

GETTING BEYOND INTUITION IN THE PROBABLE CAUSE
INQUIRY

by
Erica Goldberg*

Courts are proudly resigned to the fact that the probable cause inquiry is “nontechnical.” In order to conduct a search or make an arrest, police need to satisfy the probable cause standard, which the Supreme Court has deemed “incapable of precise definition or quantification into percentages.” The flexibility of this elusive standard enables courts to defer to police officers’ reasonable judgments and expert intuitions in unique situations. However, police officers are increasingly using investigative techniques that replace their own observational skills with test results from some other source, such as drug sniffing dogs, facial recognition technology, and DNA matching. The reliability of such practices can and should be quantified, but the vagueness of the probable cause standard renders it impossible for judges to determine which error rates are inconsistent with probable cause.

This Article confronts the intersection between quantifiable evidence and the relentlessly fuzzy probable cause standard. It proposes that the probable cause standard be assigned a numerical value as a minimum threshold, especially in cases where probable cause is based on mechanistic techniques that essentially replace a police officer’s own judgment. The Article begins by exploring how the police and courts currently apply the probable cause standard, including courts’ confrontations with probabilities. It then explains why certain evidence should require quantified error rates to establish probable cause and how to properly calculate these error rates. In the final section, the Article argues that assigning a minimum percentage to probable cause would add much-needed clarity to the law and protect against systemic abuses.

I.	INTRODUCTION.....	790
II.	THE ELUSIVE PROBABLE CAUSE INQUIRY	795
	A. <i>Probable Cause as the Referee Between Privacy and Security</i>	795
	B. <i>The Role of Probable Cause When Less Suspicion Is Required</i>	797
	C. <i>Uncertainty in the Probable Cause Standard Throughout the Stages of Review</i>	799
	D. <i>Confrontations with Probability</i>	805

* I am grateful to Colin Miller, Jacob Kreutzer, and Bryan Kreykes for reading drafts of this Article. I also wish to thank David Kaye for his helpful comments throughout the writing process, and for his keen editing.

	1. <i>Background Attributes and Group Statistics</i>	806
	2. <i>Error Rates and Individualized Suspicion</i>	808
III.	QUANTIFYING EVIDENCE RELEVANT TO PROBABLE CAUSE.....	810
	A. <i>Requiring Error Rates</i>	813
	1. <i>Cases Where Quantified Error Rates Are Not Necessary</i>	813
	2. <i>Cases Where Quantified Error Rates Are Necessary</i>	814
	B. <i>Calculating Error Rates</i>	815
	1. <i>False Positive Rates</i>	816
	2. <i>Accuracy Rates</i>	819
	C. <i>Problems with Quantifying Error Rates</i>	821
IV.	QUANTIFIABLE EVIDENCE AND THE PROBABLE CAUSE	
	STANDARD	826
	A. <i>Quantifying Probable Cause</i>	827
	1. <i>Cases Where Quantifiable Evidence Is the Critical Fact in</i> <i>Satisfying Probable Cause</i>	828
	2. <i>When Quantifiable Evidence Is Presented Alongside Other</i> <i>Factors</i>	832
	3. <i>When All of the Evidence Is Qualitative</i>	833
	B. <i>Selecting the Number</i>	834
	C. <i>Other Solutions</i>	837
V.	CONCLUSION	837

I. INTRODUCTION

Courts are proudly resigned to the fact that the probable cause inquiry is “nontechnical.”¹ To conduct a search or make an arrest, a police officer must establish probable cause based on the totality of the circumstances. This approach allows for great flexibility in the application of facts to the standard. Adding uncertainty to this flexibility, the actual legal standard of probable cause remains undefined. The Supreme Court has deemed probable cause “incapable of precise definition or quantification into percentages”² and, just this year, overturned the Florida Supreme Court’s efforts to add a more rigorous framework to the inquiry.³ Judges, prosecutors, and scholars display varying understandings as to the degree of suspicion that probable cause requires.⁴ The deferential standard that reviewing courts apply to probable cause determinations further exacerbates the confusion.

The murkiness of the probable cause standard has many virtues, primarily in situations requiring the appraisal of many different facts, some tending towards suspicion and some undermining it. A flexible standard allows police to rely on their expertise, intuition, and

¹ *Maryland v. Pringle*, 540 U.S. 366, 370 (2003) (quoting *Illinois v. Gates*, 462 U.S. 213, 231 (1983)).

² *Id.* at 371.

³ *Florida v. Harris*, 133 S. Ct. 1050, 1053 (2013).

⁴ *See infra* Part II.C.

observational skills to decide whether suspicious behavior warrants further action, without the constraints of a rigid test. In complex cases, quantification of the standard by assigning it a numerical percentage would obscure the true, qualitative nature of the inquiry and create a false sense of precision.⁵ Plus, judges, concerned about institutional incompetence, may simply be “afraid of math.”⁶

There is a subset of cases, however, where the virtues of the imprecise definition of probable cause all but disappear. Consider a real case where Moose the drug detection dog, undergoing a training exercise at a storage facility, alerted to the presence of drugs in one of the storage bins where the police had not planted drugs.⁷ The police sought to obtain a warrant to search the locker based solely on Moose’s alert and the fact that the lessee of the storage bin, whose name was given to police after Moose’s alert, had a prior record of drug convictions.⁸ Whether or not probable cause exists to issue the warrant depends largely on Moose’s reliability, which can be quantified based on Moose’s error rate in detecting drugs.

There are an increasing number of situations, like the one involving Moose, where the police rely on machines or tools (such as a dog) to create their suspicion.⁹ As a result, the likelihood of criminal activity can be quantified. In such situations, the virtues of having an undefined probable cause standard are outweighed by its vices, including the lack of uniformity in application and susceptibility to abuse. In cases where the critical evidence at issue is produced through a mechanistic process, where police essentially outsource their intuition, this Article argues that reasonable officers cannot differ greatly on the strength of the evidence, and the arguments opposing quantifying the probable cause standard are not compelling. Further, even in cases where the evidence demonstrating probable cause cannot be quantified, having a well-defined legal standard does not detract from the discretion and flexibility the police would have in assessing the totality of the circumstances of the facts of a case.

⁵ Orin Kerr, *Why Courts Should Not Quantify Probable Cause*, in *THE POLITICAL HEART OF CRIMINAL PROCEDURE: ESSAYS ON THEMES OF WILLIAM J. STUNTZ* 131, 131–32 (Michael Klarman et al. eds., 2012) (discussing how quantifying probable cause would lead to less accurate probable cause determinations by undermining a judge’s logical intuition).

⁶ *Id.* at 132 (joking that this is one reason for the refusal to quantify the probable cause standard). If judges are afraid of math, this is perhaps rightfully so. See *infra* Part III for a discussion of the difficulties involved in incorporating statistical evidence into the probable cause inquiry.

⁷ *United States v. Cook*, 904 F.2d 37, *1–2 (6th Cir. 1990) (No. 89-5947) (unpublished disposition).

⁸ *Id.*

⁹ Dogs are considered tools like other machines used by the police or forensic scientists. See Colin Miller, *Hear Spot Bark: Washington Court of Appeals Rejects Argument that Dog Barking Constitutes Hearsay*, EVIDENCEPROF BLOG (Nov. 15, 2007), <http://lawprofessors.typepad.com/evidenceprof/2007/11/hear-spot-bark.html>.

Currently, the probable cause inquiry provides room for reasonable officers to differ on the strength of their suspicion. Courts making probable cause assessments consider whether the facts and circumstances warrant a person of “reasonable caution in the belief that’ an offense has been or is being committed.”¹⁰ When a magistrate judge undertakes the probable cause inquiry to determine whether a warrant should issue, she must decide only if there is a “fair probability that contraband or evidence of a crime will be found in a particular place.”¹¹ Later, if contraband or evidence of a crime is found, the reviewing court at the criminal trial need only “ensure that the magistrate had a ‘substantial basis for . . . conclud[ing]’ that probable cause existed.”¹²

At these increasingly deferential levels of inquiry, it is unclear what degree of suspicion is required before a magistrate can ratify an officer’s judgment of probable cause.¹³ Does a reasonable police officer have to “believe” by greater than 50% that a crime is being committed? And, do we look at whether *any* reasonable officer could believe by greater than 50% in the existence of criminal activity? Then, in defining a “fair probability,” should magistrate judges allow some band of error around their own determination of 50% likelihood to provide room for reasonable police officers to differ? If a magistrate judge himself deems that there is an approximately 35% chance that evidence could be found in a home, but appreciates that reasonable officers could believe the chance to be 60%, should the warrant issue? Further, does the substantial basis review of a magistrate’s decision place an even broader band around a police officer’s greater than 50% belief that a crime is occurring so as to allow for reasonable magistrate judges to differ?

Courts have not confronted these questions. Yet, as technology progresses, more tools are available that replace a police officer’s intuition and judgment with mechanistic findings, where the degree of suspicion can and should be quantified.¹⁴ Courts have determined, for example, that both positive alerts from drug sniffing dogs and fingerprint matches are sufficient on their own, without any other evidence, to satisfy probable cause.¹⁵ To establish that a police officer reasonably relied on

¹⁰ *Brinegar v. United States*, 338 U.S. 160, 175–76 (1949) (quoting *Carroll v. United States*, 267 U.S. 132, 162 (1925)).

¹¹ *Illinois v. Gates*, 462 U.S. 213, 238 (1983).

¹² *Id.* at 238–39 (quoting *Jones v. United States*, 362 U.S. 257, 271 (1960)). Not all courts fashion these standards the same way. See *Cook*, 904 F.2d 37, *2; *infra* Part II.C.

¹³ See *infra* Part II.C.

¹⁴ According to one scholar, “just on the horizon” are “handheld gun detectors, software protocols that sift through e-mails searching for illegal material, or facial recognition technology.” Ric Simmons, *The Two Unanswered Questions of Illinois v. Caballes: How To Make the World Safe for Binary Searches*, 80 TUL. L. REV. 411, 413 (2005).

¹⁵ See, e.g., *Florida v. Harris*, 133 S. Ct. 1050, 1057 (2013) (“If a bona fide organization has certified a dog after testing his reliability in a controlled setting, a court can presume (subject to any conflicting evidence offered) that the dog’s alert

these types of techniques, quantifiable error rates are often incorporated into the probable cause inquiry. Some courts have even held that certain investigative techniques must be assessed using quantified error rates.¹⁶

Once the reliability of practices like the use of drug sniffing dogs, fingerprint analysis, partial DNA matching,¹⁷ or facial recognition technology is quantified, however, the vagueness of the probable cause standard renders it impossible to determine which, if any, error rates are inconsistent with probable cause. Police cannot predict whether a particular investigative technique is reliable enough on its own to satisfy the probable cause inquiry. Judges do not know if probable cause requires a 40%, 50%, or 51% likelihood that a suspect has committed or will commit a particular crime.¹⁸ Our privacy rights vary not only because judges and police officers assess facts differently, an unavoidable and often beneficial consequence of a system administered by human beings, but also because they are using different legal standards. This creates a system prone to inconsistency and vulnerable to abuse.

provides probable cause to search.”); *United States v. Scarborough*, 128 F.3d 1373, 1378 (10th Cir. 1997) (“It is undisputed that a drug sniffing dog’s detection of contraband in itself establishes probable cause for a search warrant.”); *United States v. Waltzer*, 682 F.2d 370, 372 (2d Cir. 1982) (holding that a drug sniffing dog’s positive alert to the smell of contraband in luggage “itself establish[es] probable cause, enough for the arrest, more than enough for the stop”); *People v. Moon*, No. 1-09-3106, 2011 Ill. App. Unpub. LEXIS 745, at *7 (Ill. App. Ct. Apr. 26, 2011) (“The evidence at trial established that the police had probable cause to arrest defendant when they determined that the kitchen window was the point of entry, gathered finger and palm prints from the outside of the kitchen window that was behind a storm window, determined that the prints belonged to defendant and ascertained from Turner that she did not know defendant or give him permission to enter her residence.”).

¹⁶ Recently, the Florida Supreme Court held that a positive alert by a certified drug sniffing dog does not presumptively establish probable cause to conduct a search, and accuracy rates are needed to assess a dog’s reliability. *Harris v. State*, 71 So. 3d 756, 758–59 (Fla. 2011). This decision was reversed by the Supreme Court in an opinion reaffirming the flexible, “totality of the circumstances” test applied to all probable cause determinations. *See Harris*, 133 S. Ct. at 1055. The Court’s unanimous opinion, which presumes that sufficient training and certification render a dog sniff reliable enough to establish probable cause, *id.* at 1056–59, fundamentally misunderstands the problems inherent in relying on false positive rates and recognizing alerts of “residual odors” as correct alerts. *See infra* Part III.C. *Cf. State v. Anez*, 738 N.E.2d 491, 497 (Ohio Com. Pl. 2000) (excluding evidence of a field breath sobriety test because “the results of a scientific test may not be introduced unless there is foundational testimony that the device used was operating with a sufficient degree of accuracy”).

¹⁷ DNA matching is most relevant to the probable cause inquiry in the context of partial matching of DNA found at a crime scene to DNA in a databank, thus implicating possible family members of the person whose DNA has been stored in the databank. *See infra* Part III.B; *see generally* Erin Murphy, *Relative Doubt: Familial Searches of DNA Databases*, 109 MICH. L. REV. 291 (2010).

¹⁸ Some judges have held that probable cause is satisfied even when error rates were between 40% and 50%. *United States v. Donnelly*, 475 F.3d 946, 955 (8th Cir. 2007) (probable cause established despite dog’s error rate close to 50%).

This Article argues that when quantifiable evidence is the primary basis for an officer's suspicion, that evidence should be held to a minimum threshold of required certainty. Although scholars have touched upon whether fallible tools like drug sniffing dogs and facial recognition technology can yield evidence that satisfies the probable cause standard¹⁹ and have debated the appropriate level of flexibility with which probable cause should be defined,²⁰ this Article confronts the intersection of probabilistic evidence and the probable cause standard and proposes assigning a numerical value to the probable cause standard in cases where error rates are largely dispositive of the probable cause inquiry.

Part II of the Article details how the police and courts currently apply the concept of probable cause, and how the increasingly deferential standard of review used to assess probable cause contributes to the standard's elusiveness. Part II also explores how probabilistic evidence is currently incorporated into the probable cause inquiry. Part III contends that certain evidence should demand quantified error rates to establish probable cause. This section argues that statistical evidence is necessary to establish probable cause for investigative techniques that essentially replace a police officer's own judgment, skill, and intuition. Part III further analyzes how to calculate error rates in a way most relevant to the probable cause inquiry and exposes the limitations of quantification of evidence.

In the final section, the Article addresses how courts might incorporate quantifiable evidence into the probable cause inquiry. In cases in which quantifiable evidence is the critical factor in satisfying probable cause, assigning a numerical value to probable cause would add uniformity to the law and would better protect Fourth Amendment interests. In cases where quantifiable evidence is presented along with other qualitative evidence tending towards suspicion, judges can

¹⁹ Douglas A. Fretty, *Face-Recognition Surveillance: A Moment of Truth for Fourth Amendment Rights in Public Places*, 16 VA. J.L. & TECH. 430, 459–62 (2011) (discussing the “tolerable error” in the context of drug sniffing dogs, field sobriety tests, and facial recognition technology); Richard E. Myers II, *Detector Dogs and Probable Cause*, 14 GEO. MASON. L. REV. 1, 12–18 (2006) (using Bayes’ Rule to demonstrate that a positive alert by a drug sniffing dog cannot be sufficient on its own to establish probable cause).

²⁰ *E.g.*, Bruce A. Antkowiak, *Saving Probable Cause*, 40 SUFFOLK U. L. REV. 569, 586 (2007) (“Mathematics . . . will fail us in our pursuit of a definition for probable cause, and, indeed, we should fear any attempt to co-opt mathematical concepts to solve the probable cause riddle.”); Ronald J. Bacigal, *Making the Right Gamble: The Odds on Probable Cause*, 74 MISS. L.J. 279, 338 (2004) (exploring the tension between the precision of mathematics and the appeal of intuition in defining the probable cause standard and creating a “tiered model of the levels of certainty required for searches and seizures”); Joseph D. Grano, *Probable Cause and Common Sense: A Reply to the Critics of Illinois v. Gates*, 17 U. MICH. J.L. REFORM 465, 469 (1984) (preferring a “commonsense approach to probable cause”); Kerr, *supra* note 5 (arguing that probable cause should not be quantified).

qualitatively assess any suspicion created by a police officer's observational skills, but will know whether or not the quantifiable evidence comes close to meeting the probable cause hurdle. Finally, when quantifiable evidence is not presented as part of the probable cause inquiry, judges can adhere to prior case law without quantifying the inquiry. While confronting the arguments against quantifying probable cause, the final section further offers insights into selecting a number that is easy to conceptualize and proposes alternative solutions to quantifying probable cause.

II. THE ELUSIVE PROBABLE CAUSE INQUIRY

"The overriding function of the Fourth Amendment is to protect personal privacy and dignity against unwarranted intrusion by the State."²¹ The probable cause standard is so significant because it serves as the gatekeeper between individuals and these unwarranted intrusions. Yet, the standard has not been defined with sufficient precision.²² An examination of the role that probable cause plays in regulating governmental intrusions of various types is first necessary to understand why probable cause requires flexibility in applying the facts but greater precision in defining the legal standard.

The standard's elusiveness is exacerbated by relatively recent decisions diminishing judicial review of probable cause determinations. Quantifying the standard is now a more urgently needed solution. Although the Supreme Court has held that the probable cause standard is incapable of quantification, courts already incorporate quantifiable evidence into the inquiry; they are just not doing so in a standardized way.

A. *Probable Cause as the Referee Between Privacy and Security*

In the Fourth Amendment context, probable cause is particularly important because it resolves particular issues in the ubiquitous tension between individual liberty and public security/crime prevention. Serious intrusions on privacy and liberty are unconstitutional when undertaken without probable cause, and the evidence obtained as a result of the intrusion is often suppressed at trial.²³ If probable cause exists, however, the need to protect the public and enforce the laws permits more serious intrusions.

Probable cause serves as the referee between privacy and security because the Fourth Amendment protects individuals only from

²¹ *Schmerber v. California*, 384 U.S. 757, 767 (1966).

²² *See infra* Part II.C.

²³ The exclusionary rule requires that evidence obtained in violation of the Fourth Amendment be excluded from admission at trial, although the exclusionary rule has several exceptions. *United States v. Pope*, 467 F.3d 912, 916 (5th Cir. 2006) (discussing the exclusionary rule and its "limits").

“*unreasonable* searches and seizures.”²⁴ In defining whether a search or arrest is reasonable, courts look to the degree of suspicion motivating the police and the degree of intrusion into the individual’s privacy or liberty. For full searches and arrests by the police, probable cause is usually the requisite level of suspicion.²⁵ To establish probable cause, police need not demonstrate a certainty that criminal activity is afoot, or that evidence will be found in a particular place, but only a “fair probability” of a successful search or seizure.²⁶ This fair probability remains a mystery, but courts have consistently held that probable cause “exists if a law enforcement official, on the basis of the totality of the circumstances, has sufficient knowledge or reasonably trustworthy information to justify a person of reasonable caution in believing that an offense has been or is being committed.”²⁷

The most serious intrusions into an individual’s privacy require not only probable cause but also the prior blessing of the courts. This court intervention provides extra protection against unreasonable searches because a neutral magistrate judge is deemed more capable of making a probable cause determination than a police officer, who has a stake in the outcome of each case.²⁸ Searches or arrests that occur inside an individual’s home require a magistrate judge to render the probable cause determination and issue a warrant prior to the search or seizure.²⁹

Other types of searches are considered intrusive enough to require probable cause, but exigencies or diminished expectations of privacy render it unnecessary for the police to first obtain a warrant. The “automobile exception” permits police to search a vehicle on a public

²⁴ U.S. CONST. amend. IV (emphasis added).

²⁵ See *United States v. \$149,442.43 in U.S. Currency*, 965 F.2d 868, 876 (10th Cir. 1992) (“The test for determining probable cause for forfeiture purposes is the same as applies in arrests, searches and seizures.”).

²⁶ See *Walczyk v. Rio*, 496 F.3d 139, 156–57 (2d Cir. 2007) (discussing the probable cause standard).

²⁷ *United States v. Gagnon*, 373 F.3d 230, 236 (2d Cir. 2004).

²⁸ See *Malley v. Briggs*, 475 U.S. 335, 353 (1986) (“By requiring that conclusions concerning probable cause . . . be drawn by a neutral and detached magistrate instead of being judged by the officer engaged in the often competitive enterprise of ferreting out crime, we minimize the risk of unreasonable assertions of executive authority.” (quoting *Johnson v. United States*, 333 U.S. 10, 14 (1948)) (citation omitted)).

²⁹ See *Kentucky v. King*, 131 S. Ct. 1849, 1865 (2011) (“In no quarter does the Fourth Amendment apply with greater force than in our homes, our most private space which, for centuries, has been regarded as ‘entitled to special protection.’” (quoting *Georgia v. Randolph*, 547 U.S. 103, 115 (2006))); *Payton v. New York*, 445 U.S. 573, 590 (1980) (“[T]he Fourth Amendment has drawn a firm line at the entrance to the house. Absent exigent circumstances, that threshold may not reasonably be crossed without a warrant.”). The warrant clause of the Fourth Amendment provides that “no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.” U.S. CONST. amend. IV.

road without a warrant, so long as probable cause is satisfied.³⁰ Arrests in public are permissible without a warrant.³¹ Police make these probable cause determinations themselves, but they are reviewed later if evidence found as a result of the search is used in a criminal trial.³²

B. The Role of Probable Cause When Less Suspicion Is Required

Although probable cause is the critical standard for the most intrusive searches, there are other categories of searches and seizures that are constitutionally permitted based on less suspicion or justification. Even in these contexts, however, the probable cause standard is important, either for setting the bar against which courts measure “reasonable suspicion,” the level of suspicion needed for less intrusive searches, or for preventing intrusions that are not considered searches from becoming full searches.

A “reasonable suspicion” justifies brief detentions and superficial pat-downs of individuals over their clothing,³³ searches of students in public schools,³⁴ and searches of parolees’ homes.³⁵ Because lesser degrees of suspicion are also considered “not ‘readily, or even usefully, reduced to a neat set of legal rules,’”³⁶ courts define reasonable suspicion in relation to probable cause. Reasonable suspicion requires some degree of certainty, which is less than probable cause, and police must articulate the grounds for that suspicion.³⁷

³⁰ *Maryland v. Dyson*, 527 U.S. 465, 466 (1999) (discussing the “automobile exception to the warrant requirement”).

³¹ *United States v. Garcia*, 179 F.3d 265, 268 (5th Cir. 1999) (“[L]aw enforcement officials may arrest an individual in a public place without a warrant if they have probable cause to believe that the individual committed a felony.”).

³² More information on how courts review probable cause determinations is presented *infra* Part II.C.

³³ *United States v. Madrid*, 30 F.3d 1269, 1275 (10th Cir. 1994) (discussing a stop permitted by the Supreme Court in *Terry v. Ohio*, 392 U.S. 1 (1968), which “involv[es] only a brief, non-intrusive detention and frisk for weapons when officers have a reasonable suspicion that the defendant has committed a crime or is about to do so”).

³⁴ *Reynolds v. City of Anchorage*, 379 F.3d 358, 370 (6th Cir. 2004) (“Under the ‘special needs’ doctrine, a search of a particular student, with the exception of drug testing, must be supported by reasonable suspicion.”).

³⁵ *United States v. Loney*, 331 F.3d 516, 520–21 (6th Cir. 2003) (“[I]t is now beyond question that a state statute survives Fourth Amendment scrutiny if it authorizes searches of parolees based on a reasonable suspicion that an individual is violating the terms or conditions of parole.”).

³⁶ *United States v. Sokolow*, 490 U.S. 1, 7 (1989) (quoting *Illinois v. Gates*, 462 U.S. 213, 232 (1983)). In a 1981 survey, 164 federal judges provided responses revealing that they quantified reasonable suspicion at an average certainty of 31%. See C.M.A. McCauliff, *Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?*, 35 VAND. L. REV. 1293, 1324–25, 1328 (1982).

³⁷ See *Sokolow*, 490 U.S. at 7 (holding that the certainty needed to satisfy reasonable suspicion “is considerably less than proof of wrongdoing by a preponderance of the evidence” and less than the “fair probability” needed to satisfy probable cause).

Other types of government intrusions, due to heightened governmental interests or diminished privacy interests, either require no individualized suspicion or are not even considered searches that implicate Fourth Amendment protections.³⁸ In those cases, the probable cause inquiry prevents the police from converting a superficial intrusion into a greater intrusion.

One category of non-search that has great import for the probable cause inquiry, because it yields results that are usually more susceptible to quantification, is the so-called “binary search.” A binary search reveals only the presence or absence of illegal substances.³⁹ The Supreme Court has held that individuals do not have a legitimate expectation of privacy in contraband, so intrusions that reveal only the presence or absence of drugs, if conducted in ways that do not seriously implicate privacy interests or physically intrude upon constitutionally protected areas, do not trigger Fourth Amendment scrutiny.⁴⁰ When an agent of the Drug Enforcement Administration tested a small amount of white powder for cocaine, for example, it was not considered a search or a seizure.⁴¹ The police may also use drug sniffing dogs to search for scents in public places, such as outside a vehicle during a lawful traffic stop⁴² or the exterior of luggage at an airport,⁴³ without triggering the Fourth Amendment.

The major intrusion to privacy from using binary technology occurs after the technology has produced a positive alert. Police either conduct a full search themselves or obtain a warrant from a magistrate judge on the basis of the positive alert, but probable cause must be satisfied. Thus, the Fourth Amendment Analysis requires the judge to determine whether the positive alert provides the police with probable cause to

In *Alabama v. White*, 496 U.S. 325, 330 (1990), the Supreme Court explained that reasonable suspicion can be established with information that is different in quantity and less reliable than that required to show probable cause.

³⁸ See *Vernonia Sch. Dist. 47J v. Acton*, 515 U.S. 646, 657 (1995) (upholding school’s suspicionless testing of student athletes); *Nat’l Treasury Emps. Union v. Von Raab*, 489 U.S. 656, 677 (1989) (upholding mandatory urinalysis tests of employees who work for the United States Customs Service).

³⁹ The term “binary” refers to the fact that the search reveals either a “yes” or a “no” answer to the question of whether illegal substances are present and does not reveal any other potentially personal information. *Simmons*, *supra* note 14, at 413.

⁴⁰ Binary searches may trigger Fourth Amendment scrutiny if performed in a physically intrusive way. See *Florida v. Jardines*, 133 S. Ct. 1409, 1415–17 (2013).

⁴¹ *United States v. Jacobsen*, 466 U.S. 109, 123 (1984) (“A chemical test that merely discloses whether or not a particular substance is cocaine does not compromise any legitimate interest in privacy.”).

⁴² *Illinois v. Caballes*, 543 U.S. 405, 409 (2005).

⁴³ *United States v. Place*, 462 U.S. 696, 707 (1983) (“[W]e conclude that the particular course of investigation that the agents intended to pursue here—exposure of respondent’s luggage, which was located in a public place, to a trained canine—did not constitute a ‘search’ within the meaning of the Fourth Amendment.”). Part of the rationale in *Place* was that a sniff of a closed suitcase did not reveal the contents of that suitcase, and was therefore minimally intrusive. *Id.* at 706–07.

search further. A “false positive”⁴⁴ alert to an area that does not contain drugs may ultimately reveal items and information in which individuals *do* have a legitimate expectation of privacy, and forces individuals to submit to prolonged interaction with the police.⁴⁵

The probable cause standard thus protects the public from serious intrusions upon privacy, sets a standard against which reasonable suspicion is measured, and prevents binary searches from automatically becoming full-scale searches. However, at all levels of review, the level of confidence needed to satisfy the probable cause standard is unknown, rendering it vulnerable to erosion and abuse.

C. *Uncertainty in the Probable Cause Standard Throughout the Stages of Review*

Currently, courts apply a “nontechnical conception”⁴⁶ of probable cause at all levels of review, allowing for flexibility in application of the standard but providing little guidance for police officers and magistrate judges. The result is, as one scholar noted, that “[a]lthough the Court has stressed the importance of a ‘single uniform standard’ of probable cause for criminal investigatory conduct, it has not defined that standard in a manner that is particularly illuminating to those charged with enforcing and interpreting the criminal law.”⁴⁷ Relatively recent developments in Fourth Amendment law have compounded this problem because they have undermined judicial review of probable cause determinations, making it even more necessary to clarify the standard.

As mentioned above, law enforcement officers must make practical commonsense judgments in areas of uncertainty, when they do not know if a suspect is actually committing a crime. Because police officers are not, as the Supreme Court often reminds us, “legal technicians,”⁴⁸ the probable cause standard must allow police officers to make educated guesses. If such a guess is reasonable but incorrect, and the police ultimately find either no evidence of the criminal activity they were looking for or evidence of entirely unrelated criminal activity,⁴⁹ then the

⁴⁴ A false positive occurs when the search technology indicates the presence of criminal activity incorrectly. For a more detailed explanation, *see infra* Part III.C.

⁴⁵ For this reason, Justice Souter dissented in *Caballes*, which held that the use of drug sniffing dogs at routine traffic stops is not a search. 543 U.S. at 411 (Souter, J., dissenting) (“The infallible dog, however, is a creature of legal fiction.”).

⁴⁶ *Brinegar v. United States*, 338 U.S. 160, 176 (1949).

⁴⁷ Barry Jeffrey Stern, *Warrants Without Probable Cause*, 59 BROOK. L. REV. 1385, 1436–37 n.172 (1994) (citation omitted). Professor Ronald Bacigal had described probable cause as a “fixed, but undefined, percentage.” Bacigal, *supra* note 20, at 324.

⁴⁸ *Illinois v. Gates*, 462 U.S. 213, 231 (1983).

⁴⁹ *See United States v. Biglow*, 562 F.3d 1272, 1281 (10th Cir. 2009) (“Probable cause ‘requires only a probability or substantial chance of criminal activity,’ rather than an actual showing of such activity.” (quoting *New York v. P.J. Video, Inc.*, 475 U.S. 868, 877–78 (1986))).

standard must also provide room for courts to defer to a police officer's expertise,⁵⁰ but prevent unreasonable intrusions.⁵¹

To accommodate the myriad interactions between police and individuals without creating rules that unduly stifle a police officer's exercise of her intuition, courts generally assess probable cause given the totality of the circumstances.⁵² This flexibility allows courts to take all of the facts into account and make an intuitive judgment when issuing a warrant or upholding a search instead of having to strictly adhere to rigid rules about what constitutes probable cause.⁵³ For example, the totality-of-the-circumstances test that governs whether an informant's tip can supply probable cause balances factors like the reliability of the informant, the basis for the informant's information, and the extent to which the police have corroborated the tip.⁵⁴ A judge may disregard the fact that a confidential informant's criminal record or drug addiction undermines her reliability if other factors point towards her truthfulness.⁵⁵ There are many ways to interpret the same set of facts, and an innocent explanation for the evidence presented does not necessarily negate probable cause.⁵⁶

With great flexibility, however, comes great uncertainty.⁵⁷ The Supreme Court has remarked that "[r]easonable minds frequently may differ on the question whether a particular [warrant] affidavit establishes

⁵⁰ *Ornelas v. United States*, 517 U.S. 690, 700 (1996) ("[O]ur cases have recognized that a police officer may draw inferences based on his own experience in deciding whether probable cause exists."). According to the Supreme Court in *Ornelas*, where a layperson might conclude that a "loose panel below the back seat armrest in [an] automobile . . . may suggest only wear and tear, . . . to [a police officer], who had searched roughly 2,000 cars for narcotics, it suggested that drugs may be sequestered inside the panel." *Id.*

⁵¹ See *supra* Part II.A.

⁵² See *Gates*, 462 U.S. at 238.

⁵³ Professor Ronald Bacigal noted that scholars opposed to prescribing precise mathematical rules to govern the probable cause inquiry "prefer to trust the subjective judgment of experienced decision makers who evaluate each individual situation in light of accumulated experience" and "rely on perceptions that they cannot articulate as explicit rules or mathematical probabilities." Bacigal, *supra* note 20, at 305.

⁵⁴ *Gates*, 462 U.S. at 230–34.

⁵⁵ *United States v. McKinney*, 919 F.2d 405, 415 (7th Cir. 1990) ("The magistrate was thus presented with specific details of a crime; neither Brown's drug addiction nor criminal record warranted disregarding her report. Although an informant's reliability is a factor to be considered by a court, it is just one relevant consideration in the totality of the circumstances analysis.").

⁵⁶ *United States v. Fama*, 758 F.2d 834, 838 (2d Cir. 1985).

⁵⁷ See Bacigal, *supra* note 20, at 339 (discussing the "unacceptable complexity of balancing the totality of the circumstances"). See generally Symeon C. Symeonides, *Choice of Law in the American Courts in 2009: Twenty-Third Annual Survey*, 58 AM. J. COMP. L. 227, 238 (2010) ("The tension between the need for legal certainty and predictability, on the one hand, and the desire for flexible, equitable, individualized solutions on the other, is as old as law itself.").

probable cause.”⁵⁸ Scholars have noted “wildly different outcomes” based on similar fact patterns when determining probable cause and reasonable suspicion.⁵⁹ Part of the problem is that no one knows how high a hurdle the standard actually presents. The Supreme Court explicitly refuses to assign probable cause a numerical value, equating it instead to a “fair probability” that evidence will be found.⁶⁰ Judges, scholars, and practitioners hold varying views as to the burden imposed by probable cause, with the largest number of judges clustering in the range between 30% and 60%.⁶¹ Disagreement among scholars and practitioners even exists as to whether probable cause is a lighter or equivalent burden to the preponderance of the evidence standard.⁶²

The recognition that reasonable minds may easily differ on whether probable cause is satisfied often makes reviewing courts loathe to second-guess probable cause determinations by either the police or magistrate

⁵⁸ United States v. Leon, 468 U.S. 897, 914 (1984).

⁵⁹ Margaret Raymond, *Down on the Corner, Out in the Street: Considering the Character of the Neighborhood in Evaluating Reasonable Suspicion*, 60 OHIO ST. L.J. 99, 121–22 (1999) (“In a high crime area, sitting in a car in a parking lot late at night may create reasonable suspicion [to detain someone] in Georgia, but not in Tennessee, and standing on a street corner may create reasonable suspicion in Louisiana, but not in Pennsylvania, even though these jurisdictions apply the same standard.” (footnotes omitted)).

⁶⁰ Illinois v. Gates, 462 U.S. 213, 235, 238 (1983); see Florida v. Harris, 133 S. Ct. 1050, 1055 (2013).

⁶¹ In the survey of federal judges in 1981, mentioned *supra* note 36, 166 provided responses revealing that they quantify “probable cause to believe,” on average, at 45.78%. Two judges even associated probable cause with 10% and one judge with 90%. See McCauliff, *supra* note 36, at 1324–25, 1327. Most of the judges believed probable cause was around 40% or 50%, but 52 of the judges thought probable cause was either around 30% or around 60%. *Id.* at 1327.

⁶² See, e.g., 2 WAYNE R. LAFAVE, SEARCH AND SEIZURE: A TREATISE ON THE FOURTH AMENDMENT § 3.2(e), at 60–64 (3d ed. 1996) (discussing cases and commentaries indicating that probable cause does not mean “more-probable-than-not”); John H. Blume et al., *Every Juror Wants a Story: Narrative Relevance, Third Party Guilt and the Right To Present a Defense*, 44 AM. CRIM. L. REV. 1069, 1104 n.240 (2007) (“Probable cause does not mean more probable than not.”); Christopher Slobogin, *Government Data Mining and the Fourth Amendment*, 75 U. CHI. L. REV. 317, 329 (2008) (describing probable cause as “a level of certainty akin to the civil preponderance standard”); Kimberly A. Burdge, Comment, *Legislating the Fourth Amendment: Can Congressional Legislation Make the Unconstitutional Legal?*, 50 HOW. L.J. 899, 901 (2007) (“[A] more detailed, modern version of probable cause would not necessarily require a more likely than not scenario.”). One particular federal prosecutor told his supervisee that probable cause meant 40%, Lawrence Rosenthal, *The Crime Drop and the Fourth Amendment: Toward an Empirical Jurisprudence of Search and Seizure*, 29 N.Y.U. REV. L. & SOC. CHANGE 641, 680 (2005), but an attorney for the FBI believed that “probable cause mean[t] . . . more likely than not, or if quantified, a 51% likelihood,” Daniel Richman, *Prosecutors and Their Agents, Agents and Their Prosecutors*, 103 COLUM. L. REV. 749, 783 (2003) (quoting Memorandum from Coleen Rowley to Robert Mueller, FBI Director (May 21, 2002)). A plurality of the Supreme Court has stated that the probable cause standard “does not demand any showing that such a belief [of criminal wrongdoing] be correct or more likely true than false.” Texas v. Brown, 460 U.S. 730, 742 (1983).

judges.⁶³ This has led to increasingly deferential review of probable cause decisions, which then in turn contributes to the elusiveness of the probable cause standard. For example, in the case of searches where a warrant is required, the police affidavit must contain information allowing a person of reasonable caution to believe that evidence will be found in the place to be searched.⁶⁴ Then, the magistrate judge must find that, “given all the circumstances set forth in the affidavit . . . there is a fair probability that contraband or evidence of a crime will be found in a particular place.”⁶⁵ Great deference is later given to the magistrate’s issuance of a warrant by a reviewing court,⁶⁶ which determines whether the magistrate’s decision is supported by “a substantial basis.”⁶⁷

This deference, combined with a vague standard, does not promote uniformity in magistrate judges’ decisions.⁶⁸ The deference might be appropriate if reviewing courts were simply deferring to a magistrate judge’s assessment as to whether the facts at issue, based on the totality of the circumstances, surpassed a known probable cause threshold. However, because the probable cause hurdle is so vague, reviewing courts cannot know the degree of suspicion on which the magistrate judge relied in issuing a warrant.⁶⁹ Is that magistrate basing his ruling on the

⁶³ See *United States v. Leon*, 468 U.S. 897, 914 (1984) (holding that a magistrate’s issuance of a warrant is entitled to “great deference” due to the strong preference for warrants and the reality of “doubtful or marginal case[s]” (quoting *Spinelli v. United States*, 393 U.S. 410, 419 (1969))); *United States v. Ventresca*, 380 U.S. 102, 106 (1965) (internal quotation marks omitted).

⁶⁴ See *United States v. Harris*, 403 U.S. 573, 593 (1971) (Harlan, J., dissenting).

⁶⁵ *Gates*, 462 U.S. at 238.

⁶⁶ See *United States v. Allen*, 625 F.3d 830, 840 (5th Cir. 2010) (“A magistrate’s determination of probable cause is entitled to great deference by reviewing courts.”).

⁶⁷ *Id.* In *Allen*, the court made clear that the duty of a reviewing court is simply to ensure that the magistrate had a “substantial basis for concluding that probable cause existed.” *Id.* at 834; see, e.g., *United States v. McClellan*, 350 F. App’x 767, 769 (3d Cir. 2009) (“Our review of the initial probable cause determination, however, is deferential as we scrutinize the affidavit to determine if ‘there is a substantial basis for a fair probability’ that the wiretaps would obtain communications regarding illegal drug transactions.” (quoting *United States v. Conley*, 4 F.3d 1200, 1205 (3d Cir. 1993))); *United States v. Reed*, 195 F. App’x 815, 822 (10th Cir. 2006) (“A reviewing court should accord great deference to a magistrate’s determination of probable cause; its duty is ‘simply to ensure that the magistrate had a substantial basis for . . . conclud[ing] that probable cause existed.’ This deference is appropriate to further the Fourth Amendment’s strong preference for warrants.” (quoting *Gates*, 462 U.S. at 238–39) (citation omitted)).

⁶⁸ The Supreme Court has warned against too much deference to a trial court’s probable cause determination, remarking that “[a] policy of sweeping deference would permit, ‘[i]n the absence of any significant difference in the facts,’ ‘the Fourth Amendment’s incidence [to] tur[n] on whether different trial judges draw general conclusions that the facts are sufficient or insufficient to constitute probable cause.’ Such varied results would be inconsistent with the idea of a unitary system of law.” *Ornelas v. United States*, 517 U.S. 690, 697 (1996) (quoting *Brinegar*, 338 U.S. 160, 171 (1949)) (citation omitted).

⁶⁹ In his excellent piece on the tension between a statistical versus an intuitive

fact that a reasonable officer has to believe that evidence will be found by greater than 50%, and a “fair probability” determination by the magistrate judge allows for some room for disagreement among reasonable police officers?⁷⁰ Reviewing courts, which decide whether a “substantial basis” supports the magistrate’s decision,⁷¹ do not even know how much suspicion the magistrate believed was required before rendering his decision.

The probable cause standard’s imprecision at various levels of review is also partially responsible for the “good-faith exception” to the exclusionary rule,⁷² which, in turn, contributes to the indeterminacy of the standard. Although the exclusionary rule requires the “fruits” of unlawful searches to be suppressed at trial, the good faith exception permits the admission of evidence found pursuant to a search warrant not supported by probable cause, so long as law enforcement reasonably relied on the warrant.⁷³ A reviewing court can therefore find under certain circumstances that even if a substantial basis did not exist for the magistrate’s determination, any evidence found may still form the basis of a criminal conviction.⁷⁴ The good faith exception means all evidence found pursuant to a warrant will be admitted unless the warrant is “so lacking in indicia of probable cause as to render official belief in its existence entirely unreasonable.”⁷⁵

The good faith exception thus almost entirely insulates a magistrate judge’s issuance of a warrant from review at the trial or appellate levels. Many reviewing courts simply find that the good faith exception applies

approach to probable cause, Ronald Bacigal argues that the Supreme Court has maintained the “fiction” that probable cause is a “single, familiar standard” that applies equally to every crime. Bacigal, *supra* note 20, at 323–24 (quoting *Dunaway v. New York*, 422 U.S. 200, 213 (1979)). According to Bacigal, “[t]his rigid view of probable cause requires the police to pinpoint their level of certainty to some fixed, but undefined, percentage.” *Id.*

⁷⁰ At least one court has even conflated the standards at various levels of review, holding that probable cause exists if a reasonable person could “believe there is a fair probability that contraband or evidence of a crime would be found in a particular place.” *United States v. Fladten*, 230 F.3d 1083, 1085 (8th Cir. 2000).

⁷¹ See *supra* Part I.

⁷² Police are permitted to rely on a search warrant, even if later found lacking in probable cause, where a search warrant affidavit “provided evidence sufficient to create disagreement among thoughtful and competent judges as to the existence of probable cause.” *United States v. Leon*, 468 U.S. 897, 926 (1984).

⁷³ The Supreme Court has held that the good faith exception does not apply to a search warrant if the police lied to procure the warrant, or the warrant was so lacking in probable cause that the magistrate judge essentially rubber-stamped a warrant request. See *Leon*, 468 U.S. at 921–24 (“Penalizing the officer for the magistrate’s error, rather than his own, cannot logically contribute to the deterrence of Fourth Amendment violations.”).

⁷⁴ *United States v. Bynum*, 293 F.3d 192, 195 (4th Cir. 2002) (allowing the admission of evidence even if “it is clear that . . . Agent Peterson’s affidavit does not provide a substantial basis for determining the existence of probable cause”).

⁷⁵ *Id.* (quoting *Leon*, 468 U.S. at 923) (internal quotation marks omitted).

without considering whether the warrant was supported by probable cause.⁷⁶ This stymies the development of the law surrounding the probable cause standard in favor of admission of a great deal more evidence.⁷⁷ Given that the review of a magistrate judge's probable cause determination is already so deferential, it is difficult to determine in a particular case if probable cause is actually lacking unless it is so obviously lacking that a court must exclude the evidence despite the good faith exception.

Although *Leon's* exception applies only in the warrant context, one scholar has found that searches conducted pursuant to warrants are much more likely to produce evidence than searches conducted where no warrant is required.⁷⁸ *Leon's* good faith exception, and its progeny,⁷⁹ may lead to an erosion of the extra protections that warrants offer.⁸⁰ The good faith exception focuses on what a reasonable officer would believe, not whether the probable cause standard was actually met based on the warrant application,⁸¹ thereby bypassing the magistrate's oversight as intermediary between law enforcement's own determination of probable cause and the resulting search.⁸²

Unscrupulous or overextended police officers and magistrate judges can easily exploit the uncertainty in the probable cause standard,

⁷⁶ *E.g.*, *United States v. Houston*, 665 F.3d 991, 994 (8th Cir. 2012) ("A court may consider whether the good-faith exception applies before conducting probable cause analysis.").

⁷⁷ Indeed, as the dissent in *Leon* predicted, "After today's decisions, there will be little reason for reviewing courts to conduct such a conscientious review [of the probable cause determination]; rather, these courts will be more likely to focus simply on the question of police good faith." *Leon*, 468 U.S. at 956–57 n.15 (Brennan, J., dissenting).

⁷⁸ One scholar examined available empirical data to conclude that searches pursuant to warrants recover evidence at rates that may exceed 80%. The recovery rate on warrantless searches, even when police allege probable cause, is significantly less, perhaps as low as 12%. Max Minzner, *Putting Probability Back Into Probable Cause*, 87 TEX. L. REV. 913, 923–24 (2009).

⁷⁹ *E.g.*, *Herring v. United States*, 555 U.S. 135 (2009) (extending *Leon's* good faith exception to a situation where police negligence led to an arrest pursuant to a revoked warrant).

⁸⁰ *See, e.g.*, *Houston*, 665 F.3d at 995–96 (upholding police officers' reliance on a search warrant based on the "common sense appeal" of the inference that because defendant molested a child and looked at child pornography six years ago, he might have child pornography on his current computer at a different location (quoting *United States v. Perry*, 531 F.3d 662, 665 (8th Cir. 2008))).

⁸¹ In fact, in applying the good faith exception, some courts consider information known to the police officer that was inadvertently omitted from the warrant application. *See United States v. McKenzie-Gude*, 671 F.3d 452, 460–61 (4th Cir. 2011).

⁸² *See United States v. McKinney*, 919 F.2d 405, 413 (7th Cir. 1990) (discussing the "far less reliable procedure of an after-the-event justification for the arrest or search, too likely to be subtly influenced by the familiar shortcomings of hindsight judgment" (quoting *Beck v. Ohio*, 379 U.S. 89, 96 (1964) (internal quotation mark omitted))).

especially when combined with the exceptional deference offered to the initial probable cause determination.⁸³ Given the lack of oversight over probable cause decisions, the flexibility afforded by the current application of the probable cause standard may not always be a virtue, and may place too much discretion in the hands of police officers at the expense of privacy interests.⁸⁴ According to Professor Ronald Bacigal, “[t]he inability to formulate clear rules or precise probability levels governing probable cause has lead [sic] the Court to adopt one overarching rule for the police—just use your common sense and act reasonably.”⁸⁵ To some extent, this standard contravenes the purpose of the Fourth Amendment, which is to limit police discretion.⁸⁶ Although reasonableness is part of the standard, an undefined legal hurdle leads to variability in how much suspicion is deemed “reasonable,” allowing for perhaps unjustified amounts of police discretion.

While there is ample reason to allow great flexibility in application of the facts to the probable cause standard, there is less justification for preserving a vague, undefined legal standard. This is especially true because courts are already incorporating probabilities into the probable cause inquiry.

D. *Confrontations with Probability*

Part of the reason that courts are loathe to attach rules or numbers to the probable cause inquiry is a concern that a “probabilistic” approach to probable cause would replace detective work that yields individualized suspicion.⁸⁷ This type of probabilistic approach, some fear, would

⁸³ Although many scholars assume that most police officers would not abuse their power to establish probable cause, empirical evidence indicates a serious problem with perjury in warrant affidavits. See Stephen W. Gard, *Bearing False Witness: Perjured Affidavits and the Fourth Amendment*, 41 SUFFOLK U. L. REV. 445, 447–48 (2008). Clarifying the probable cause standard or applying a more stringent standard of review would not solve the problem of perjury, but it would impede police attempts to establish probable cause where none exists, which police might prefer to outright lying.

⁸⁴ Scholars have commented that the cynical explanation for flexible or inconsistent Fourth Amendment standards is to make it easier for police to conduct searches without worrying about constitutional protections. *E.g.*, Susan F. Mandiberg, *Reasonable Officers vs. Reasonable Lay Persons in the Supreme Court’s Miranda and Fourth Amendment Cases*, 14 LEWIS & CLARK L. REV. 1481, 1525 (2010) (“So why should police not be held to higher standards suitable to the trained professionals they presumably are? . . . The cynical answer is that lower courts would be more likely to find a Fourth Amendment or *Miranda* violation if police were held to a higher standard . . . than is currently the case.”).

⁸⁵ Bacigal, *supra* note 20, at 318.

⁸⁶ *Id.* at 320.

⁸⁷ See *id.* at 297 (“An unavoidable feature of probabilistic thinking is that it treats people as members of a group rather than as individuals. Some commentators suggest that this focus violates an ethical command to treat citizens as unique individuals and to judge them only on evidence about their own conduct or matters within their

penalize people based on group attributes instead of individual characteristics.⁸⁸ Yet, this is not the only way probabilities have been incorporated into the probable cause inquiry. Courts also review probabilities relevant to individualized suspicion, which is not as problematic as finding probable cause based on group statistics. These two types of probabilistic evidence—data that measures group attributes and data that gives rise to individualized suspicion—have different roles in the probable cause inquiry, and both illustrate how courts evaluate statistical evidence in assessing probable cause.

1. *Background Attributes and Group Statistics*

Many courts and scholars have expressed an aversion to using non-individualized statistical information to sanction intrusive police conduct or treat people as potential criminals.⁸⁹ For example, presence in a high-crime area is not sufficient by itself to justify stopping an individual on the theory that most people in that area are engaged in criminal activity.⁹⁰ Courts do not want to punish individuals for where they live, or for the demographic groups to which they belong.⁹¹ However, presence in a high-crime area can be considered as a relevant factor in determining whether the police have reasonable suspicion, a lower hurdle than probable cause,⁹² to briefly detain someone.⁹³

It is in such situations that courts consider group statistics on criminality within an area. Although one article has noted that most courts are content to rely on “an officer’s testimony that an area is a ‘high-crime area’ without much analysis as to the basis of that

own control.”); *see also* William J. Stuntz, *Local Policing After the Terror*, 111 *YALE L.J.* 2137, 2164 (2002) (“Current Fourth Amendment law discourages group seizures. . . . Aggregate justifications—[where] no one person is reasonably suspected, but the odds are high that some members of the group are criminals—do not suffice.”).

⁸⁸ *See* Bacigal, *supra* note 20, at 295–97; *see also* David A. Harris, *Using Race or Ethnicity as a Factor in Assessing the Reasonableness of Fourth Amendment Activity: Description, Yes; Prediction, No*, 73 *MISS. L.J.* 423, 451 (2003) (“Group guilt can be both immensely powerful and deeply troubling when used to punish people or focus suspicion on them merely for their associations.”).

⁸⁹ *E.g.*, *United States v. Montero-Camargo*, 208 F.3d 1122, 1129–30, 1138 (9th Cir. 2000); *see* Raymond, *supra* note 59, at 105–06 (discussing a hypothetical where one in three individuals in a certain area is likely to possess drugs and arguing that, although this probability likely satisfies reasonable suspicion, “reasonable suspicion would not exist as to each individual in the neighborhood”). According to Raymond, “[s]ome particularized observations—proof that implicates an identified individual—must also be offered in support of the claim of reasonable suspicion.” *Id.* at 106 (footnote omitted).

⁹⁰ *Brown v. Texas*, 443 U.S. 47, 52 (1979) (“The fact that appellant was in a neighborhood frequented by drug users, standing alone, is not a basis for concluding that appellant himself was engaged in criminal conduct.”).

⁹¹ *See generally* Raymond, *supra* note 59.

⁹² *See supra* Part II.B.

⁹³ *See Illinois v. Wardlow*, 528 U.S. 119, 124 (2000) (“[O]fficers are not required to ignore the relevant characteristics of a location in determining whether the circumstances are sufficiently suspicious to warrant further investigation.”).

conclusion,”⁹⁴ some courts have required prosecutors to provide quantitative evidence demonstrating the existence of a high-crime area.⁹⁵ Courts must then review statistical evidence on whether a particular location is either a high-crime area or undergoing a crime wave with respect to the criminal activity at issue.⁹⁶

Courts have also examined group statistics when assessing the efficacy of random searches and seizures that do not require a showing of probable cause, to determine whether these intrusions are reasonable. Sobriety roadblocks, airport screening programs, and the practice of ordering motorists to exit their vehicle during traffic stops have all been assessed using percentages of criminal activity found per search or seizure.⁹⁷ In *Delaware v. Prouse*,⁹⁸ the Supreme Court invalidated random stops made by the Delaware Highway Patrol in part because “no empirical evidence indicated that such stops would be an effective means of promoting roadway safety.”⁹⁹ Later, distinguishing *Prouse*, the Supreme Court upheld a sobriety checkpoint pilot program where “the detention of [each of] the 126 vehicles that entered the checkpoint resulted in the arrest of two drunken drivers. Stated as a percentage, approximately 1.6 percent of the drivers passing through the checkpoint were arrested for alcohol impairment.”¹⁰⁰

It is important to note that statistics that measure the likelihood of criminality within a certain group, like the ones mentioned above, are usually insufficient to satisfy probable cause, which is a much higher bar than reasonable suspicion. Satisfying probable cause almost always requires some individualized suspicion.¹⁰¹ As one court of appeals held,

⁹⁴ 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, 1607 (2008).

⁹⁵ *E.g.*, *United States v. Bonner*, 363 F.3d 213, 216 (3d Cir. 2004) (finding not “clearly erroneous” the district court’s determination that a location could not be categorized as a high-crime area where “an average of 1.3 arrests [were made] per week [near the place of the stop], and that most of the arrests were for misdemeanors and summary offenses” due to the large number of people who frequent the area); *United States v. Montero-Camargo*, 208 F.3d 1122, 1139 n.32 (9th Cir. 2000) (requiring district courts to “examine with care the specific data underlying any such assertion” of the existence of a high-crime area). For a detailed analysis of how courts have attempted to standardize quantification of statistics relevant to the existence of a high-crime area, see Ferguson & Bernache, *supra* note 94, at 1607–23.

⁹⁶ *See United States v. Wright*, 485 F.3d 45, 53–54 (1st Cir. 2007) (articulating the test for determining if a neighborhood’s crime profile can contribute to reasonable suspicion).

⁹⁷ Bacigal, *supra* note 20, at 311–14.

⁹⁸ 440 U.S. 648 (1979).

⁹⁹ *Mich. Dep’t of State Police v. Sitz*, 496 U.S. 444, 454 (1990) (distinguishing *Prouse*).

¹⁰⁰ *Id.* at 454–55. The Court also relied on data from other states that “on the whole, sobriety checkpoints resulted in drunken driving arrests of around 1 percent of all motorists stopped.” *Id.* at 455.

¹⁰¹ *Chandler v. Miller*, 520 U.S. 305, 313 (1997).

“[s]tanding alone, a high incidence of child molestation by persons convicted of child pornography crimes may not demonstrate that a child molester is likely to possess child pornography.”¹⁰²

2. Error Rates and Individualized Suspicion

When probable cause is the requisite standard, courts often assess statistical evidence to indicate a likelihood of criminality particular to an individual suspect. Devices such as drug sniffing dogs and fingerprint analysis are used by police to generate suspicion specific to an individual. Instead of using statistics to determine how likely it is that individuals in a group are criminals, statistical evidence here quantifies the error rates of investigative tools that indicate whether a particular individual is behaving criminally. The quantification of error rates for such devices, therefore, does not ascribe guilt to any class of individuals or otherwise violate the notion that probable cause must be based on individualized suspicion.

For example, quantitative data has been used to assess the reliability of tipsters in providing the police with accurate information about criminal activity, based on how many tips lead to the discovery of criminal activity.¹⁰³ In fact, some courts have held that “[w]hen an informant has a proven track record for providing reliable information, corroboration of the information he provides is not necessarily essential.”¹⁰⁴ However, courts have not quantified in a standardized way what constitutes a reliable track record, likely because the nature of the information matters more than the bare percentage of times an informant has given accurate information.¹⁰⁵

Statistical data is also used to bolster the demonstration of probable cause in cases where a binary search leads to a full search.¹⁰⁶ This is a particularly salient issue in the context of drug sniffing dogs, who are constitutionally permitted to sniff the outside of an individual’s vehicle or suitcase in a public place without any suspicion.¹⁰⁷ Courts must then determine if a drug sniffing dog’s positive alert provided the police with

¹⁰² *United States v. Hodson*, 543 F.3d 286, 293 (6th Cir. 2008) (quoting *United States v. Adkins*, 169 F. App’x 961, 967 (6th Cir. 2006)) (internal quotation marks omitted). As another example, courts have held that merely finding drugs in a defendant’s car does not lead to probable cause that he has drugs in his house. *E.g.*, *State v. Kahn*, 555 N.W.2d 15, 19 (Minn. Ct. App. 1996).

¹⁰³ *E.g.*, *United States v. Shepherd*, 714 F.2d 316, 317 (4th Cir. 1983) (“[O]n approximately thirty earlier occasions the informant had provided information that had led to over twenty-five convictions.”).

¹⁰⁴ *United States v. Smith*, 182 F.3d 473, 479 (6th Cir. 1999).

¹⁰⁵ *See United States v. Bynum*, 293 F.3d 192, 203 (4th Cir. 2002) (“The bar for establishing an informant’s veracity through references to his track record is not high. For example, courts have uniformly held that an informant’s veracity is adequately established when the affiant asserts that the informant has supplied information leading to arrests and convictions.”).

¹⁰⁶ *See supra* Part II.A.

¹⁰⁷ *Id.*

probable cause to search an area. Prior to *Florida v. Harris*,¹⁰⁸ as one scholar noted, the Supreme Court had “not yet addressed the specific qualifications that a ‘trained canine’ possesses . . . [or] questioned the accuracy record of a particular canine in practice. As a result, lower courts have attempted to fill the gap and determine when a canine alert may be accepted as reliable.”¹⁰⁹

In a decision that was recently reversed in *Harris*, the Florida Supreme Court held that “field performance records” and evidence regarding the dog handler’s training must be considered when a judge assesses whether a dog’s positive alert provided probable cause.¹¹⁰ Many courts across the country have also examined a drug sniffing dog’s “accuracy rate”¹¹¹ in determining whether the dog’s positive alert satisfied the probable cause standard.¹¹² However, some courts have held that accuracy rates are not essential to the probable cause determination.¹¹³ Police in those jurisdictions may simply demonstrate that the dog that

¹⁰⁸ 133 S. Ct. 1050 (2013).

¹⁰⁹ Robert C. Bird, *An Examination of the Training and Reliability of the Narcotics Detection Dog*, 85 Ky. L.J. 405, 417 (1996–1997).

¹¹⁰ *Harris v. State*, 71 So. 3d 756, 775 (Fla. 2011), *overruled by Harris*, 133 S. Ct. 1050 (2013).

¹¹¹ Accuracy rates are usually measured by dividing the number of times drugs were found after a particular dog alerted to their presence by the number of times that dog alerted to the presence of drugs. For a full analysis on how error rates should be measured and the issues that arise when compiling statistics in the context of drug sniffing dogs, see *infra* Part III.B.

¹¹² *E.g.*, *United States v. Anderson*, 367 F. App’x 30, 33 (11th Cir. 2010) (holding that a canine alert was sufficient to establish probable cause where, “[e]ven assuming [defendant]’s view of the statistics, Aron [the dog] had a 55% accuracy rate in finding measurable amounts of drugs”); *United States v. Kennedy*, 131 F.3d 1371, 1378 (10th Cir. 1997) (holding that “a 70–80% success rate [by the drug detection dog Bobo] meets the liberal standard for probable cause established in *Gates*”); *United States v. Scarborough*, 128 F.3d 1373, 1378 (10th Cir. 1997) (“The affidavit states that [the dog] Pogo is certified, was trained to alert to drug odors . . . and had alerted approximately 65 times to controlled substances and contraband in four-and-one-half years on the job. Drugs were not found five of those times—for an overall reliability rate of 92 percent.”); *United States v. Huerta*, 247 F. Supp. 2d 902, 910 (S.D. Ohio 2002) (holding that the court was “unwilling to conclude that [the dog] Abby’s alert on the van, *alone*, established probable cause to search that vehicle” when Abby had an approximately 65% success rate); *Commonwealth v. Ramos*, 894 N.E.2d 611, 613–14 (Mass. App. Ct. 2008) (holding that probable cause was not satisfied where, in the past four-and-a-half years, the dog had only five to six correct alerts and two false alerts).

¹¹³ *See, e.g.*, *United States v. Diaz*, 25 F.3d 392, 394–96 (6th Cir. 1994) (holding that evidence of training and certification was sufficient but evidence of the dog’s performance was admissible on the issue of “credibility” of the dog); *United States v. Venema*, 563 F.2d 1003, 1007 (10th Cir. 1977) (holding that an affidavit describing Chane as a “trained, certified marijuana sniffing dog” was sufficient to give police probable cause to obtain a warrant to search “defendant’s locker in the Poco Storage Company”).

made the positive alert is well trained and certified,¹¹⁴ and sometimes do not even need to show that.¹¹⁵

In *Harris*, the Supreme Court rejected the idea that a “checklist” of evidence, including field accuracy rates, was necessary to establish probable cause.¹¹⁶ A unanimous Court concluded that lower courts may consider the totality of the evidence presented in favor of a dog’s reliability, and should particularly rely on a dog’s certification with a legitimate training company.¹¹⁷ The Court’s conclusion was partially based on the fact that certification organizations require dogs to be tested in controlled, simulated environments. As discussed in Part IV.B, this reliance on controlled testing over recordkeeping of field accuracy rates relies on the mistaken assumption that data on false positive rates is particularly helpful in answering the probable cause inquiry.

Whether by trusting false positive data or when conducting a more rigorous scrutiny of other binary devices, courts increasingly recognize the need to rely on quantifiable evidence relevant to individualized suspicion when conducting a probable cause inquiry. The Supreme Court’s resistance to quantifying probable cause, however, has resulted in assessments of probabilistic evidence that are non-uniform and haphazard. Questions concerning which evidence should be quantified, how to best quantify that evidence, and how to interpret the statistics, remain largely unresolved. Resolution of these questions is essential to bringing uniformity to the probable cause inquiry and protecting against unreasonable searches.¹¹⁸

III. QUANTIFYING EVIDENCE RELEVANT TO PROBABLE CAUSE

The vices of a vague probable cause standard, including the lack of clarity and susceptibility to abuse mentioned above, are not outweighed by its virtues when law enforcement replaces investigative techniques that harness a police officer’s expertise or judgment with nonhuman tools or

¹¹⁴ See 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, 761 (2007) (“Often, courts are willing to accept assertions of the dog’s training and certification as prima facie evidence of a dog’s accuracy.”); see also *id.* at 761 n.96 (citing *United States v. Berry*, 90 F.3d 148, 153 (6th Cir. 1996); *Venema*, 563 F.2d at 1007; *United States v. Stanley*, 4 F. App’x 148, 150 (4th Cir. 2001)).

¹¹⁵ See *United States v. Williams*, 69 F.3d 27, 28 (5th Cir. 1995) (per curiam) (holding that “[t]he fact that the dog alerted provided probable cause to search” even in the absence of a showing of training and reliability).

¹¹⁶ *Florida v. Harris*, 133 S. Ct. 1050, 1056 (2013).

¹¹⁷ *Id.* at 1056–57.

¹¹⁸ Although the Supreme Court, in deciding *Florida v. Harris*, missed an opportunity to properly ensure that a drug sniffing dog’s alert is sufficiently reliable to establish probable cause under the Fourth Amendment, state legislatures can create laws that require more evidence of reliability, or more record-keeping of a dog’s field accuracy rate.

forensic methods.¹¹⁹ Despite their perceived precision, investigative techniques that rely on forensic analysis or mechanized devices are fallible.¹²⁰ Indeed, techniques such as drug sniffing dogs, fingerprint analysis,¹²¹ facial recognition technology,¹²² DNA matching,¹²³ and child-pornography-detecting software¹²⁴ have variable rates of error, depending on the quality of the tool being used and the skill of the human operator.¹²⁵ There is a growing literature on the unreliability of positive

¹¹⁹ Police are increasingly investing in binary search technology, described *supra* Part II.A, because in many cases, no suspicion is required to conduct a binary search. Michael Bell, Note, Caballes, Place, and *Economic Rin-tin-tincentives: The Effect of Canine Sniff Jurisprudence on the Demand for and Development of Search Technology*, 72 BROOK. L. REV. 279, 284–88 (2006) (discussing “law enforcement agencies’ market demand for technologies—such as thermal imaging devices, x-ray scanners, stationary radar detectors, and drug sniffing canines—that enhance police officers’ abilities to detect unlawful activity, make arrests, or issue citations” (footnotes omitted)).

¹²⁰ See, e.g., U.S. DEP’T OF COMMERCE, NAT’L INST. OF STANDARDS & TECH., THE REPORT OF THE EXPERT WORKING GROUP ON HUMAN FACTORS IN LATENT PRINT ANALYSIS, LATENT PRINT EXAMINATION AND HUMAN FACTORS: IMPROVING THE PRACTICE THROUGH A SYSTEMS APPROACH (2012), available at http://www.nist.gov/customcf/get_pdf.cfm?pub_id=910745 (studying errors in fingerprint analysis); Andrew E. Taslitz, *Does the Cold Nose Know? The Unscientific Myth of the Dog Scent Lineup*, 42 HASTINGS L.J. 15, 20–25, 52–59 (1990) (discussing problems with dog scent lineups, where a dog smells evidence at a crime scene and is then used to detect an identical smell on a suspect); Natalie Ram, *The Mismatch Between Probable Cause and Partial Matching*, 118 YALE L.J. POCKET PART 182, 185 (2009), <http://thepocketpart.org/2009/04/13/ram.html> (describing how “partial [DNA] matching methods presently have a substantial rate of false positives—supposed relatives who, upon analysis, turn out not to be related”). Justice Souter largely based his dissent in *Illinois v. Caballes*, 543 U.S. 405 (2005), which held that dogs may sniff vehicles during routine traffic stops, on the fact that “[t]he infallible dog . . . is a creature of legal fiction.” *Id.* at 411.

¹²¹ U.S. DEP’T OF COMMERCE, *supra* note 120, at 32 (discussing research that shows that error rates in fingerprint matching “may not be trivial” (quoting Simon A. Cole, *More Than Zero: Accounting for Error in Latent Fingerprint Identification*, 95 J. CRIM. L. & TECH. 985, 1034 (2005))).

¹²² The New York City Police Department has invested in facial recognition technology that would, according to one court, “use thousands of government and private cameras to capture vehicle license plates, suspicious behavior and the faces of pedestrians using facial recognition technology, all of which would be fed into a central database maintained at an NYPD command center where the information would be immediately analyzed to determine whether any security threat was imminent or whether there was any suspicious pattern in the movement of persons and vehicles.” *N.Y. Civil Liberties Union v. N.Y.C. Police Dep’t*, N.Y. L.J., July 6, 2009, at 18 (N.Y. June 26, 2009).

¹²³ Familial DNA matching occurs when a DNA sample from a crime scene partially matches the DNA stored in a DNA databank. This type of partial match does not implicate the person whose DNA is stored, but relatives of that person. See Ram, *supra* note 120, at 183.

¹²⁴ See Simmons, *supra* note 14, at 453 (describing hypothetical software that sifts through millions of emails and files to uncover child pornography, boasting a low false positive rate, but a similarly low accuracy rate).

¹²⁵ Fingerprint analysis, presumed highly reliable by courts, is subject to a variety of human errors. U.S. DEP’T OF COMMERCE, *supra* note 120, at 23 (discussing human-factor errors “that are specific to an individual examiner, those that are fea-

alerts by drug sniffing dogs due to handlers misreading the dog's signals or influencing the dog's behavior.¹²⁶

Despite the potential for false positives, courts have held that positive results from some of these techniques, including drug sniffing dogs, can satisfy probable cause without any other basis for suspicion.¹²⁷ Other scholars believe that positive results from certain techniques should never be enough on their own to satisfy probable cause.¹²⁸ Rather than presume that certain types of investigative devices are always reliable enough or never reliable enough to satisfy probable cause, courts should hold that alerts from these devices may or may not be sufficient to yield probable cause. Consideration of the error rates of certain investigative tools should be deemed necessary to determine whether positive alerts by such tools are reliable enough to satisfy probable cause to conduct a full search.¹²⁹

Other investigative techniques, usually those relying on a police officer's skill or experience, would not benefit as much from a rigorous quantification of error rates. Separating these two types of techniques is critical because data collection requires resources, and because an overly technical assessment of probable cause may distort the role that police expertise and intuition plays in the inquiry.¹³⁰ It is also critical that error rates are measured in ways that can be most easily translated into the

tures or products of the work environment, and those that are the result of failures of supervision and the structure of the organization"). Drug sniffing dogs also have highly variable error rates. As Justice Souter noted in *Caballes*, "a study cited by Illinois in this case for the proposition that dog sniffs are 'generally reliable' shows that dogs in artificial testing situations return false positives anywhere from 12.5% to 60% of the time, depending on the length of the search." *Caballes*, 543 U.S. 405, 412 (2005) (Souter, J., dissenting) (citing Kelly J. Garner et al., *Duty Cycle of the Detector Dog: A Baseline Study* 12 (Apr. 2001) (prepared by Auburn Univ. Inst. for Biological Detection Sys., funded by the Fed. Aviation Admin.)); see also Bird, *supra* note 109, at 411 n.41.

¹²⁶ E.g., Katz & Golembiewski, *supra* note 114, at 762 (explaining that the handlers of drug detection dogs can "cause an error by influencing a dog's sniff or misconstruing a dog's reaction"); Lisa Lit et al., *Handler Beliefs Affect Scent Detection Dog Outcomes*, ANIMAL COGNITION (Jan. 12, 2011), <http://www.npca.net/Files/SWGDOG/LIT%20Study.pdf> (demonstrating that a dog is more likely to positively alert to drugs where handlers believe drugs are located). A lawsuit was recently filed alleging that the Nevada Highway Patrol was training its dogs to respond to cues instead of the smell of drugs. See Nicole Lucht, *NHP Troopers Sue Department Over K-9 Program*, 8 NEWS NOW (June 26, 2012), <http://www.8newsnow.com/story/18886948/nhp-troopers-sue-department-over-k-9-program>.

¹²⁷ See *supra* note 15 and accompanying text.

¹²⁸ One scholar has argued that a positive alert by a drug sniffing dog, on its own, should never satisfy probable cause. See *infra* note 164 and accompanying text.

¹²⁹ Some of these techniques are intrusive enough to already be considered searches, requiring probable cause or reasonable suspicion before they can be used. Binary searches may even trigger Fourth Amendment protections if they intrude upon privacy interests. See Simmons, *supra* note 14, at 459–62.

¹³⁰ See *infra* Part IV for arguments against quantifying probable cause generally.

probable cause inquiry, and that judges are aware of the limitations of quantification.

A. *Requiring Error Rates*

Rather than presume that positive results from mechanized or nonhuman investigative techniques are always (or never) sufficient to satisfy probable cause, quantified error rates should be included in the probable cause inquiry. This section proposes some guidelines for courts to determine when to require quantified error rates and when a rigorous assessment of error rates is unnecessary.

1. *Cases Where Quantified Error Rates Are Not Necessary*

As one scholar has remarked, “[a]ll proof is ultimately ‘probabilistic’ in the sense that no conclusion can ever be drawn from empirical data without some step of inductive inference. The question is whether to bring this probabilistic element of inference to explicit attention in a quantified way, i.e., in overtly probabilistic evidence.”¹³¹ Some of the evidence relevant to probable cause can best be assessed without overtly probabilistic attention, using testimony that need not be reduced to numbers. For example, a police officer’s testimony that a driver had a large air freshener hanging from his rearview mirror satisfied probable cause to stop the car for violating a statute prohibiting mirror obstructions.¹³² Probable cause was also satisfied when a police officer searched a duffel bag belonging to a man arrested for burglary after the police officer recognized the bag from the bank robbery.¹³³

Statistical evidence could have been used in these cases, as in every case, to assess the reliability of a police officer’s suspicion. Courts could require data on how often a police officer actually recovers evidence from her searches or, in the case of arrests, how often a targeted individual has actually committed the crime suspected by police. However, in situations where the police can testify to their direct observations, statistical evidence is not necessary to assess reliability. In these cases, statistical evidence measures only how effectively and how scrupulously officers are performing their jobs, and police officers can testify directly.

For example, the collection of data on a particular informant’s accuracy rate (or “track record”) in leading the police to criminal activity can assist the courts in determining if a particular informant is likely to be lying or misinformed in a particular case. However, discerning the accuracy of an informant is something law enforcement learns to do over time when dealing with informants. In some situations, a police officer may have reason to believe that an informant who has been unreliable in

¹³¹ Bacigal, *supra* note 20, at 295.

¹³² United States v. Garcia-Garcia, 633 F.3d 608, 612–16 (7th Cir. 2011).

¹³³ United States v. Potts, 459 F. App’x 455, 456 (5th Cir. 2012).

the past is now highly credible. A police officer's expert intuition might actually be more accurate in that case than any past statistical evidence of the informant's track record.¹³⁴ In the case of informants, police can perform the same task as statistical evidence in determining reliability, and in fact can sometimes perform the task better. Thus, while data on an informant's track record is relevant to probable cause, it should not be required.¹³⁵

As another example, one scholar has proposed that success rates in recovering evidence be reported for individual police officers (not just for their informants) as part of the probable cause determination.¹³⁶ This data would be helpful in ferreting out officers who lie to obtain warrants or who frequently misjudge informants,¹³⁷ but should not be deemed necessary to the probable cause analysis. If officers seeking to establish probable cause had to present evidence of their own track records, this might over-deter them from following their instincts in close cases for fear of being accused of malfeasance. Moreover, in cases where police are making judgment calls in unique or fast-developing situations, the factors influencing probable cause are less susceptible to statistical quantification.¹³⁸

2. Cases Where Quantified Error Rates Are Necessary

In contrast, where a police officer uses a device or investigative technique to perform a task that she cannot perform based on her observation, intuition, or judgment, statistical evidence would not measure how well the police are performing their jobs or whether they are lying to magistrate judges, but would evaluate the accuracy of a device that has essentially replaced the police officer. In addition, because the tools being measured largely do not rely on a police officer's expertise in appraising criminal behavior, and operate somewhat independently of the police officer, the officer cannot accurately assess that tool's reliability based on observation or skill. Error rates should be required, for instance, where the police use drug sniffing dogs outside vehicles at routine traffic stops because the police cannot tell from direct observation in that one instance the accuracy of a particular dog.¹³⁹

¹³⁴ For more on how courts have used informants' track records, see *supra* Part II.C.

¹³⁵ The Supreme Court's determination that the reliability of a drug sniffing dog can be assessed in the same way as a confidential informant thus ignores the ability of humans to judge each other's credibility in a way that cannot be accomplished when assessing a dog's reliability. See *Florida v. Harris*, 133 S. Ct. 1050, 1056 (2013).

¹³⁶ Minzner, *supra* note 78, at 920–21.

¹³⁷ See *id.* at 922; see generally Melanie D. Wilson, *An Exclusionary Rule for Police Lies*, 47 AM. CRIM. L. REV. 1 (2010).

¹³⁸ See *Watson v. Kansas City*, 857 F.2d 690, 695 (10th Cir. 1988) (“Whether or not probable cause exists is not susceptible to statistical quantification. It represents a judgment call on the part of the officer or officers at the scene taking into account the particular circumstances.”).

¹³⁹ The Florida Supreme Court came to a similar conclusion. See *supra* Part

Consider a case in which a witness tells a police officer that someone discarded a gun behind a dumpster after being involved in a murder. Police recover the gun, and a fingerprint analyst finds that a print left on the gun matches a suspect. The testimony of the police officer and witness (with cross-examination if a court is later reviewing probable cause at a criminal trial) speaks directly to whether the suspect touched the gun. No statistics are needed to establish the police officer's reliability. Statistical evidence on the print examiner's accuracy rate, however, is needed to assess whether his determination that the prints matched was reliable. The police cannot assess this themselves. The same would be true for DNA matches and positive alerts by drug sniffing dogs or facial recognition technology.

A more difficult case involves a police officer's testimony that a suspect was located in a "high-crime area." As mentioned above, presence in a high-crime area cannot establish probable cause or reasonable suspicion on its own, but may be relevant to the inquiry.¹⁴⁰ Suspicion based on presence in a "high-crime area" exists on the border between evidence whose reliability can be established with direct testimony and evidence that requires statistical verification. Unlike the case of an informant, where a police officer can provide testimony on why he trusted a particular informant in a particular situation, it is far more difficult for courts to assess whether, *in general*, an area has certain crime characteristics. Courts might consider, as some scholars have proposed, requiring statistical evidence to demonstrate that a particular location is a high-crime area.¹⁴¹ However, because police officers learn from experience which neighborhoods have particular profiles for criminality, police officers' testimony about a neighborhood being a high-crime area is more well-informed and instinctively based than testimony that a particular investigative device is accurate. Because evidence that an area is high in criminal activity cannot on its own satisfy probable cause, courts could simply decide that rigorous statistical evidence is less important in this context.

B. *Calculating Error Rates*

Once a court determines that a particular investigative technique requires statistical evidence to assess its reliability, the statistics must be properly calculated. False positive rates, which measure how often a device positively alerts given an innocent situation, can assess a device's

II.C. Most courts, however, presume that a drug sniffing dog's positive alert is reliable if that dog has been trained and certified, and some even refuse to hear challenges to a dog's training. *See, e.g.*, Amicus Brief of the Commonwealth of Virginia et al. in Support of Petitioner at 4, *Florida v. Harris*, 133 S. Ct. 1050 (2013) (No. 11-817) (citing cases that support the proposition).

¹⁴⁰ *See supra* Part II.D.

¹⁴¹ *See Ferguson & Bernache, supra* note 94, at 1593.

consistent, unchanging reliability.¹⁴² However, false positive rates are not directly translatable into the probable cause inquiry because they do not account for how often a device actually detects true criminal activity. This number will vary based on the population that is targeted by an investigative device. The question most appropriate to the probable cause inquiry is: What percentage of positive alerts by a particular investigative device, either alone or when combined with other evidence, actually lead to the discovery of criminal behavior?

As explained in this subsection, in order to best assist a judge in making a probable cause determination, statistical evidence must account for both a device's accuracy and the odds of criminality prior to using the device. These prior odds of criminality represent both the frequency of criminal activity among those targeted for the device's use (often referred to as a "base rate")¹⁴³ and any extra suspicion a particular suspect has elicited prior to use of the device. A positive alert and a device's false positive rate then "updates"¹⁴⁴ these prior odds of criminality. The prior odds and the consistent device accuracy combine to answer the probable cause question—or, the odds that criminal behavior is occurring.

1. False Positive Rates

To ensure that a device is accurate enough to be probative of the probable cause inquiry, courts first need to know how an investigative tool's positive alert affects the prior probability that a suspect is behaving criminally.¹⁴⁵ A device with a high false positive rate provides information that is not particularly probative in affecting the probable cause inquiry, and does not greatly influence the prior likelihood of criminality. If a device is infallible, then every time it indicates criminal activity, courts can be sure there is a 100% chance that criminal activity is occurring. In measuring the accuracy of a fallible device, courts should be concerned with "false positives," or incorrect indications of criminal behavior.¹⁴⁶

¹⁴² Simmons, *supra* note 14, at 452–53.

¹⁴³ Jonathan J. Koehler & Daniel N. Shaviro, *Veridical Verdicts: Increasing Verdict Accuracy Through the Use of Overtly Probabilistic Evidence and Methods*, 75 CORNELL L. REV. 247, 247 n.2 (1990) ("Base rates describe the frequency with which a relevant attribute occurs among members of a reference population. A base rate may also be thought of as the probability that a randomly selected member of a reference population will have the relevant attribute.").

¹⁴⁴ For a thorough discussion of how prior odds of criminal behavior can be updated with new information using a formula called Bayes' Theorem, *see generally id.* at 255–56.

¹⁴⁵ The probative value of a piece of evidence is defined by the degree to which it "tends to prove or disprove a point an issue." BLACK'S LAW DICTIONARY 639 (9th ed. 2009) (definition of "probative evidence"); *see also id.* at 1323 ("probative" defined as "[t]ending to prove or disprove").

¹⁴⁶ In the surveillance context, for example, false positives are defined as "interceptions that prove upon investigation to have no intelligence value." Richard A. Posner, *Privacy, Surveillance, and Law*, 75 U. CHI. L. REV. 245, 252 (2008). False pos-

False positives are particularly troubling in the context of the Fourth Amendment, as they lead to intrusive searches of innocent individuals.¹⁴⁷

False positive rates can be measured by testing a device in simulated situations where the device should not positively alert. For example, if facial recognition technology is exposed to 100 faces that do not actually match a suspect's face, and the technology indicates a match for 50 of those faces, the technology has a 50% false positive rate. If Bandit the cocaine-sniffing dog, in a training test, is exposed to 1000 cars that do not contain cocaine, and Bandit alerts to the presence of drugs in two of those cars, Bandit has a fairly low false positive rate of 0.2%. Bandit's accuracy means that his positive alert is highly probative in determining the likelihood that a car contains cocaine, if he also has a good specificity, meaning if he has a high "true positive" percentage of alerting when there are actually drugs in the car.

The probative value of a device in demonstrating probable cause can be measured using the ratio of the true positive to the false positive rate, or what some scholars have termed a "likelihood ratio."¹⁴⁸ The likelihood ratio measures the ratio of how often an alert would occur given that there are drugs in the car to how often an alert would occur given that there are no drugs in the car.¹⁴⁹ A dog that always alerts when there are drugs in the car, but has a 50% false positive rate has a likelihood ratio of two, or 1/0.5. A likelihood ratio of two means that the odds of drugs being in a particular location are twice as likely, given an alert, than if there is no alert.

False positive rates do not give a complete picture of probable cause, however. Even if a device has a very low rate of false positives, this does not always translate to a high likelihood of criminal activity, the essence of the probable cause inquiry.¹⁵⁰ In other words, Bandit's 0.2% false positive rate does not necessarily mean that there is now a 99.8% chance that a car that Bandit alerts to contains cocaine. If there is a low incidence of cocaine in vehicles targeted by Bandit, say one out of every 10,000 vehicles, Bandit's 0.2% false positive rate will not translate to a high likelihood of drugs in a vehicle given a positive alert. For every 10,000 vehicles, Bandit will falsely alert to the presence of cocaine 20 times (10,000 x .002), and he will correctly alert to the presence of drugs

itives are called a Type I error in statistics. See Alexander A. Reinert, *Public Interest(s) and Fourth Amendment Enforcement*, 2010 U. ILL. L. REV. 1461, 1492-93 and n.179 (2010).

¹⁴⁷ See Myers, *supra* note 19, at 12-13.

¹⁴⁸ D. H. Kaye & Jonathan J. Koehler, *The Misquantification of Probative Value*, 27 LAW & HUM. BEHAV. 645, 648 (2003).

¹⁴⁹ See DAVID H. KAYE, *THE DOUBLE HELIX AND THE LAW OF EVIDENCE* 20-22 (2010).

¹⁵⁰ See Myers, *supra* note 19, at 12-16; see also Simmons, *supra* note 14, at 452 (explaining that although a false positive rate is constant, the actual likelihood of criminality "will vary widely depending on the actual frequency of the illegal activity that is being investigated").

once, assuming he always alerts when drugs are present. Thus, given a positive alert by Bandit, there will be only a 1/21 chance, or about a 5% chance, that a particular vehicle alerted to by Bandit will contain drugs. This should not satisfy probable cause—a 5% likelihood of criminal activity can hardly be said to warrant a person of “‘reasonable caution in the belief that’ an offense has been or is being committed.”¹⁵¹ In short, Bandit’s false positive rate means that his positive alert is very *probative* in a probable cause inquiry, and it highly alters the prior probability of criminality of 1/10,000. However, the low percentage of drugs in the population targeted by Bandit means that his positive alert likely does not cross the probable cause hurdle.¹⁵²

This example illustrates why statistical evidence that accounts solely for false positive rates does not actually measure the odds of criminal behavior.¹⁵³ Professor Richard Myers has explained, in the context of drug sniffing dogs and probability theory, why the odds that a dog will alert when there are no drugs in a vehicle (the false positive rate) is not the same as the odds that there are no drugs in a vehicle given that a dog has alerted (the information needed to answer the probable cause question).¹⁵⁴ Using an equation called Bayes’ Theorem, Myers explains that what is missing from the false positive rate to convert it into the odds of criminal behavior is the prior probability of criminal activity.¹⁵⁵ This prior probability can reflect either the degree to which an officer already suspects an individual,¹⁵⁶ or the frequency of the criminal behavior within a population targeted by a device.¹⁵⁷ As with the Bandit example, the false positive rate of a device is said to update the original odds of criminality with new odds for a particular suspect who has been identified as engaging in criminal behavior by the device.¹⁵⁸

Because base rates are so variable, the Supreme Court in *Florida v. Harris* was unduly hasty in concluding that “evidence of a dog’s satisfactory performance in a certification or training program can itself provide sufficient reason to trust his alert.”¹⁵⁹ These programs use simulated testing that does not account for the variability of base rates of

¹⁵¹ *Brinegar v. United States*, 338 U.S. 160, 175–76 (1949) (quoting *Carroll v. United States*, 267 U.S. 132, 162 (1925)).

¹⁵² See Myers, *supra* note 19, at 14 (“If the probability [of criminal activity] was low to begin with, even a really good test will still result in a relatively low [likelihood of that criminal activity].”). Although we do not know what the probable cause hurdle is, it is likely greater than 5% suspicion.

¹⁵³ See *id.* at 12–16. For another example of why false positives cannot answer that probable cause question, using software that searches for child pornography, see Simmons, *supra* note 14, at 453.

¹⁵⁴ Myers, *supra* note 19, at 12–16.

¹⁵⁵ *Id.*

¹⁵⁶ *Id.*

¹⁵⁷ See Simmons, *supra* note 14, at 451–53.

¹⁵⁸ See Myers, *supra* note 19, at 12–16.

¹⁵⁹ *Florida v. Harris*, 133 S. Ct. 1050, 1057 (2013).

drugs within the population. As explained in the next subsection, the Court's elevation of "a dog's real false positives"¹⁶⁰ over field accuracy rates misunderstands the statistics involved, and places the probative value of false positives over the actual probable cause inquiry: the likelihood that drugs will actually be present in a particular car given a positive alert.

2. Accuracy Rates

To properly measure the actual likelihood of criminal behavior based on a positive alert by an investigative tool, statistical evidence must therefore account for both the tool's accuracy (the false positive rate and the true positive rate) and either the frequency of criminality of those targeted for a device's use, or some measure of a police officer's suspicion prior to the positive alert (the prior odds of criminality).¹⁶¹ The frequency of criminal behavior in a population is difficult to measure, and it is basically impossible to quantify many of the intangible factors tending towards suspicion.¹⁶² Because even a low false positive rate is not a good indicator of likelihood of criminal activity, especially with largely innocent populations or where the police have no other reason to suspect someone of wrongdoing,¹⁶³ Professor Myers argues that a positive alert by a drug sniffing dog should never be sufficient on its own, without knowing the prior odds of criminality, to satisfy probable cause.¹⁶⁴

However, there is a way to collect data that accounts for both a tool's accuracy and the prior odds of criminality within the population targeted for use by a device. Police officers could record statistics in the field instead of measuring false positives in simulations. Field statistics measure how often a device's positive alert leads to finding criminal activity, the key number for the probable cause analysis. Many courts already use field statistics in the context of drug sniffing dogs,¹⁶⁵ although law enforcement does not always appear to be calculating them correctly.¹⁶⁶ To translate most directly into the probable cause inquiry,

¹⁶⁰ *Id.*

¹⁶¹ In the area of drug detection dogs, according to one scholar, "[p]olice usually use drug detection canines only after a suspect or item is sufficiently suspicious to warrant the use of a dog. Therefore, the relevant 'population' is not the general populace, but a narrowed group that police have determined through their training and experience are more likely to possess narcotics than the general public." Bird, *supra* note 109, at 429 (footnote omitted).

¹⁶² For example, how does one quantify the added suspicion gained when a suspect appears nervous, especially given that whether or not a suspect appears nervous is a vague and fairly subjective assessment?

¹⁶³ See *supra* note 150 and accompanying text.

¹⁶⁴ Myers, *supra* note 19.

¹⁶⁵ See, e.g., *United States v. Scott*, 610 F.3d 1009, 1014 (8th Cir. 2010) ("Officer Berg testified that Naton's accuracy rating is 98 or 99 percent in training and 85 percent in the field."). For more examples, see *supra* note 112.

¹⁶⁶ See, e.g., *United States v. Smith*, 448 F. App'x 936, 939 (11th Cir. 2011) ("Moreover, even considering [the dog] Dusty's alleged 70% accuracy rate in the

courts must assess the number of times drugs were found in the car out of the times that the dog had positively alerted, not given all the times a dog had been used.

As an example, if Lucy the drug sniffing dog positively alerts to the presence of drugs in 100 cars targeted for drug sniffs in New York City, and 90 of those cars actually contain drugs, Lucy has a 90% field accuracy rate. If those field statistics can be extrapolated generally, which is a difficult question that will be discussed in the next subsection,¹⁶⁷ courts can then conclude that if Lucy alerts to the presence of drugs in a vehicle, there is a 90% chance that drugs will be found in that vehicle. This answers the probable cause question, as a 90% likelihood of criminality will certainly satisfy probable cause. While 10 out of 100 car searches guided by an alert will likely not contain drugs, probable cause does not require that no innocents are searched. In contrast with a 50% *false positive rate*, if Lucy positively alerts to the presence of drugs 100 times, and 50 of those cars actually contain drugs, we might then also be able to say that there is a 50% chance that drugs are actually in the vehicle if Lucy alerts. Depending on the actual probable cause hurdle, probable cause may or may not be satisfied in that case.

For many cases, field statistics can directly translate into the probable cause inquiry. Field statistics reflect the degree of suspicion associated with a suspect's behavior and the likelihood of criminality prior to using a device because, if officers target only those who are more likely to be behaving criminally, field accuracy rates will increase dramatically.¹⁶⁸ Of course, police officers will then be using their intuitive judgment to determine which suspects to target, but that intuition is then entirely

field, a positive alert would be sufficient to establish a 'fair probability' that drugs would be found in the car."). At the district court level, Dusty's accuracy rate should have been 56% but was calculated as 30% because the court did not simply account for the times that drugs were found in the vehicle given a positive alert, but also added the times that Dusty did not alert at all. *See* United States v. Smith, No. 2:09-cr-59-FtM-29SPC, 2009 U.S. Dist. LEXIS 123516, *13-14 (M.D. Fla. 2009). According to the district court, "[The investigator] also testified to reviewing the records of the K-9 used in this case specifically for the years 2008 and 2009. During that time period, Dusty, the K-9, was used to potentially detect narcotics a total of 89 times. Thirty-two (32) of those times there was no alert and 57 times Dusty actually alerted. No narcotics or narcotics paraphernalia were found as a result of 25 of those alerts. . . . In calculating the percentages, [the investigator] testified that almost 30 percent of the time Dusty alerted, there were no drugs present." *Id.* (citations and footnotes omitted).

¹⁶⁷ Extrapolation of field data to formulate a general likelihood is difficult because the base rate of criminality varies, so questions will arise as to how specifically statistics need to be collected and broken down with respect to different locations, types of suspects, etc. For example, do field statistics need to account for use of the device among middle-aged men in Central Park, or can youths in the Upper West Side generally be aggregated with that data? What about all people in Manhattan? This question will be confronted *infra* Part III.C.

¹⁶⁸ Thus, requiring officers to collect field statistics may render the officers less willing to use particular devices on subjects in a largely innocent population without any prior suspicion.

reflected in an increase or decrease in the device's accuracy. The device thus both reflects and replaces the officer's intuition. In the Lucy example, any prior suspicion a police officer has before using Lucy is essentially transferred into her nose in altering her accuracy rate. Moreover, once the device has essentially replaced the police officer's intuition, only statistical evidence can attest to the device's accuracy rate. Ultimately, it is the positive alert by the device, and not the police officer, that makes a prediction relevant to probable cause.

Field statistics even account for human error in using a device, as decreased accuracy rates in the field will result if law enforcement is using a device incorrectly. In the case of drug sniffing dogs, a dog's handler may unconsciously cue the dog to alert when the handler thinks a vehicle is likely to contain drugs.¹⁶⁹ This interaction between dog and handler will be incorporated into the field accuracy rates. If an officer is incorrect most of the time, the accuracy rate will decrease. If the interaction between dog and handler makes it more likely that a positive alert will lead to the discovery of drugs, this will also be reflected in the accuracy rate.

Field statistics thus incorporate the prior odds of criminal wrongdoing and the accuracy of an investigative tool, which translate to the probability of criminal activity. Taking field statistics has some costs, however, and it may be difficult to discern which field statistics can be extrapolated so as to be relevant to a particular probable cause inquiry. These issues will be addressed in the next subsection.

C. *Problems with Quantifying Error Rates*

An obvious difficulty in mandating that quantified error rates be included as part of the probable cause inquiry is the time and cost of collecting this data. Although federal, state, and local legislatures routinely require that law enforcement collect certain statistics,¹⁷⁰ the collection of data on accuracy rates relevant to the probable cause inquiry would, necessarily, be an unfunded mandate—courts can require the collection of such statistics, but cannot allocate money to fund the effort. However, many law enforcement groups already collect statistics testing the accuracy of particular investigative devices in order to more effectively perform their jobs.¹⁷¹ In many ways, the collection of this data would redound to the benefit of the law enforcement community: skeptical judges can be satisfied when devices are accurate, and law

¹⁶⁹ See Myers, *supra* note 19, at 34.

¹⁷⁰ See, e.g., Jordan Blair Woods, Comment, *Addressing Youth Bias Crime*, 56 UCLA L. REV. 1899, 1902 (2009) ("Bias crime statistics legislation allocates resources to facilitate the collection of data on bias crimes.").

¹⁷¹ See, e.g., Myers, *supra* note 19, at 33 (explaining that the United States Army mandates the collection of drug sniffing dogs' field accuracy rates and proposing that, "[a]t the very least, the courts should mandate the collection of data on the use of the dogs and their accuracy rates in the field").

enforcement can have more faith in its investigations. Requiring statistical evidence on a device's accuracy rate may deter police officers from using devices in random, haphazard ways, which would otherwise result in more false alerts in largely innocent populations.¹⁷²

Of course, asking law enforcement, or forensic scientists,¹⁷³ to collect statistical evidence on the accuracy of its investigative tools could be likened to letting the fox guard the henhouse. Random checks and supervisory oversight can prevent police officers from improperly reporting statistics, but unscrupulous police officers may fail to report searches that do not result in uncovering evidence of illegal activity. That said, police officers are asked to collect statistics in areas that impugn their integrity far more than recording the accuracy of their investigative devices.¹⁷⁴

Even if police scrupulously record statistics, whether a device has actually produced a false alert may be a complicated question. For example, law enforcement officers may argue that if a drug sniffing dog gives a positive alert, but no drugs are found, then the area recently contained drugs that were removed prior to the search.¹⁷⁵ For the purposes of the probable cause inquiry, the best approach would be to assume an inaccurate alert in all cases where illegal activity is not ultimately found. Because there is usually no way to tell whether drugs have been removed or if they were never there in the first place, assuming that the dog has alerted falsely best measures the percentage of cases in which a positive alert actually leads to the detection of criminal behavior, the essence of the probable cause inquiry. The point of collecting accuracy rates is ultimately not to test the quality of the dog, but the meaning of an alert by the dog—an alert that means “there used to be cocaine here” is not enough for a search.¹⁷⁶

¹⁷² See Bird, *supra* note 109, at 430–33.

¹⁷³ The study and collection of data on accuracy rates for finger print examiners is explored thoroughly in *The Report of the Expert Working Group on Human Factors in Latent Print Analysis*. See generally U.S. DEP'T OF COMMERCE, *supra* note 120.

¹⁷⁴ For example, the United States Department of Justice (USDOJ) has entered into agreements with several law enforcement agencies requiring “the collection and analysis of demographic data on police stops. USDOJ typically monitors agencies operating under these agreements for continuing statistical evidence of bias.” Michael R. Smith, *Depoliticizing Racial Profiling: Suggestions for the Limited Use and Management of Race in Police Decision-Making*, 15 GEO. MASON U. C.R. L.J. 219, 221 (2005).

¹⁷⁵ See 1 WAYNE R. LAFAVE, *SEARCH AND SEIZURE: A TREATISE ON THE FOURTH AMENDMENT* § 2.2(f) n.239, at 450 (3d ed. 1996) (“It has been estimated that most of the cash in circulation (the estimates range from 70% to 97% of all bills) contains sufficient quantities of cocaine to alert a trained dog. Thus it sometimes happens in practice that a drug dog alert will lead to nothing but currency.” (citation omitted)).

¹⁷⁶ For this reason, the Supreme Court was incorrect to conclude, in *Florida v. Harris*, that because of residual odors, a positive alert that leads to no discovery of evidence does not undermine a dog's reliability in the probable cause inquiry. 133 S.

Problems will also arise in deciding which field accuracy data should be considered in the probable cause determination. Courts must consider whether all field data collected with respect to a particular dog is relevant, or only field data that is specific to a dog/law enforcement handler pair. Field sobriety tests, as another example, show a wide variance in accuracy depending on the officer administering the test.¹⁷⁷ Thus, data should be recorded for each officer individually.

Because both the baseline frequency of criminal behavior in a population and odds with respect to a particular suspect based on her behavior contribute to field accuracy rates, courts must decide whether data from different areas in the same city can be combined, and whether data from the use on someone who behaved suspiciously can be combined with data for a suspicionless use of the device.¹⁷⁸ If the police always use a device when there is some reason to suspect an individual of criminal behavior, this will compensate for any differences in the frequency of criminality in different locations and produce a consistent accuracy rate. In that instance, data from different locations in the same city can be combined and extrapolated into the probable cause inquiry.

Questions regarding the basis for accuracy rates may mire courts in battles over whether the statistical evidence offered can be directly translated into the probable cause inquiry. However, this has not been overly problematic for the jurisdictions that already include statistical evidence in the probable cause inquiry.¹⁷⁹ Currently, as mentioned above, courts already use field accuracy statistics and extrapolate to the probable cause inquiry in the areas of informants' track records, efficacy of checkpoints, and high-crime areas; they are just not doing so in a particularly informed or systematized way.¹⁸⁰

The concern over which evidence is relevant to the probable cause inquiry can be resolved by simply giving more specific evidence more weight. Professor Max Minzner proposed a similar solution in the context of arguing that police should be required to present their accuracy rates as part of the probable cause inquiry.¹⁸¹ According to Minzner, the way in which the evidence was gathered, and how specific it is, should influence how much weight the evidence merits in the probable cause inquiry.¹⁸²

Ct. 1050, 1056–57 & n.2 (2013). Probable cause considers the likelihood that evidence will actually be *found*, not the likelihood that a suspect actually committed the crime involved. An alert that does not lead to the detection of evidence should therefore be considered a false alert, even if the dog is correctly responding to the former presence of drugs.

¹⁷⁷ See, e.g., *United States v. Horn*, 185 F. Supp. 2d 530, 539–40 (D. Md. 2002).

¹⁷⁸ This might be problematic because individuals whom the police have reason to suspect of wrongdoing may have higher prior odds of criminality, affecting the Bayesian updating of a device's accuracy.

¹⁷⁹ See *supra* note 112.

¹⁸⁰ *Supra* Parts II.D and III.B.

¹⁸¹ See *supra* note 78 and accompanying text.

¹⁸² Minzner, *supra* note 78, at 955. How courts can incorporate quantifiable

For example, if evidence is collected pertaining to the field accuracy rate of a particular drug sniffing dog/handler pair in a specific area in Philadelphia, and the dog is generally used only when police suspect an individual for other reasons insufficient to satisfy probable cause on their own, that data is highly relevant to the likelihood that an individual in that neighborhood, when targeted based on the usual level of suspicion by that particular dog/handler pair, is carrying drugs.

In other words, statistical evidence collected in a very specific way applies to a particular instance that matches the other nearly identical field data. In contrast, evidence that is supported only by a heterogeneous sample carries less weight and may not be directly extrapolated into the probable cause likelihood, but can still be considered. If an accuracy rate is not measured specifically enough, for example, a court could hold that the prosecution has not met its burden of proof to satisfy probable cause. Or, a court could conclude that, in showing that a particular dog has a 70% accuracy rate with less specific statistical evidence, that the evidence is sufficient to meet the 50% hurdle.

Those who argue against overly quantified legal standards also present the “capacity objection,” which asserts that courts, not generally experts in mathematics, are not well equipped to assess statistical evidence.¹⁸³ The capacity objection provides another reason why courts should require statistical evidence only when there is no other valid way to assess whether a positive alert by that device gave rise to probable cause. However, this objection has much more salience when juries, not judges, are examining probabilistic evidence.¹⁸⁴ According to Minzner, “magistrate judges . . . are repeat players in the search process. If statistical learning is possible, the search process presents the ideal test case. Furthermore, there is little evidence that judges are better at making a probability calculation from nonstatistical evidence.”¹⁸⁵

Judges, repeat players, can be instructed as to the pitfalls of weighing explicitly probabilistic evidence. The evolution of the increasingly nuanced and scientifically accurate role of DNA evidence in criminal

evidence into the probable cause inquiry will be discussed *infra* Part IV.

¹⁸³ One scholar has explained that “capacity-based objections are claims that legal decision makers simply cannot make good decisions based on statistical data because they are not trained in drawing inferences from empirical data and, in any event, good, relevant data is generally unavailable.” Minzner, *supra* note 78, at 951. For further discussion of this objection, see *infra* Part IV.

¹⁸⁴ See Colin Miller, *Diamond Dogs, Take 2: Etan Patz, Florida v. Harris & Quantifying Probable Cause*, EVIDENCEPROF BLOG (Apr. 22, 2012), <http://lawprofessors.typepad.com/evidenceprof/2012/04/cadaver-dogs-are-back-in-the-news-with-the-revelation-that-a-cadaver-sniffing-dog-detected-the-odor-of-human-remains-in-a-base.html> (discussing the differences between judges’ and juries’ weighing of probabilistic evidence).

¹⁸⁵ Minzner, *supra* note 78, at 954–55; see *supra* Part III.B.

trials demonstrates how this learning occurs.¹⁸⁶ Although *Daubert*-type hearings, which ensure that scientific evidence satisfies the requirements for expert testimony at trial, would not be necessary at the probable cause stage,¹⁸⁷ the role of judge as gatekeeper to admit only scientifically valid testimony at trial should allay fears about a judge's ability to conduct a flexible inquiry that examines error rates and is probabilistic in nature.¹⁸⁸

In some cases, courts may decide they must use false positive data instead of field accuracy rates. Using data on false positives is an attractive option because false positive rates are more easily measured and do not change based on the frequency of criminality within a population. New devices or recently developed techniques may require consideration of false positives, as there will be insufficient field statistics to measure accuracy. Alternatively, courts could use statistics from other devices; they could presume, for instance, that a well-trained dog that received a similar type of training is generally a certain percentage accurate until the dog has a large enough number of alerts.

False positive data, generated by simulations, would also have to substitute for field data in cases where police cannot ultimately verify whether a suspect was behaving criminally. This occurs in cases of fingerprint or DNA analysis, where a forensic scientist may find a match, but there is no way to ultimately know if the suspect is guilty. Testing can still be performed on individual fingerprint examiners using artificial experiments to measure an examiner's false positive rate.¹⁸⁹ However, as mentioned above, false positive data does not speak directly to the probable cause inquiry.¹⁹⁰

In cases where statistics truly cannot be collected because of extreme hardship, cost, or complexity of data collection, courts could mandate standardized training in using a device.¹⁹¹ One of the reasons that the

¹⁸⁶ See generally KAYE, *supra* note 149.

¹⁸⁷ As one federal circuit court remarked with respect to dog sniffs, "the dog's alert . . . would serve not as actual evidence of drugs, but simply to establish probable cause to obtain a warrant to search for such substantive evidence." *United States v. Allen*, 159 F.3d 832, 840 (4th Cir. 1998).

¹⁸⁸ The Supreme Court in *Daubert* listed several factors for judges to consider before admitting expert testimony under Federal Rule of Evidence 702. The Court instructed judges that, "in the case of a particular scientific technique, the court ordinarily should consider the known or potential rate of error, and the existence and maintenance of standards controlling the technique's operation." *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 594 (1993) (citations omitted).

¹⁸⁹ See U.S. DEP'T OF COMMERCE, *supra* note 120, at 33–38.

¹⁹⁰ See *infra* Part IV for incorporation of different types of error rates into the probable cause inquiry.

¹⁹¹ Richard Myers has proposed this with respect to drug sniffing dogs. See Myers, *supra* note 19, at 34 ("If we are really interested in protecting the public's Fourth Amendment interests, we need to set state and federal standards for training dogs, rather than leaving them to the private sector. The process of drafting standards will move that debate out of the courts, where they are enforced by the exclu-

Florida Supreme Court required field evidence of drug sniffing dogs' accuracy as part of the probable cause inquiry was because "[i]n the absence of a uniform standard [of training of drug sniffing dogs], the reliability of the dog cannot be established by demonstrating only that a canine is trained and certified."¹⁹² Standardized training and uniform protocols have also been suggested for fingerprint examiners.¹⁹³

Civil liberties are compromised when law enforcement uses fallible devices that sanction intrusive searches with no indication of their reliability. Ultimately, it is better to quantify accuracy rates than simply presume that a particular device's positive alert gives rise to probable cause. Especially with new technology on the horizon,¹⁹⁴ courts should establish standards for measuring an investigative tool's reliability.

IV. QUANTIFIABLE EVIDENCE AND THE PROBABLE CAUSE STANDARD

As mentioned earlier, some courts have already begun mandating admission of quantifiable error rates for specific investigative techniques. Although the Supreme Court has declared that field accuracy rates are not absolutely required by the Fourth Amendment in determining the reliability of drug sniffing dogs, many prosecutors will continue to admit this type of evidence, as they have done in the past.¹⁹⁵ Further, the prosecution may wish to proffer this type of evidence as part of the totality of the circumstances in order to rebut evidence by the defendant that a particular alert is unreliable or a particular dog's certification is insufficient to establish probable cause.¹⁹⁶ Courts are already confronting explicitly probabilistic evidence in many forms. When evidence sufficient to satisfy probable cause is admitted in probabilistic terms, the reasons for not assigning a minimum probability of certainty necessary to satisfy probable cause are far less compelling.

There have not been many proposals to quantify probable cause. According to Professor Craig Lerner, "[f]ew courts have summoned the courage, or foolhardiness, to propose . . . a number (e.g., thirty percent probability) for probable cause."¹⁹⁷ Lerner instead offered a

sionary rule, and into police standards commissions and legislatures.").

¹⁹² *Harris v. State*, 71 So. 3d 756, 768 (Fla. 2011), *overruled by* *Florida v. Harris*, 133 S. Ct. 1050 (2013).

¹⁹³ See U.S. DEP'T OF COMMERCE, *supra* note 120, at 90, 164–67.

¹⁹⁴ See, e.g., Sam Kamin, *Law and Technology: The Case for a Smart Gun Detector*, 59 LAW & CONTEMP. PROBS. 221, 221 (1996).

¹⁹⁵ See *supra* Part I.

¹⁹⁶ In *Florida v. Harris*, the Supreme Court held that a defendant may counter the prosecution's evidence that a dog is well trained "by disputing the reliability of the dog overall or of a particular alert." 133 S. Ct. 1050, 1058 (2013). This may lead particular states to require record keeping on field accuracy rates anyway, as these accuracy rates are certainly relevant to any totality of the circumstances inquiry.

¹⁹⁷ Craig S. Lerner, *The Reasonableness of Probable Cause*, 81 TEX. L. REV. 951, 995 (2003).

mathematical formula to evaluate whether there is sufficient probable cause based on the probability of a successful search and its social value against a search's social cost, or privacy intrusion.¹⁹⁸ Another scholar, Ronald Bacigal, came closer to quantifying probable cause, by assigning probable cause a range of approximately 40–49%,¹⁹⁹ but stopped short of “endorsing precise mathematical expressions of probable cause.”²⁰⁰ Bacigal, wishing to add some clarity to the probable cause inquiry, still cautions that “[t]oo much precision diminishes the resources of ambiguity.”²⁰¹

There are situations, however, where the utility of ambiguity is outweighed by the benefits of clarifying the standard, especially in light of recent developments that have greatly diminished review of probable cause decisions.²⁰² Quantifying probable cause would add both integrity and accountability to the jurisprudence, allowing citizens to evaluate courts' decisions and reducing cynicism over application of the fuzzy, “malleable probable cause analysis.”²⁰³ Quantifying probable cause can also provide police with a more certain standard to guide their decisions in the field.

This final section examines how assigning a minimum numerical value to the probable cause hurdle would function in a variety of scenarios, and explores how courts should choose the numerical value. In so doing, this section rebuts the myriad objections by courts and scholars to quantifying probable cause.

A. *Quantifying Probable Cause*

Quantifying probable cause presents varying benefits and drawbacks, depending on the types of evidence used to satisfy probable cause. There are three basic scenarios in which judges would confront a numerical probable cause standard. In cases where a positive alert from a mechanized device is the critical factor in producing probable cause, many of the objections to assigning probable cause a value disappear. The quantifiable evidence can be used directly to determine whether the probable cause standard is satisfied. In cases where both quantitative and

¹⁹⁸ *Id.* at 1019–21. Lerner's formula looks like this: $P \times V > (1 - P) \times (C \times m)$, where P is the probability a search will be successful, V is the value of the search, C is the cost of the search, and m is the privacy multiplier for particularly stigmatizing searches. *Id.* at 1020–21.

¹⁹⁹ Bacigal, *supra* note 20, at 338–39.

²⁰⁰ *Id.* at 339.

²⁰¹ *Id.* Bacigal also argued that the requisite degree of suspicion should depend somewhat on the severity of the crime and the exigencies of the circumstances. *See id.* at 323–32. These views will be addressed *infra* Part IV.B.

²⁰² *See supra* Part II.B.

²⁰³ *See* John M. Burkoff, *The Fourth Amendment and Terrorism*, 109 PENN ST. L. REV. 911, 920–21 (2005) (describing the “fluid’ and open-ended conception of probable cause” articulated in *Maryland v. Pringle*, 540 U.S. 366, 370–72 (2003)).

qualitative evidence are necessary to satisfy probable cause, judges cannot rely solely on the number produced by the quantifiable evidence and compare it to the probable cause number. However, judges can use the value associated with probable cause to discern how close the quantifiable evidence comes to satisfying probable cause and can then assess whether the qualitative evidence, combined with the quantitative evidence, crosses the probable cause threshold. In cases where only qualitative evidence is considered, assigning probable cause a numerical percentage is valuable only to the extent that it provides a conceptual benchmark, but judges should not attempt to actually quantify the inquiry.

1. Cases Where Quantifiable Evidence Is the Critical Factor in Satisfying Probable Cause

In situations where quantifiable evidence is the sole or critical factor creating probable cause, the benefits of added clarity in the law outweigh the downsides of quantifying probable cause. Consider a common scenario. The police pull over a vehicle and briefly detain an individual for reasons unrelated to drug possession, such as making an improper lane change,²⁰⁴ or having an improperly illuminated license plate.²⁰⁵ While detained, the individual behaves suspiciously, perhaps appearing nervous when answering questions,²⁰⁶ but the behavior is not sufficiently suspicious to give the police probable cause to search the vehicle for evidence of illegal activity.²⁰⁷ However, the suspicion suffices to permit the police to continue detaining the individual to wait for a drug sniffing dog to arrive and sniff the exterior of the vehicle.²⁰⁸ The dog positively alerts to the presence of drugs inside the vehicle, giving the police probable cause to search the vehicle, and drugs are ultimately found.²⁰⁹ A court upholds the search, largely based on the dog's alert.²¹⁰ The court will cite language such as “[a] positive identification by a dog during a canine search following a lawful stop of a vehicle provides probable cause that drugs are present in the vehicle, thereby justifying a search of the vehicle.”²¹¹

²⁰⁴ See, e.g., *United States v. Lopez-Gutierrez*, 334 Fed. App'x. 880, 881 (10th Cir. 2009).

²⁰⁵ See, e.g., *United States v. Mohamed*, 600 F.3d 1000, 1002 (8th Cir. 2010).

²⁰⁶ *Id.*

²⁰⁷ See *id.* (“Trooper Frisby told Mohamed that he had reasonable suspicion to run a drug dog around the car.”).

²⁰⁸ See *Lopez-Gutierrez*, 334 Fed. App'x. at 883 (“A traffic stop may be extended if the officer has reasonable articulable suspicion of criminal activity.”).

²⁰⁹ See *Mohamed*, 600 F.3d at 1002–03.

²¹⁰ See *Lopez-Gutierrez*, 334 Fed. App'x. at 883 (“An alert by a certified drug-sniffing dog during a sweep of the exterior of the vehicle can give probable cause to search the interior.”).

²¹¹ *Mohamed*, 600 F.3d at 1004.

In this scenario, the sniff is the critical factor in satisfying probable cause. The police would likely concede that probable cause did not exist prior to the positive alert, and a court would likely justify the search almost exclusively based on that positive alert. Now assume that this dog and handler's field accuracy rate is 50%.²¹² This court's probable cause assessment would be greatly assisted if the court knew that probable cause requires a 35% probability, a 40% probability, a 50%, or a 51% probability of criminal activity.

Accepting the field accuracy rate as the best evidence of the likelihood that drugs are in the vehicle, the court can compare the field accuracy rate to the assigned probable cause standard to determine if probable cause is met solely based on the positive alert. Courts make this comparison already.²¹³ Courts holding that the accuracy rates of certain devices are sufficient to satisfy probable cause must have some fuzzy minimum threshold in their minds, but that number is not uniform and is usually not transparent to the citizenry, the police, or other courts. As one court held in the context of drug sniffing dogs, "based on historical performance, this dog's alert suggested a 58% chance of finding a seizable quantity of drugs. While we hesitate to get into the business of affixing figures on probable cause, if we were pushed to do so we would hold this to be enough."²¹⁴

This example demonstrates a situation where many of the objections to quantifying the probable cause standard are overblown. When the evidence in favor of suspicion is already quantified in a form that directly translates into the probable cause question,²¹⁵ it seems silly to doggedly contend that probable cause is incapable of quantification. Although some scholars have argued that "using math to enlighten discussions of probable cause"²¹⁶ gives the illusion of precision, the sense of precision gleaned from quantifying probable cause is no longer false when field accuracy rates are collected and calculated correctly. Further, the concern that assigning a minimum probable cause probability will cause courts to give undue weight to probabilistic evidence over nonscientific or case-specific evidence is greatly diminished in cases where probable cause is satisfied based on quantifiable evidence.²¹⁷

²¹² For the difference between false positive rates and field accuracy rates, see *supra* Part III.B.

²¹³ *United States v. Anderson*, 367 Fed. App'x 30, 33 (11th Cir. 2010) (allowing a 55% accuracy rate to satisfy probable cause); *United States v. Koon Chung Wu*, 217 Fed. App'x 240, 246 (4th Cir. 2007) ("[A]n accuracy rate of 60% is more than reliable enough for [the dog's] alert to have established probable cause.").

²¹⁴ *United States v. Ludwig*, 641 F.3d 1243, 1252 (10th Cir. 2011) (footnote omitted).

²¹⁵ For the form in which this evidence should be submitted, see *supra* Part III.B.

²¹⁶ Antkowiak, *supra* note 20, at 586 (arguing that "math seduces lawyers with the siren song of precision, certainty, and constancy").

²¹⁷ Scholars have vigorously debated whether, in the context of jury trials, jurors give undue weight to probabilistic evidence, like the probability associated with

Moreover, in cases like the one presented above, even if there are other factors besides the quantified evidence that also create suspicion, these factors can be directly translated into the accuracy rate of the device. The qualitative evidence essentially becomes quantitative. Thus, if field statistics are collected in situations that closely resemble the case at issue,²¹⁸ the court should not even factor in other suspicious behavior, like the failure to maintain eye contact with the police. The field accuracy rate will already reflect the fact that most suspects targeted for the drug sniff display some amount of suspicious behavior.²¹⁹

However, there will be instances where the data on field performance is not sufficiently numerous or specific to extrapolate the likelihood of criminal behavior from the field accuracy rate alone.²²⁰ In those cases, there is a legitimate concern that assigning probable cause a numerical value will cause a judge to nonetheless feel compelled to use the insufficient or inapposite accuracy rate to answer the probable cause inquiry. A more generalized version of this fear has been articulated by Professor Orin Kerr, who contends that if probable cause is quantified into a hard number, judges will be less likely to heed their logical, yet unexpressed, intuitions and instincts about a case.²²¹

Professor Kerr uses an example of a fictional “scientific study by top Harvard scientists showing that marijuana can be found in 60 percent of Harvard dormitory rooms.”²²² He argues that quantifying probable cause may stifle a judge’s intuition that the 60% suspicion created by the study, although a high enough percentage to satisfy probable cause, feels insufficient. This instinct is based on the subconscious but logical skepticism as to why additional evidence was not collected to rule out or place specific suspicion upon the particular dorm room that police sought to search.²²³ Instead of allowing judges to follow their intuitions

a suspect’s DNA matching evidence left at a crime scene, at the expense of nonscientific or circumstantial evidence. See, e.g., Brenda Inman Rowe, Note, *A Possible Solution for the Problem of Juries Slighting Nonscientific Evidence: A Bayesian-Like Judicial Instruction*, 24 AM. J. CRIM. L. 541, 541–47 (1997) (“An important question in this age of increasing use of scientific evidence in criminal trials is whether jurors consider scientific evidence to the exclusion, or near-exclusion, of nonscientific evidence.”); Brian C. Smith et al., *Jurors’ Use of Probabilistic Evidence*, 20 LAW & HUM. BEHAV. 49, 54–78 (1996) (concluding that mock jurors can reasonably assess probabilistic evidence). Some scholars have actually concluded that juries do not attach enough weight to probabilistic evidence over case-specific evidence. E.g., Julie A. Singer et al., *The Impact of DNA and Other Technology on the Criminal Justice System: Improvements and Complications*, 17 ALB. L.J. SCI. & TECH. 87, 108 (2007) (“[J]urors . . . often seriously underestimate the weight of probabilistic evidence.”).

²¹⁸ See *supra* Part III.B for an example of this.

²¹⁹ See *supra* Part III.B.

²²⁰ See *supra* Part III.C. for problems with measuring accuracy rates.

²²¹ Kerr, *supra* note 5, at 138–39 (explaining that not all instincts are emotional and illogical, and some may render the probable cause determination for accurate).

²²² *Id.* at 135.

²²³ *Id.* at 138–139; see also David H. Kaye, *The Paradox of the Gatecrasher and*

about information not provided to the courts, quantifying probable cause might compel a judge to simply hold that a 60% likelihood satisfies probable cause.

Professor Kerr, however, does not address situations where individualized, quantifiable evidence is incorporated into the probable cause inquiry. His objection reflects the problems inherent in using group statistics instead of individualized suspicion to satisfy probable cause.²²⁴ Indeed, many of the objections to quantifying probable cause provide examples involving group probabilities where there is no individualized suspicion.²²⁵ These examples do not pertain to a situation where the quantifiable evidence is sufficiently particularized on its own to answer the probable cause inquiry.

Yet, Professor Kerr's objections may be valid even in cases where individualized suspicion is quantified. Professor Kerr explains that quantification of probable cause could provide a "fertile ground for cognitive biases,"²²⁶ such as the "representativeness heuristic," or "the tendency to measure probability by reference to data that seem to resemble the probability to be estimated."²²⁷ A judge may have some intuition, that she cannot express mathematically, that the statistical evidence tending toward suspicion has not been properly collected or calculated, but she may feel compelled to compare a 49% field accuracy rate presented to a 45% established hurdle and find that probable cause is satisfied.

A thorough understanding of the issues at play would allow a judge to discount certain field accuracy rates as less relevant. For example, if a dog's field accuracy rate is 60%, but a judge finds that the field accuracy data was measured in a way that was not specific enough, she can decide that she is confident enough in that data to allow it to satisfy a 40% probable cause hurdle, if probable cause is defined that way, but not a 50% probable cause hurdle. Alternatively, judges need not accept all field accuracy rates as dispositive of probable cause if there is reason to doubt them. For example, a Massachusetts court of appeals upheld a trial court's finding that a canine alert did not satisfy probable cause because the dog's training was deficient in some areas, and because he had properly detected drugs only five or six times and had improperly alerted

Other Stories, 1979 ARIZ. ST. L.J. 101 (1979).

²²⁴ For the distinction between probabilities relevant to individualized suspicion and group probabilities, see *supra* Part I.D.

²²⁵ Professor Joseph Grano has argued that probable cause should exist where police know that one of a group of ten people have committed a crime, because the 1/10 chance of criminality based on this group statistic "causes us to overlook the success of the police in narrowing their investigation from the universe of all possible suspects, which may include much of the population, to ten individuals." Grano, *supra* note 20, at 496.

²²⁶ Kerr, *supra* note 5, at 132.

²²⁷ *Id.* at 139.

to drugs twice.²²⁸ Defining the probable cause standard would likely not have affected this outcome, as there was so little data that the two false positives alarmed the judge.

Moreover, judges are already finding probable cause based on field accuracy rates, often without fully assessing whether these accuracy rates can be extrapolated into the probable cause inquiry.²²⁹ A judge may just as easily harbor *illogical* intuitive views about a case, such as that certain investigative devices are nearly infallible and their positive alerts should always satisfy probable cause. Requiring quantifiable evidence and quantifying probable cause would undermine judges' discretion to import their unfounded assumptions into a standardless probable cause inquiry. Quantifying probable cause when the evidence is already in probabilistic form would allow a good balance between judicial discretion and judicial transparency/uniformity.

Of course, in cases where a field accuracy rate obviously exceeds the probable cause hurdle, assigning a numerical value to probable cause is less important. Judges do not need to know that a "fair probability"²³⁰ is 48% in order to determine that probable cause is met by a drug sniffing dog with a properly calculated 90% field accuracy rate. However, the Supreme Court should reject the notion that probable cause cannot be defined numerically in anticipation of the marginal cases, which do exist, especially as more technology replaces traditional police work.²³¹ Once the view that probable cause cannot be quantified is rejected, the Supreme Court could either allow lower courts to experiment with selecting a number before ultimately assigning one, or the Court could simply give guidance on what that number is when it confronts a probable cause case where the critical evidence is quantifiable.

2. *When Quantifiable Evidence Is Presented Alongside Other Factors*

When quantifiable evidence must be considered by a court alongside several other significant factors, assigning a numerical value to probable cause has a greater chance of distorting the probable cause inquiry. The concern that courts will overvalue probabilistic evidence becomes more salient when the probabilistic evidence is insufficient on its own to satisfy probable cause. However, assigning probable cause a numerical value would assist judges in understanding the significance of the quantifiable evidence and then determining how much more qualitative evidence is needed to cross the probable cause hurdle.

²²⁸ Commonwealth v. Ramos, 894 N.E.2d 611, 613–14, 617 (Mass. App. Ct. 2008).

²²⁹ See *supra* Part III.

²³⁰ See *supra* Part I.A.

²³¹ See United States v. Sanchez-Tamayo, No. 1:10-CR-0532, 2011 WL 7767740, *14 (N.D. Ga. 2011) ("Courts have held that an accuracy rate of approximately 50%–60% 'in finding measurable amounts of drugs' does not undermine the reliability of the K-9 to the extent of negating probable cause." (quoting United States v. Anderson, 367 F. App'x 30, 33 (11th Cir. 2010))).

For example, consider a case where facial recognition technology's accuracy rate is 42%. The police had a positive alert to a pedestrian by this technology, but the police also had reason to believe that this pedestrian was a wanted bank robber because he appears nervous and walks briskly past police officers. Assuming that a 42% likelihood of criminal activity is not sufficient to satisfy probable cause, assigning a numerical value to probable cause allows a judge to assess how closely the quantifiable evidence comes to being sufficient. The quantifiable evidence's contribution to the probable cause inquiry can be resolved, and now it is up to the judge's discretion and prior case law to determine if the qualitative factors pointing towards suspicion cross the probable cause threshold.²³² Probable cause functions as more of a benchmark to guide judges. If probable cause is set at 60%, a judge will want to see a lot more qualitative evidence than if probable cause is set at 45%.

In cases like these, it is important to note that any extra suspicion, like the suspect's nervousness or brisk walking, may already be reflected in the accuracy rate of the facial recognition technology. Judges must be careful not to double count suspicion, which would occur if the statistical evidence collected on the facial recognition device already reflects the fact that most suspects targeted for use by the technology also appear nervous. However, if this technology is usually used on random individuals, or whenever a passerby resembles a wanted suspect without regard to whether he appears suspicious, then a judge should consider the suspect's nervousness and brisk walking to augment the suspicion calculated by positive alert by facial recognition technology. This type of analysis also applies to drug sniffing dogs and all other technology whose accuracy is measured in the field.

3. *When All of the Evidence Is Qualitative*

When all of the evidence presented is qualitative, or when judges give little weight to the quantifiable evidence, courts should not feel compelled to use the percentage assigned to probable cause to dispose of the probable cause inquiry. However, assigning a numerical value to probable cause can still assist judges in making probable cause determinations, so long as they appreciate that this number serves only as a reference.

This might occur in a case involving an informant with a track record that is reliable but that does not reflect enough data points to be directly extrapolated into the probable cause inquiry. If the court determines based on experience, intuition, and prior case law that the

²³² The *Daubert* inquiry, although not necessary during the probable cause analysis, could serve as a model of how judges may conduct a flexible inquiry, examining several factors, some of which are quantifiable and some of which are not. See *supra* Part III.C; see also *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 141 (1999) (applying *Daubert* "to the testimony of engineers and other experts who are not scientists" because *Daubert's* inquiry is "flexible," and some of its factors may be disregarded or given extra weight by judges).

informant's statements are sufficiently corroborated by the police to satisfy probable cause, the court need not trouble itself with the actual quantum of suspicion assigned as a minimum value to probable cause.²³³ Prior case law would control regardless of the likelihood assigned to probable cause.

Yet, having some conception of the magnitude of the probable cause hurdle can assist judges in evaluating whether an informant's testimony has met the standard. Without feeling bound by the number, or compelled to quantify the inquiry, a judge's assessment of whether a unique set of facts satisfies probable cause may well change depending on whether probable cause is 30% or 50%.

Alternatively, in cases where most or all of the evidence is qualitative, and assigning a precise number to probable cause is unhelpful, judges could conceptualize probable cause as a range, as Professor Bacigal proposes.²³⁴ This would alleviate some of the confusion surrounding probable cause, without fostering many of the cognitive biases that arise when flexible inquiries are made overly rigid. This range could place as its center the number assigned to probable cause for cases where quantifiable evidence is the critical factor demonstrating probable cause.

When quantifiable evidence and unquantifiable evidence are both essential to the probable cause inquiry, or when all of the evidence is qualitative, flexibility in application of the probable cause standard is important. Quantifying probable cause may undermine this flexibility if judges ignore their intuitions or convert evidence that is not susceptible to quantification into fictional percentages. Courts must be advised about the pitfalls inherent in quantifying the standard.

B. *Selecting the Number*

The selection of the number associated with the probable cause hurdle will be a difficult endeavor. The number assigned must reflect the fact that probable cause balances the degree of suspicion at which law enforcement interests trump privacy interests.²³⁵ Additionally, the numerical threshold must be a percentage that judges can easily abstract and conceptualize in their minds.

Although there is wide variance regarding what this percentage is, a significant number of courts and scholars assume that probable cause is within the 40% to 51% range.²³⁶ In case assigning probable cause a

²³³ See *United States v. Searcy*, 664 F.3d 1119, 1121–22 (7th Cir. 2011) (upholding the search of defendant's home where confidential informant's direct observations of defendant and reliable track record compensated for the absence of detail provided by the informant).

²³⁴ See Bacigal, *supra* note 20, at 338–39 (placing a “fair probability” within the 40% to 49% range).

²³⁵ See *supra* Part I.A.

²³⁶ See *supra* note 61 and accompanying text; see also Daniel A. Crane, *Rethinking Merger Efficiencies*, 110 MICH. L. REV. 347, 356 (2011) (“The Supreme Court

numerical value will dilute its power and mystery, and consequently the privacy protections that it affords, probable cause should be placed towards the higher end of this range. Many courts have assumed that probable cause does not require a preponderance of the evidence in favor of suspicion,²³⁷ but this standard can be conceptually helpful in selecting a benchmark.

In Fourth Amendment jurisprudence, in contexts unrelated to probable cause, judges frequently apply the preponderance-of-the-evidence standard.²³⁸ This standard has an approximate numerical percentage of just above 50%; some describe it as akin to “[50%] and a feather,”²³⁹ others deem it to be 51%.²⁴⁰ One solution to assigning probable cause a number that is easy to conceptualize is to select the other side of the preponderance standard, so that probable cause is satisfied at just slightly lower than 50%. For courts, then, a “fair probability” would occur where the evidence tending towards suspicion is almost at the point where it is in equipoise, but may be slightly less. Because judges already have a sense of when evidence tips slightly above 50%, they can also appreciate when evidence slips slightly below 50%, or the other side of the preponderance coin.

Through the stages of review, slightly below 50% would be used as a benchmark. At the lowest level of the inquiry, an “officer of reasonable caution”²⁴¹ must be able to believe that evidence of a crime would be found by at least 50% (this is, after all, what it usually means to have a belief), but a magistrate judge issuing a warrant or a court examining a warrantless search of an automobile would assign as the “fair probability”²⁴² a value slightly below 50%. Alternatively, judges could place some band of error around 50% to determine whether a reasonable police officer could have considered the likelihood of

has denied that probable cause is susceptible of ‘precise definition or quantification into percentages,’ but practitioners and commentators often understand it to lie in the 40–45 percent range.” (quoting *Maryland v. Pringle*, 540 U.S. 366, 371 (2003)) (footnote omitted)).

²³⁷ See *supra* note 62.

²³⁸ For example, in *Nix v. Williams*, the Supreme Court held that courts should apply a preponderance-of-the-evidence test to the “inevitable discovery rule,” which allows the admission of evidence obtained after a Fourth Amendment violation if it would have been discovered eventually through lawful means. 467 U.S. 431, 444 (1984) (“If the prosecution can establish by a preponderance of the evidence that the information ultimately or inevitably would have been discovered by lawful means . . . the evidence should be received.”).

²³⁹ E.g., Gordon Shemin, *Mercury Rising: The Omnibus Autism Proceeding and What Families Should Know Before Rushing Out of Vaccine Court*, 58 AM. U. L. REV. 459, 476 n.84 (2008).

²⁴⁰ E.g., Jean Macchiaroli Eggen, *Toxic Reproductive and Genetic Hazards in the Workplace: Challenging the Myths of the Tort and Workers’ Compensation Systems*, 60 FORDHAM L. REV. 843, 893 (1992).

²⁴¹ See *supra* Part I.

²⁴² *Id.*

criminal behavior to be 50%.²⁴³ Then, reviewing courts decide if there is a “substantial basis”²⁴⁴ for concluding that this threshold of slightly below 50%, the inferiority of the evidence, is met.

A serious consequence of assigning probable cause one particular percentage is that it will force courts to apply the same probable cause value regardless of the severity of the crime. The Court currently maintains, according to Professor Bacigal, “the fiction of one uniform definition of probable cause.”²⁴⁵ However, some justices and lower courts have insinuated that the probable cause standard should be a sliding scale, depending on the gravity of the offense.²⁴⁶ Justice Robert Jackson wrote in a dissenting opinion that

[i]f we assume, for example, that a child is kidnaped and the officers throw a roadblock about the neighborhood and search every outgoing car, it would be a drastic and indiscriminating use of the search. The officers might be unable to show probable cause for searching any particular car. However, I should candidly strive hard to sustain such an action, executed fairly and in good faith, because it might be reasonable to subject travelers to that indignity if it was the only way to save a threatened life and detect a vicious crime. But I should not strain to sustain such a roadblock and universal search to salvage a few bottles of bourbon and catch a bootlegger.²⁴⁷

There are important reasons why the probable cause hurdle should not vary depending on the severity of the crime uncovered. A variable probable cause standard might enable courts to use hindsight judgment to uphold or invalidate searches, depending on how severe the criminal activity ultimately turned out to be. Having probable cause depend on balancing the severity of the crime against the nature of the intrusion would also convert probable cause into a more general “reasonableness” standard, whose malleability is even more susceptible to manipulation based on the concerns of the day than the harder line of the probable cause standard.²⁴⁸

²⁴³ An analogue occurs in the context of conditional probabilities, where evidence is probative only if a precondition is met. For example, the fact that a murder victim cheated on the defendant is relevant to prove motive only if the defendant was aware of the cheating. In those cases, a judge will admit evidence of the cheating if a reasonable jury could have found that the defendant was aware of this fact by a preponderance of the evidence. *See* FED. R. EVID. 104(b).

²⁴⁴ *See supra* Part I.

²⁴⁵ Bacigal, *supra* note 20, at 323.

²⁴⁶ *See id.* at 324 (quoting Judge Posner’s comment in *Llaguno v. Minge*, 763 F.2d 1560, 1565 (7th Cir. 1985), that “probable cause—the area between bare suspicion and virtual certainty—describes not a point but a zone, within which the graver the crime the more latitude the police must be allowed.”).

²⁴⁷ *Brinegar v. United States*, 338 U.S. 160, 183 (1949) (Jackson, J., dissenting).

²⁴⁸ According to several scholars and one retired judge, “the malleable concept of ‘reasonableness’ once unleashed from its textual moorings, found in the

There are also arguments for allowing flexibility in the probable cause hurdle depending on the nature of the crime, as Justice Jackson's kidnapping example illustrates.²⁴⁹ However, if the Court wants the probable cause standard to vary depending on the situation, it should hold so explicitly, instead of purporting to apply a "single, familiar standard."²⁵⁰ The Supreme Court should truly grapple with whether the severity of the crime should affect probable cause, instead of enabling courts to alter the standard at their discretion, without transparency.²⁵¹

C. Other Solutions

There are, of course, other solutions to incorporating quantifiable evidence besides assigning a minimum percentage of suspicion necessary to satisfy probable cause. Courts could, for instance, assign maximum false positive error rates and conditions of usage in order to allow a positive alert to be relevant to the probable cause inquiry. Or, courts could simply mandate that the use of all devices that essentially replace police officers must be preceded by uniform, standardized training and testing. These solutions would enable a more rigorous probable cause inquiry to preclude what many courts are currently doing, which is presuming that positive alerts by particular devices and investigative techniques are sufficient, even in the absence of data, to satisfy probable cause.

Alternatively, if the Supreme Court decides that courts require the flexibility to alter the probable cause standard depending on the severity or exigency of the crime, a range can be assigned to probable cause. Professor Ronald Bacigal proposes quantifying probable cause into a range of 40% to 49% for precisely this reason.²⁵² Even providing a range to probable cause would add more uniformity to the inquiry and assist judges in better assessing quantifiable evidence presented as part of the probable cause inquiry.

V. CONCLUSION

As more quantifiable evidence is incorporated into the probable cause inquiry, courts must grapple seriously with how they assess that

Amendment's probable cause and warrant requirements, can all too quickly and unnecessarily diminish the compass of liberty and privacy. What is 'reasonable,' after all, is subject to the fears and prejudices of the day." Laurence A. Benner et al., *Criminal Justice in the Supreme Court: An Analysis of United States Supreme Court Criminal and Habeas Corpus Decisions (October 2, 2000–September 30, 2001)*, 38 CAL. W. L. REV. 87, 187 (2001).

²⁴⁹ See *supra* note 247 and accompanying text.

²⁵⁰ *Dunaway v. New York*, 442 U.S. 200, 213 (1979).

²⁵¹ For an interesting proposal on how to incorporate crime severity into the probable cause analysis, see Jeffrey Bellin, *Crime-Severity Distinctions and the Fourth Amendment: Reassessing Reasonableness in a Changing World*, 97 IOWA L. REV. 1 (2011).

²⁵² See Bacigal, *supra* note 20, at 338–39.

evidence. Courts often make the mistake of presuming that certain fallible investigative devices are sufficient to establish probable cause without properly calculated evidence of their reliability. Or, if accuracy rates are admitted as evidence to demonstrate reliability, courts often do not understand how to evaluate these accuracy rates to best answer the probable cause question. This Article has endeavored to explain which types of evidence require accuracy rates, how to calculate and assess those accuracy rates, and finally how to approach the probable cause inquiry using quantifiable evidence. In the process, it has argued that probable cause should be assigned a numerical value in cases where quantifiable evidence is the critical basis on which probable cause is satisfied.