judge 3	49%
Judge 4	34%
Judge 5	47%
Judge 6	16%
Judge 7	57%
Judge 8	89%
Judge 9	96%
Judge 10	10%

At first, estimates of hassle will be just as unwieldy as estimates of hit rates (like the ones we see in the Rachlinski study). But in time, if the importance of both hit and hassle rates is emphasized, courts will begin to demand better data. They may come to expect record keeping about the reasons that a stop or search was conducted and its ultimate outcome, information that would give magistrate judges something to depend on, apart from their own intuitions, in estimating the reliability of evidence and the impact of a program on a community's hassle rate. Law-enforcement officers could also have an incentive to keep careful records, which would enable them to rebut a judge's sense that a stop or search was unjustified or likely to cause too much hassle.

Attention to hassle may also increase tolerance for police experimentation. In time, if courts become comfortable digesting criminal-investigation data, they will also understand that hit and hassle estimations are best determined after appropriate small-scale police-investigation experiments. Courts may provisionally accept a theory about predictors of crime, and they could then revisit that theory's Fourth Amendment status after the police learn more about the accuracy of those predictors and the effects on innocent people targeted by the program. This provides some space for law enforcement to develop new predictive profiles.²⁰⁹ New models might incorporate

208. I generated this data using Excel by randomly pulling numbers from a distribution with mean 55.08019 and standard deviation 24.12281. These numbers were the mean and standard deviation for the sample of judges prospectively estimating a hit rate for the hypothetical posed by Rachlinski, Guthrie, and Judge Wistrich. E-mail from Jeffrey J. Rachlinski, *supra* note 207.

209. "Profiling" has become a dirty word among academics, in part because the mind races directly to racial profiling. Even apart from noting their potential racially disparate effects, scholars have criticized profiles for becoming stale and outlasting their predictive utility. HARCOURT, *supra* note 5, at 109. And yet resisting crime profiling is futile as long as law

factors that the police learned from mining their own data, or the models could be informed at first by the low-tech common sense and on-the-ground experiences that have informed profiles for decades, only to be verified (or abandoned) once better data are available.²¹⁰

While experimentation may seem far-fetched for some crimes, there is at least one area where both the documentation and the results lend themselves very well to hit and hassle analysis: DUIs.

Before the police draw blood from a suspected drunk driver, they often must secure a warrant. The Supreme Court recently clarified that a DUI does not create an automatic exigency excusing the police from applying for a warrant,²¹¹ and therefore police are more motivated than ever to secure warrants using application forms that allow officers quickly to check off the conditions and behaviors they observe.²¹²

These prefab forms are ideal for data analysis. Some of the observable behaviors, like vomiting and urinating, need no validation by data analysis. But the long list of other factors (for example, odors, stutters, sways, and hiccups) could be correlated with blood-alcohol levels and used in a model to predict drunkenness with increasing accuracy. The same information can also be mined to see which types of behaviors occur on their own or in combination with others much more often than we would expect—so often that, if the police used these behaviors to make stops or searches, they would run the risk of creating too much hassle. In time, the form could be modified to direct police to use only the factors or combinations of factors that turn out to be predictive when deciding whether to apply for a warrant.

These types of experiments would go a long way toward preparing for a future with more automated law enforcement. They would also give society some relief from the overreliance on police narratives, a topic that we will visit next.

B. Fewer Vague Narratives

In 1986, the DEA established Operation Pipeline, a program designed to provide training in drug-interdiction profiling to DEA agents as well as local

enforcement is charged with the task of detecting crimes that are underreported or inchoate. Section I.E describes the limits of crime-out investigations. All other investigations must use profiling. As one example, detecting tax evasion will depend on profiling. SCHAUER, *supra* note 15, at 160–65 (describing the Internal Revenue Service's use of the discriminate function to select returns for audit). If profiling is inevitable, courts should welcome new methods to improve stale profiles.

210. Drug-courier profiles were formed in precisely this way—based on the observations and hypotheses of both DEA agents and airport employees. HARCOURT, *supra* note 5, at 15.

211. *Missouri v. McNeely*, 133 S. Ct. 1552 (2013).

212. See, e.g., *Affidavit in Support of Telephonic Search Warrant*, MONTANA DEP'T TRANSP., http://www.mdt.mt.gov/tsrp/docs/telephonic_search_warrant.pdf (last visited Aug. 27, 2014) (allowing officer to check off characteristics to indicate drunk driving).

highway patrol.²¹³ Operation Pipeline developed a running list of indicators that could be used to convert a routine traffic stop into a car search based on probable cause to believe that the driver was carrying distribution-level amounts of drugs.²¹⁴ Many of the factors listed in the training manuals are objective. Officers were instructed to pay attention when the amount of luggage was not appropriate for the trip, when fast-food wrappers were scattered throughout the car, when the driver had only a single key on his key ring (a sign of a borrowed car), and when the car had radar detectors, high mileage relative to the vehicle's age, deodorizers, or parts of the door strewn about the floor or seat (a sign that contraband may be stashed inside the door).²¹⁵ All of these factors are sufficiently clear to provide the officers with real guidance.

Other factors, though, are less objective. For example, the list of signs that could indicate excessive nervousness included some concrete factors (if the driver repeatedly rubs his hands together, has a bad case of goose bumps, or vomits) and some hopelessly subjective factors (a "frowning" driver, a "distant look," or just "eyes: the window to the soul").²¹⁶

Ever since the Supreme Court decided *Illinois v. Wardlow*, allowing the combination of "high-crime area" and "unprovoked flight" to justify a *Terry* stop,²¹⁷ the police have regularly used factors like the officer's training and experience,²¹⁸ excessive nervousness,²¹⁹ a "[s]uspicious [b]ulge,"²²⁰ and "[f]urtive [m]ovements"²²¹ to justify their stops and searches. Criminal-procedure scholars have interpreted these subjective factors to be the product of a power-hungry law-enforcement agency that prefers to keep as much discretion for itself as possible,²²² but the courts have been complicit and share some of the responsibility for the frequent use of vague factors. The existing individualization case law encourages and even commands the use of these vague and subjective factors.

213. David Kocieniewski, *New Jersey Argues that the U.S. Wrote the Book on Race Profiling*, N.Y. TIMES, Nov. 29, 2000, at A1, available at <http://www.nytimes.com/2000/11/29/nyregion/new-jersey-argues-that-the-us-wrote-the-book-on-race-profiling.html>.

214. E.g., KAN. HIGHWAY PATROL, OPERATION PIPELINE: PASSENGER VEHICLE DRUG INTERDICTION, available at http://norml.org/pdf_files/brief_bank/Operation_Pipeline_Manual.pdf. The Operation Pipeline program is now infamous for allegedly encouraging racial profiling, but the training manuals avoid referring to race and obvious proxies for race. See Kocieniewski, *supra* note 213.

215. KAN. HIGHWAY PATROL, *supra* note 214.

216. *Id.*; see also Kate Shatzkin & Joe Hallinan, *Highway Dragnets Seek Drug Couriers—Police Stop Many Cars for Searches*, SEATTLE TIMES, Sept. 3, 1992, at B6, available at <http://community.seattletimes.nwsourc.com/archive/?date=19920903&slug=1510873>.

217. 528 U.S. 119, 124 (2000).

218. *United States v. Tussa*, 816 F.2d 58, 62 (2d Cir. 1987).

219. *United States v. Sokolow*, 490 U.S. 1, 3–5 (1989).

220. Harcourt & Meares, *supra* note 19, at 820.

221. *Id.*

222. E.g., Sandra Guerra, *Domestic Drug Interdiction Operations: Finding the Balance*, 82 J. CRIM. L. & CRIMINOLOGY 1109 (1992).

Courts have rejected short, objective profiles out of a respect for case-by-case particularization. In New York, for example, police may not stop a person wearing a balaclava even in the summertime unless they are able to give some other reason to round out the narrative and make the case unique.²²³ As a consequence, police have little reason to cite objective facts and improve their objective profiles. And they have a lot of incentive to rely instead on the soft factors like “furtive movements.” Because the soft factors incorporate the impressions of the officer and the innumerable nuances of street observations, they are personalized enough to satisfy the courts’ demand for a story. And because they are vague enough to encompass a wide variety of real-world scenarios, they make reliable talking points.²²⁴ The vague factors are so efficient that police can’t *not* use them.

Changing the individualization requirement will rectify these perverse incentives. Courts should allow mechanically applied profiles (such as “balaclava + summertime = reasonable suspicion”) as long as the police are prepared to show that the hit rates are high enough and that the hassle rates are low enough. Narratives based on vague perceptions, on the other hand, should fare poorly under both the hit and hassle inquiries. Such perceptions cannot instill sufficient confidence that the stop or search will have a reliable hit rate. And after the experience with New York City’s stop-and-frisk program, courts should think twice before assuming that vague perceptions are sufficiently individualized to keep hassle rates low.

C. *More Randomness*

Back to the Harvard dorm rooms. Suppose the police officer applies for warrants to search ten dorm rooms for illicit drugs. The officer has two pieces of information to offer in support of the request. The first is the methodologically sound study from the introductory hypothetical showing that 60% of Harvard dorm rooms contain drugs. The second is an affidavit and program log showing that the ten dorm rooms were selected using a random number generator and that no other dorm rooms have been searched on the basis of the study.

223. See *People v. Giles*, 647 N.Y.S.2d 4, 6 (App. Div. 1996) (“[W]earing a long winter coat on a hot summer night . . . is no more than ‘odd’ behavior . . .”).

224. Lerner makes a compelling argument that courts are foolish if they expect police to refrain from acting on hunches or if they expect officers to be able to articulate the import of their explanations. He encourages courts to respect an officer’s hunch if his interaction with the suspect was respectful. Lerner, *supra* note 28, at 468–69. Minzner has a different proposal to limit, or at least modify, the use of narratives. He recommends that courts consider a police officer’s past track record as one of the factors contributing to probable cause. Because police officers’ personal hit rates vary and because their hit rates are stable over time, a court should be willing to accept a reliable officer’s hunch by permitting a relatively sparse record of evidence. And conversely, courts should demand more evidence from the officers who have a poor track record. Minzner, *supra* note 80, at 920–22.

Today, a judge would not respond well to this request.²²⁵ He would be appalled at the officer's blatant disregard for the individualized suspicion requirement. But if the logic of this Article is persuasive, the warrant application should be granted. The study satisfies the suspicion standard, and the affidavit and program log meet the individualization requirement. Because the program will search only ten dorm rooms, the impact on hassle rates will be very low. Whether the relevant jurisdiction is the greater Boston area, Cambridge, or even Harvard, the four or so searches of innocent dorm rooms will not pester the community with increased hassle rates.

This outcome is foreign to our current criminal-investigation system, but it is a significant improvement over business as usual. Randomly distributed²²⁶ hassle is preferable to the nonrandom distribution brought about by common police practices. Police are physically deployed in greater numbers to neighborhoods with low socioeconomic status (and reasonably so, since that is where they are most often needed).²²⁷ Because the traditional methods of building individualized suspicion rely on the perceptions of police officers, those methods are bound to target the poor disproportionately. A policeman must be physically present in order to provoke flight, observe a suspicious bulge, or craft some other tailor-made story about the suspect.

The police cannot craft that type of story about a randomly chosen Harvard dorm, but the random selection process achieves the ultimate goal of governmental constraint all the same.

Moreover, randomization has the potential to increase public trust in the fairness of criminal investigations. People are quite sensitive to their treatment relative to the treatment of others. Slobogin asked a sample of potential jurors in Gainesville, Florida, to score the intrusiveness of twenty-five different criminal-investigation techniques. Roadblocks earned the lowest average intrusiveness score, lower even than governmental access to anonymous phone and credit-card records.²²⁸ This is consistent with the Supreme Court's rationale for allowing the government to set up DUI checkpoints and airport security screens under the special-needs doctrine despite a complete lack of suspicion.²²⁹ The Court identified the evenhandedness of

225. Two scholars would *not* be appalled. Harcourt and Meares advocate for increased use of randomness in order to achieve the evenhanded outcomes that I describe here. Harcourt & Meares, *supra* note 19, at 815. Their work was tremendously influential on my thinking about this hypothetical. But Harcourt and Meares argue that courts have applied "individualized suspicion" as a single determination, whereas I think it is clear that courts and scholars apply the concepts separately and use "suspicion" exactly as Harcourt and Meares suggest. It is *individualization* that gets in the way of Harcourt and Meares's very thoughtful and appealing proposal.

226. See generally *id.* at 853.

227. Stuntz, *supra* note 6, at 2033; William J. Stuntz, *Law and Disorder: The Case for a Police Surge*, WKLY. STANDARD, Feb. 23, 2009, at 19.

228. Slobogin, *supra* note 127, at 335.

229. Mich. Dep't of State Police v. Sitz, 496 U.S. 444 (1990); United States v. Edwards, 498 F.2d 496 (2d Cir. 1974) (permitting airport security screens); United States v. Davis, 482 F.2d 893 (9th Cir. 1973) (same).

a checkpoint as one of its chief virtues.²³⁰ Harcourt and Meares also champion the benefits of checkpoints and random searches.²³¹ The indignity of a search is mitigated when we are all subjected to it and therefore cannot ascribe shame, fault, or prejudice to the experience.

The Supreme Court has not explicitly endorsed the use of randomness outside the special-needs doctrine. And the Court has twice rejected random searches when they were conducted without any suspicion at all—first in the context of random home searches to ensure compliance with fire codes²³² and then in the context of a vehicle checkpoint designed for general criminal law enforcement.²³³ But those cases do not foreclose the randomness proposed here. The randomly selected Harvard dorm search is *not* suspicionless.

As long as law-enforcement officers can meet the probable cause or reasonable suspicion standard, they should be permitted and encouraged to curb the impact of their searches through random selection.

D. *Objections*

This Article has urged a bold change in our approach to individualization. It has encouraged the use of data and randomness. Protests are inevitable.

Tribe is skeptical of replacing “intuitive tools” of legal reasoning with mathematical ones because people can be so mesmerized by the precision of quantified factors that they lose track of the soft, unquantified factors. Soft factors can contribute to the best estimates of guilt or fault. They may not be counted, but they count.²³⁴ For example, Tribe worries that a jury may be so taken by expert testimony on the high likelihood that a criminal defendant’s palm print appears on a knife used in a murder that it will forget to consider and account for the possibility that the defendant may have been framed.²³⁵

230. *Illinois v. Lidster*, 540 U.S. 419, 428 (2004) (emphasizing that police stopped vehicles *systematically* and that this reduces anxiety and alarm); *Sitz*, 496 U.S. at 451–52; *United States v. Martinez-Fuerte*, 428 U.S. 543, 559 (1976) (emphasizing that checkpoints involve less discretion).

231. Harcourt & Meares, *supra* note 19, at 851.

232. *Camara v. Mun. Court*, 387 U.S. 523, 528 (1967) (finding unconstitutional “arbitrary invasions by government officials”).

233. *City of Indianapolis v. Edmond*, 531 U.S. 32 (2000). The Court’s decision in *Edmond* may be regrettable considering the benefits of random searches. See Harcourt & Meares, *supra* note 19 (advocating a switch from suspicion-driven policing to random checkpoints).

234. Tribe, *supra* note 5, at 1333–34 (recounting that the attorneys litigating the Dreyfus affair did not understand a word of an expert witness’s mathematics-driven testimony but were nevertheless impressed by it). In modern litigation, we have seen a different problem emerge. Because both sides hire an expert witness, jurors are confronted with a battle of expertise that they have little chance of sorting out. See, e.g., Christopher Tarver Robertson, *Blind Expertise*, 85 N.Y.U. L. REV. 174, 177 (2010).

235. Tribe, *supra* note 5, at 1362–63.

Harcourt picks up these themes in his book *Against Prediction*.²³⁶ He also warns that police may become blind to context and inappropriately rely on quantification tools that stop predicting well over time.²³⁷

These criticisms are poorly suited to police stops and searches for at least four reasons. The first, identified by Minzner, is that the truth comes out.²³⁸ Unlike the decision to convict, the decision to search a suspect is immediately tested and either vindicated or repudiated by the result of the search. By the time an officer completes his stop or search, he knows whether he has uncovered evidence of a crime. By contrast, when a jury convicts a defendant, its decisions will rarely be proven wrong or right (except in the case of DNA exonerations²³⁹). The on-the-ground results of stops and searches help counteract whatever cognitive pull the *ex ante* statistics may have.²⁴⁰

Second, even if data-driven tools can and will be misused, they should not be compared to perfection. The tendency for misuse and abuse surely must be counted among a technique's costs.²⁴¹ But Tribe and Harcourt do not compare misusing statistics to misusing the "intuitive tools" traditionally used by police departments. The internal, nonspecified models that guide human intuition about whether to stop, search, or arrest a person have a pitiful track record. Eyewitness accounts are biased and unreliable.²⁴² Tips are often fruitless and vindictive.²⁴³ And the hunches of police officers vary wildly in their success.²⁴⁴ The comparative misuse costs of a mechanical system of criminal profiling and the intuitive systems used today may suggest that we need to discard the intuitive tools, not the mechanical ones.²⁴⁵

Moreover, misunderstanding and misuse of statistical tools will diminish over time. As the world becomes increasingly Moneyballed,²⁴⁶ judges,

236. Harcourt, *supra* note 5.

237. Harcourt argues that relying on a profile might be counterproductive if the population of would-be criminals learns of the profile or, in any event, discovers that their crimes are likely to be detected. *Id.*

238. Minzner, *supra* note 80, at 958.

239. See, e.g., NAT'L REGISTRY EXONERATIONS, <http://www.law.umich.edu/special/exoneration/> (last visited Aug. 27, 2014).

240. Also, the concepts of "probable cause" and "reasonable suspicion" may be the types of issues that are mathematical in nature and require explicit use of mathematical assumptions. See Tribe, *supra* note 5, at 1338; see also Cheng, *supra* note 27, at 397.

241. Tribe counts misuses of a technique among its flaws, since "the costs of abusing a technique must be reckoned among the costs of using it at all." Tribe, *supra* note 5, at 1331.

242. See Samuel R. Gross et al., *Exonerations in the United States, 1989 Through 2003*, 95 J. CRIM. L. & CRIMINOLOGY 523, 542 (2005).

243. See Alexandra Natapoff, *Beyond Unreliable: How Snitches Contribute to Wrongful Convictions*, 37 GOLDEN GATE U. L. REV. 107, 107–08 (2006).

244. Minzner, *supra* note 80, at 930–33.

245. Sentimental attachment to human judgment is also driving the debate on airplane autopilot programs. See Thom Patterson, *Who's Really Flying the Plane?*, CNN (Mar. 26, 2012, 8:15 AM EDT), <http://www.cnn.com/2012/03/24/travel/autopilot-airlines/index.html>.

246. See generally MICHAEL LEWIS, *MONEYBALL* (2003) (describing Major League Baseball teams' use of statistics to increase team performance and productivity).

lawyers, police, and even jurors are less likely to be entranced by experts.²⁴⁷ As Faigman puts it, “Science and technology today are so pervasive that the Court cannot continue its slapdash ways.”²⁴⁸

Finally, and most crucially, there is no principled difference between the quantified, “actuarial” factors that Tribe and Harcourt criticize and the “soft” or “intuitive” factors that they champion. The intuitive factors are equally probabilistic.²⁴⁹ When a juror or policeman rejects a statistical prediction of guilt because of an intuition, he is not rejecting prediction at all. He is rejecting the particular model in favor of a different one.²⁵⁰ His chosen model can go awry just as easily as any mathematical model. And since such a model operates inside the mind of the policeman or judge, it is less transparent and responsive to challenge.²⁵¹ Faith in intuitive policing places undue value on the imprecision that usually accompanies human discretion.²⁵²

CONCLUSION

Individualized suspicion has baffled courts and scholars for the last thirty years.²⁵³ While everyone seems to agree about what “suspicion” requires, “individualization” has perplexed the academy and muddled the case law. Fearful of the effects of large-scale generalizations, courts and scholars have shaped individualization to center on the suspect. They have insisted that police tell elaborate stories about the suspect’s distinct characteristics

247. This is not to say that lawyers and others will cease making mistakes. Indeed, even Tribe, whose article is overall quite careful, makes some obvious errors in probability theory. See Tribe, *supra* note 5, at 1336, 1355 (incorrectly describing the probabilities that somebody, or many people, will share a particular characteristic without using the pigeonholing principle). As the average American grows more comfortable with statistics, however, people will know enough to challenge and attack an expert’s declaration. At the very least, they will not feel intimidated.

248. DAVID L. FAIGMAN, LABORATORY OF JUSTICE 364 (2004). He continues as follows:

The Constitution’s framers were products of the Enlightenment. They expressly sought to bring the science of their time into the document that would govern the times to come. Subsequent generations of lawyers and judges have failed to carry forth this mandate. This state of affairs may change with the next generation of lawyers. It is not that lawyers and judges will suddenly volunteer to join a new enlightenment. They will be forced to do so.

Id.

249. See *supra* Section I.A (discussing the snowflake theory and its logical flaws).

250. At times, Tribe’s article acknowledges that human intuition is just a set of adjustments to an a priori model. Tribe, *supra* note 5, at 1350. For information on the cognitive biases and heuristics at play in human judgment, see generally Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124 (1974).

251. Indeed, intuition resists “articulation,” which, at least at times, is critical to Fourth Amendment requirements of suspicion. See, e.g., *Terry v. Ohio*, 392 U.S. 1, 21 (1968) (requiring articulable suspicion for a stop).

252. Schauer offers this same critique of Justice Marshall’s dissent in *United States v. Sokolow*. SCHAUER, *supra* note 15, at 173.

253. Cf. Harcourt & Meares, *supra* note 19, at 843 tbls.6 & 7 (showing a graph of the increasing usage of the term “individualized suspicion” in federal and state cases).

and behaviors, and they have trained the police to pack their affidavits with details that add lots of nuance but little substance.

Their mistake is in focusing on the particular suspect. That suspect is sufficiently protected by the suspicion requirement—the requirement that the police have reason to believe that a search has a fair probability of uncovering evidence. Instead, courts and scholars should be concerned about *everyone* who may fit the same suspicious facts and may therefore become subject to a search. If the profile casts too wide a net, the police are bound to cause a lot of hassle.

The goal of individualization is to control hassle. If suspicion protects the suspect, individualization protects everyone else. It protects the public from sweeping investigation practices that meet the relevant suspicion standard but impose too much hassle on the innocent.

