

Putting Probability Back into Probable Cause

Max Minzner*

When deciding whether a baseball player is likely to get a hit, we look at his history of success at the plate. When deciding whether to listen to the advice of a stock analyst, we look at whether the prices of her past recommendations rose or fell. But when police officers claim that they have probable cause to believe a certain location contains evidence of a crime, we do not look at whether they have been right or wrong when they have made the same claim in the past. This is a mistake.

Law-enforcement-search success rates vary widely, even when the same legal standard applies. Searches pursuant to warrants issued on a probable-cause standard recover evidence at very high rates, usually exceeding 80%. By contrast, warrantless searches, even when officers allege they have probable cause, succeed at far lower rates, recovering evidence as infrequently as 12% of the time. Similarly, some officers are far more successful than others when they conduct probable-cause searches. Some almost never recover evidence; some almost always do.

What role should these different success rates play in probable-cause analysis? The current answer is none. Judges are not presented with the success rates of the law enforcement officers who appear before them. I argue that law enforcement should be forced to present success-rate data to judges when making probable-cause claims and judges should be allowed to consider the data when deciding whether to issue a warrant or approve a previously conducted search. These success rates capture information not currently analyzed in the search process and their addition would improve the accuracy of probable-cause decisions. Most significantly, we would learn private information in the possession of law enforcement that is not currently presented to judges.

* Assistant Professor, Benjamin N. Cardozo School of Law, Yeshiva University. The author would like to thank Mitchell Berman, Stephen Morse, Margaret Lemos, Alex Reinert, and the presenters and commentators in the Yale–Stanford Junior Faculty Forum, the Cardozo School of Law Junior Faculty Forum, and the University of New Mexico Faculty Workshop.

I. Introduction

You are a magistrate judge. Officer A is seeking a search warrant for a house.¹ He knows the house belongs to an individual with a history of narcotics trafficking. When the officer went to the home and interviewed the owner, the owner seemed nervous and appeared to be concealing something behind his back. Officer A tells you that, in his view, this information constitutes probable cause to believe that the house contains narcotics.

Officer B performed a search of an automobile as part of a routine traffic stop. When he pulled the car over, the officer ran the license plates and learned that the car's owner was a known drug dealer. When he approached the car, the driver seemed nervous and appeared to be concealing something beneath his seat. Officer B searched the car and found narcotics.² He claims the evidence should be admissible because he had probable cause to believe there were narcotics in the vehicle.

These cases may appear similar and may seem close to the probable-cause threshold but are, in fact, very different. There is far more reason to rely on Officer A's judgment. When he believes he has probable cause, he is usually right—he recovers evidence almost 90% of the time.³ On the other hand, Officer B is almost always wrong—he recovers evidence just under 12% of the time.⁴ Searches pursuant to a warrant are far more likely to recover evidence than warrantless searches.

Such variation does not simply exist between different *types* of searches. Different officers conducting the *same* type of search succeed at very different rates. Here we can take as an example two real officers of the Florida State Police.⁵ Between January 2000 and September 2001, each performed a similar number of probable-cause searches of automobiles: Officer A conducted eighteen searches, while Officer B conducted fifteen searches. When Officer A thought he was likely to recover evidence, he was almost always wrong. Only one of his eighteen searches led to a seizure—a success rate of 5.6%. Officer B, by contrast, was almost always right. He

1. Law enforcement generally needs a search warrant before searching a home. “[O]nly in a few specifically established and well-delineated situations . . . may a warrantless search of a dwelling withstand constitutional scrutiny, even though the authorities have probable cause to conduct it.” *Vale v. Louisiana*, 399 U.S. 30, 34 (1970) (internal citation and quotations omitted).

2. Law enforcement generally may search automobiles as part of a routine traffic stop without a warrant as long as officers have probable cause to believe the vehicle contains contraband. *Carroll v. United States*, 267 U.S. 132, 149 (1925).

3. Laurence A. Benner & Charles T. Samarkos, *Preliminary Findings from the San Diego Search Warrant Project*, 36 CAL. W. L. REV. 221, 249–50 (2000).

4. MATTHEW R. DUROSE ET AL., BUREAU OF JUSTICE STATISTICS, CONTACTS BETWEEN POLICE AND THE PUBLIC: FINDINGS FROM THE 2002 NATIONAL SURVEY 13 (2005) [hereinafter 2002 NATIONAL SURVEY].

5. See Florida State Highway Patrol Data (on file with the author) (data for officers 669 and 1,577). This data set is discussed further in Part IV. See *infra* notes 74–75 and accompanying text. Many thanks to Nicola Persico for making this data available to me.

recovered evidence in thirteen of the fifteen searches, for an 86.7% success rate.

What role should these historic success rates play in the probable-cause analysis? The answer under present law is none at all. Law enforcement's history of prior success is irrelevant. I argue that this is a mistake. Law enforcement should be required to report the success rates of both the individual officer performing the search and the relevant law enforcement unit when making claims about probable cause. These rates include valuable information that is currently excluded from the probable-cause analysis but that could substantially affect the likelihood of recovering evidence.

Subpart II(A) provides a general background of the search process—how law enforcement officers either obtain authorization to execute a search in advance or obtain validation of a search after it has occurred. Subpart II(B) explains when prior success rates will aid the judicial determination. Success rates provide a mechanism to make use of information that currently is absent from the probable-cause analysis or is incorporated ineffectively. In particular, looking at the prior success rates of law enforcement agencies and individual officers will aid the probable-cause determination when information affecting the likelihood of recovering evidence is currently excluded from the process.

Parts III and IV examine two particular types of currently excluded information that strongly affect the likelihood that a search will recover evidence. First, Part III argues that pressures on law enforcement, particularly the cost of searches, heavily influence search success rates. Next, Part IV looks at information relating to the types of officers that would be revealed by considering success rates. Notably, hit rates provide a means to detect police perjury.

Part V reviews the potential law enforcement responses to a data-collection and reporting requirement. Part VI examines the doctrinal support for the use of success rates. While courts have not required officers to report their success rates when making probable-cause claims, they have been willing to consider the history of prior success in components of the probable-cause analysis, including informants and canines.

Part VII considers standard objections in the academic literature to the use of statistical measures in legal decision making. Commentators have generally been opposed to the use of overt statistical measures, but crucially they have only considered the effect of these measures in the trial context. I argue that the search context is fundamentally different for two reasons. First, the probable-cause determination is explicitly and exclusively a statement about the probability of a particular outcome—namely, the odds of recovering evidence from a particular location. Trials, on the other hand, have multiple goals in addition to verdict accuracy. Second, trials are a statement about events in the past and usually represent the last word on their subject matter. We seldom learn the truth about what actually happened. By

contrast, at the conclusion of the search process we learn whether the initial prediction that the location would contain contraband was correct.

Finally, Part VIII considers the consequences of a shift to an outcome-oriented model of probable cause. Because search warrants currently recover evidence at rates far higher than the probable-cause standard, such an approach would almost certainly lead to more search warrants. I argue that this would be a positive outcome.

II. Conducting Searches

A. *The Traditional Probable-Cause Calculation*

Law enforcement can perform searches in two ways. Officers can either get a warrant in advance from a magistrate or attempt to defend the decision to search after the fact by convincing the court that a warrant was unnecessary. When law enforcement is required to get a warrant, agents gather information about a particular target and determine whether they have enough evidence to convince a judge that there is probable cause to believe that the target contains evidence of a crime. If they decide they do have sufficient evidence, they will either draft a warrant application themselves or enlist the assistance of a prosecutor.⁶ Once the warrant application is completed, it is presented to a magistrate judge who approves it, denies it, or requests more information.⁷ If the warrant is approved, the officers have a certain period of time in which to execute the warrant and complete the search.⁸ If law enforcement officers rely in good faith on a validly issued search warrant, any evidence seized is protected from later suppression.⁹

Alternatively, law enforcement can proceed without a warrant and search without prior judicial authorization. While the Fourth Amendment is

6. Warrant affidavits need not be drafted by prosecutors and indeed often are not. *See* *United States v. Ventresca*, 380 U.S. 102, 108 (1965) (“[Warrant affidavits] are normally drafted by nonlawyers in the midst and haste of a criminal investigation.”). Prosecutorial involvement varies substantially based on the jurisdiction, and can range from simply reviewing the document written by the officer applying for the warrant to actually drafting the warrant affidavit itself. RICHARD VAN DUIZEND ET AL., *THE SEARCH WARRANT PROCESS* 20 (1985).

7. RICHARD VAN DUIZEND ET AL., *supra* note 6, at 22–23. Review of the warrant application by a “neutral and detached magistrate” is a core requirement of the Fourth Amendment. *See, e.g.*, *United States v. U.S. Dist. Court*, 407 U.S. 297, 318 (1972) (“Prior review by a neutral and detached magistrate is the time-tested means of protecting Fourth Amendment rights.”); *Johnson v. United States*, 333 U.S. 10, 14 (1948) (“When the right of privacy must reasonably yield to the right of search is, as a rule, to be decided by a judicial officer, not by a policeman or Government enforcement agent.”).

8. *See* FED. R. CRIM. P. 41(e)(2)(A) (imposing a ten-day limit on the validity of a warrant); *Sgro v. United States*, 287 U.S. 206, 208–12 (1932) (suppressing evidence seized in a search that was not conducted within the ten-day limit).

9. *See* *United States v. Leon*, 468 U.S. 897, 918–23 (1984) (establishing the good faith exception to the exclusionary rule).

generally described as expressing a preference for warrants,¹⁰ searches without warrants can occur in a large number of contexts. Consent of the person searched, of course, provides virtually a blanket exception to both the probable-cause and warrant requirements,¹¹ but law enforcement can engage in nonconsensual warrantless searches in certain situations. For instance, exigent circumstances can excuse the failure to obtain a warrant.¹²

If a warrantless seizure takes place, law enforcement will have to defend the decision to search to a judge in a suppression hearing. The seizing officer will testify to the facts that led him to believe that a search was appropriate. If the judge agrees that the officer possessed probable cause to believe that the search would produce evidence and that an exception to the warrant requirement applied, the evidence will be admissible. Otherwise, the exclusionary rule will require suppression of the evidence in the subsequent criminal prosecution.¹³

Whether a search takes place with or without a warrant, an essential question about the underlying facts is whether there was probable cause to believe that the location to be searched contained evidence of a crime.¹⁴ In both of these situations, judges determine the presence or absence of probable cause based on a narrative statement by the law enforcement officer. Law enforcement is expected to articulate reasons for a search and those reasons are expected to be specific to the target.

10. *E.g.*, *Ornelas v. United States*, 517 U.S. 690, 699 (1996) (“The Fourth Amendment demonstrates a ‘strong preference for searches conducted pursuant to a warrant,’ . . .” (quoting *Illinois v. Gates*, 462 U.S. 213, 236 (1983))).

11. *See* *Schneekloth v. Bustamonte*, 412 U.S. 218, 219 (1973) (“It is . . . well settled that one of the specifically established exceptions to the requirements of both a warrant and probable cause is a search that is conducted pursuant to consent.”); *Amos v. United States*, 255 U.S. 313, 317 (1921) (finding that a wife’s consent to a warrantless search was not sufficient to waive the constitutional rights of her husband).

12. *See* *Brigham City v. Stuart*, 547 U.S. 398, 403 (2006) (noting that exigent circumstances may be so compelling that a warrantless search is objectively reasonable under the circumstances); *Mincey v. Arizona*, 437 U.S. 385, 393–94 (1978) (“[W]arrants are generally required to search a person’s home or his person unless ‘the exigencies of the situation’ make the needs of law enforcement so compelling that the warrantless search is objectively reasonable under the Fourth Amendment.” (quoting *McDonald v. United States*, 335 U.S. 451, 456 (1948))).

13. *See* *Mapp v. Ohio*, 367 U.S. 643, 660 (1961) (extending the exclusionary rule to apply to state law enforcement); *Weeks v. United States*, 232 U.S. 383, 398 (1914) (applying the exclusionary rule to federal law enforcement).

14. *See, e.g.*, *New Jersey v. T.L.O.*, 469 U.S. 325, 340 (1985) (“Ordinarily, a search—even one that may permissibly be carried out without a warrant—must be based upon ‘probable cause’ to believe that a violation of the law has occurred.”); *Almeida-Sanchez v. United States*, 413 U.S. 266, 273 (1973) (distinguishing between a routine search of a vehicle at a border checkpoint and an unconstitutional search conducted without probable cause or consent by a roving patrol twenty miles from the border). Several authors, Akhil Amar chief among them, have argued that the probable-cause standard should only apply to searches on warrants and a more forgiving reasonableness standard should cover warrantless searches. Akhil Reed Amar, *Fourth Amendment First Principles*, 107 HARV. L. REV. 757, 782 (1994).

This approach extends back to *United States v. Nathanson*.¹⁵ In *Nathanson*, the magistrate authorized a Prohibition Era search warrant for “liquors of foreign origin” based on the naked allegation that the federal agent “[had] cause to suspect and [did] believe” that the contraband was located in the defendant’s home.¹⁶ The Supreme Court faced the question of whether such a spare affidavit could support a probable-cause finding. Justice McReynolds’s majority opinion rejected the government’s argument that “mere affirmation of suspicion or belief without disclosure of supporting facts” was sufficient.¹⁷ The magistrate must base the probable-cause determination on “facts or circumstances presented to him under oath or affirmation.”¹⁸

In *Aguilar v. Texas*,¹⁹ the Court extended this requirement of articulation to focus particularly on the use of confidential informants.²⁰ The Supreme Court in *Aguilar* also accepted the use of hearsay testimony from informants as a basis for a warrant application,²¹ but rejected a warrant application that merely related that law enforcement had “received reliable information from a credible person” that narcotics were located in the target premises.²² The Court stated, “Here the ‘mere conclusion’ that petitioner possessed narcotics was not even that of the affiant himself; it was that of an unidentified informant.”²³ The Court established the rule that the warrant application must include “some of the underlying circumstances from which the informant concluded that the narcotics were” present, and “some of the

15. 290 U.S. 41 (1933).

16. *Id.* at 44.

17. *Id.* at 47.

18. *Id.*; see also *Byars v. United States*, 273 U.S. 28, 29 (1927) (stating that a warrant similar to the one in *Nathanson* “clearly is bad if tested by the Fourth Amendment and the laws of the United States”).

19. 378 U.S. 108 (1964).

20. Between deciding *Nathanson* and *Aguilar*, the Supreme Court returned to the articulation issue in *Giordenello v. United States*, 357 U.S. 480 (1958). *Giordenello* involved a challenge to an arrest warrant under Federal Rule of Criminal Procedure 4, which permits the issuance of a warrant if “there is probable cause to believe that an offense has been committed.” *Id.* at 487 n.3. The Court concluded that the affidavit did not establish probable cause because the complaint “[did] not indicate any sources for the complainant’s belief; and it [did] not set forth any other sufficient basis upon which a finding of probable cause could be made.” *Id.* at 486. Lacking that information, the magistrate could not “assess independently the probability that petitioner committed the crime charged.” *Id.* at 486–87. Despite the slightly different procedural context, the *Giordenello* holding has been used in Fourth Amendment search jurisprudence as well. See, e.g., *Aguilar*, 378 U.S. at 113 (concluding that, under *Giordenello*, an affidavit was insufficient because it “contained no affirmative allegation that the affiant spoke with personal knowledge of the matters contained therein”).

21. *Aguilar*, 378 U.S. at 114; see also *Jones v. United States*, 362 U.S. 257, 269 (1960) (“An affidavit is not to be deemed insufficient [because it relies on hearsay] so long as a substantial basis for crediting the hearsay is presented.”).

22. *Aguilar*, 378 U.S. at 109.

23. *Id.* at 113.

underlying circumstances from which the officer concluded that the informant . . . was ‘credible’ or his information ‘reliable.’”²⁴

Early on, the Court struggled to apply this two-pronged test, which focused on the informant’s veracity and basis of knowledge.²⁵ The Court subsequently identified a variety of circumstances that either would or would not permit an informant’s tip to sufficiently support probable cause. For instance, officers could establish credibility through the informant’s past performance,²⁶ by the self-incriminating nature of the informant’s information,²⁷ or by partial corroboration of the information provided by the informant.²⁸ All of these methods assume that by adding this additional information to the content of the informant’s tip, the magistrate can calculate the actual probability of recovering evidence.

The Supreme Court moved away from this rigid framework in *Illinois v. Gates*²⁹ and folded informant tips into a “totality of circumstances” test that now applies to all probable-cause decisions.³⁰ Law enforcement presents a collection of facts to the magistrate in an effort to convince the judge that the standard of proof exceeds the probable-cause threshold. The magistrate considers all of the collected facts and decides whether the allegations are sufficient. In the classic language of *Gates*, the Court held:

The task of the issuing magistrate is simply to make a practical, common sense decision whether, given all of the circumstances set forth in the affidavit before him, . . . there is a fair probability that contraband or evidence of a crime will be found in a particular place.³¹

24. *Id.* at 114.

25. *See, e.g., Spinelli v. United States*, 393 U.S. 410 (1969). *Spinelli* applied the articulation requirement to a search-warrant application that included an anonymous tip that the defendant had been operating a gambling business using particular telephone numbers. *Id.* at 414–15. Despite the fact that the defendant was observed entering and leaving the apartment where the telephones were subscribed, the Supreme Court found that probable cause was lacking. *Id.* at 414, 418.

26. *See McCray v. Illinois*, 386 U.S. 300, 303–04 (1967) (holding that there was probable cause where two police officers testified that information from the informant had previously resulted in arrests and convictions).

27. *See United States v. Harris*, 403 U.S. 573, 583 (1972) (“Admissions of crime, like admissions against proprietary interests, carry their own indicia of credibility—sufficient at least to support a finding of probable cause to search.”).

28. *See Draper v. United States*, 358 U.S. 307, 313 (1959) (accepting as probative of probable cause the fact that an officer had “personally verified every facet of the information given” by the informant except whether the target actually possessed the expected contraband).

29. 462 U.S. 213 (1983).

30. *See id.* at 230–31 (“This totality-of-the-circumstances approach is far more consistent with our prior treatment of probable cause than is any rigid demand that specific ‘tests’ be satisfied by every informant’s tip.”).

31. *Id.* at 238.

The Court has assumed that by aggregating the value of the information presented, magistrates can reach an informed decision about the likelihood of recovering evidence.³²

B. *Incorporating Hit Rates*

Nothing in *Gates* precludes incorporating law enforcement's prior success rates into the Fourth Amendment analysis. Indeed, the "common sense" decision would be to use all relevant information. However, the fact that the Court has focused on the need to require law enforcement to articulate individuated reasons for each search has meant that this type of statistical evidence is not required.³³

My proposal is to impose an obligation to report success rates. Law enforcement should be obliged to report two particular success rates, and judges should incorporate the data into the probable-cause analysis. First, the court should look at the history of the officer's law enforcement unit for the relevant type of search. Second, the court should consider the history of prior success of the officer conducting the search. Judges will then increase

32. This approach essentially expects courts to make Bayesian calculations. Bayes Theorem provides the standard statistical model for updating a prior probabilistic belief when one is presented with new evidence. ANDREW GELMAN ET AL., *BAYESIAN DATA ANALYSIS* 8 (2d ed. 2004). In formal terms, it is conventional to represent the probability (P) of a location containing contraband (C) given certain evidence (E) as $P(C | E)$. Bayes Rule states:

$$P(C | E) = P(C) * \left[\frac{P(E | C)}{P(E)} \right]$$

Despite its apparent symbolic complexity, the rule is rather intuitive. Indeed, "[i]t has become common to think of Bayes Theorem as providing 'a quantitative description of the ordinary process of weighing evidence.'" Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329, 1351 n.69 (1971) (quoting IRVING JOHN GOOD, *PROBABILITY AND THE WEIGHING OF EVIDENCE* 62 (1950)). The term $P(C)$ is commonly called the "prior." It represents the probability of finding contraband without any additional evidence. The ratio of $P(E | C)$ to $P(E)$ "updates" the prior by determining how much weight the new evidence should receive in changing the prior belief. The ratio reflects the probative force of the evidence by comparing the probability of the evidence's existing if there is contraband, $P(E | C)$, to the probability of the evidence's existing whether or not there is contraband, $P(E)$. It provides the relative strength of the new evidence—how much it should affect one's previous beliefs.

33. Professor Rosenthal has put it more forcefully in his arguments for a greater use of empirical evidence in Fourth Amendment cases: "[I]t is striking that the Supreme Court, with only rare exceptions, . . . neither looks at nor insists on empirical evidence." 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). Rosenthal convincingly argues that the effectiveness of law-enforcement techniques should play an important role in determining the Fourth Amendment calculus. *See id.* at 681 (arguing that new, very successful police tactics should result in courts' being prepared to reassess their views of what amounts to reasonable search and seizure in the constitutional sense, based on the empirical evidence of police efficacy compiled over the past decade). Professors Meares and Harcourt have made a similar argument about the importance of empirical evidence in constitutional criminal procedure, using both *Miranda v. Arizona*, 384 U.S. 436 (1966), and *Terry v. Ohio*, 392 U.S. 1 (1968), as case studies. *See* Tracey L. Meares & Bernard E. Harcourt, 90 J. CRIM. L. & CRIMINOLOGY 733, 735 (1999) (calling for a mode of judicial decision making that incorporates empirical assessment as an important factor, thereby making criminal-procedure decisions more transparent).

their scrutiny of search types that are particularly unlikely to recover evidence and officers who, especially in relative terms, are less capable than their colleagues. By analyzing these data points, both together and separately, I argue that the probable-cause decision can become far more accurate.

When will this data improve the accuracy of the search process? The answer is straightforward: The data will improve the accuracy of the probable-cause determination if two conditions hold. First, there must be a relationship between these data and the likelihood of recovering evidence in the current case.³⁴ Second, the reasons for that relationship must currently be excluded from the probable-cause analysis.³⁵ Put more simply, the differences in success rates must be driven by the same causes in case after case, and those causes must not be already apparent to magistrate judges.

This first condition is relatively intuitive. Historic success rates will help if and only if they tell us which officers are likely to recover evidence in the future. If officers or law enforcement units with high prior success rates are more likely to recover evidence than those with historically lower prior success rates, these data will improve the probable-cause analysis. The second requirement may be less obvious. If the reasons for the differences in the past success rates are already apparent to magistrate judges when they engage in the probable-cause determination, adding success rates will not matter. For example, consider two officers, one who succeeds on 60% of his search warrants and another who succeeds on 90% of her warrants. If the reason for the differential success rates is simply that the second officer always has stronger evidence and the difference in evidentiary strength is always apparent on the face of the warrant application, including these success rates will not improve the probable-cause decision. All of the relevant information is already before the judge.³⁶

Another way to ask this question is to inquire whether information that affects success rates is currently excluded from the judicial analysis. Imagine the two searches described in the Introduction. Law enforcement

34. In Bayesian terms, this requirement means that the probability of recovering evidence is higher if the magistrate knows about the historic success rate than if he does not, i.e., $P(C | R) > P(C)$, where R represents the relevant success rate. If $P(C | R) = P(C)$, we say that C is *independent* of R . If they are independent, knowing the success rate does not change the probability of recovering contraband.

35. In Bayesian terms, this requirement means that knowing R , along with all of the other evidence in the case, increases the likelihood of recovering contraband when compared to simply knowing the other evidence, i.e., $P(C | R, E) > P(C | E)$. If $P(C | R, E) = P(C | E)$, we say that the probability of recovering contraband is *conditionally independent* of the success rate given the evidence. Once we know the evidence, the recovery rate no longer matters. See, e.g., A.P. David, *Conditional Independence in Statistical Theory*, 41 J. ROYAL STAT. SOC'Y 1, 4 (1979) (explaining that when $P(C | R, E) = P(C | E)$, the parameter E "encapsulates the *sufficiency principle*," meaning no information is contained in C over and above that contained in E).

36. For this reason, if C is conditionally independent of R once E is known, it is sometimes said that the evidence "screens off" the success rate from the recovery rate.

presents similar information to a magistrate, but the actual likelihood of recovering evidence is much higher under one set of facts than the other. Alternatively, if law enforcement either lacks information relevant to the likelihood of recovering evidence or possesses information and fails to disclose it, the current model does not have an effective way to incorporate the missing information. As an example, imagine an officer who claims falsely that a suspect engaged in surveillance-conscious behavior and, alternatively, imagine an officer who makes the same claim in subjectively good faith but is a poor judge of what constitutes surveillance-conscious behavior. The former problem is revealed only in the unlikely case where the perjury is discovered. There is no mechanism for additional scrutiny of the latter problem.

Analyzing success rates of search types and officers, on the other hand, will reveal these problems to the extent that they consistently appear for the same search type or the same officers. Officers who rely on perjured testimony or poor judgment will see their success rates fall and the scrutiny of their warrant applications increase. Search types with low success rates, even if the apparent evidence is equally strong, will get the additional scrutiny they deserve.

The following Parts analyze two types of omitted information that are overlooked by the current probable-cause calculation but revealed by considering hit rates. Part III argues that pressures at the law enforcement level, especially budget pressures, are a primary determinant of the likelihood of a search recovering evidence. The disparity in success rates between searches with a warrant and warrantless searches is a function of budget constraints. Law enforcement officers rarely obtain search warrants because searches are so expensive. Officers husband their resources carefully and seek warrants only when they are particularly likely to succeed. As a result, law enforcement success rates vary based on the cost of the search and the budget of the law enforcement entity. This information is highly relevant to the probable-cause determination because it implies that cheap searches need more judicial scrutiny than more expensive ones; however, this information is currently excluded from the probable-cause process.

Part IV argues that information specific to the law enforcement officer in question may substantially affect the likelihood of recovering evidence. In particular, success rates provide a way to detect racial profiling or police perjury, problems currently difficult or impossible to identify.

III. Hidden Facts: Types of Searches

A. *Searches with and Without Warrants*

As discussed in subpart II(A), the core probable-cause question for searches, with or without warrants, is essentially the same. However, while the probable-cause analysis is similar, the actual outcomes of the searches are very different. There are two primary studies about the effectiveness of

searches that take place when a warrant has been obtained. In 1985, Richard van Duizend published a seminal study on the search-warrant processes of seven American jurisdictions.³⁷ Identified only by pseudonyms, the jurisdictions studied had very high recovery rates for search warrants, ranging from 84% to 97%.³⁸ In 2000, Laurence Benner and Charles Samarkos published a similar study of narcotics search warrants in San Diego³⁹ and found similarly high recovery rates for search warrants. Eighty-eight percent of the narcotics warrants in their data set led to a seizure.⁴⁰ Note that the search warrants did not merely lead to the seizure of “something”—the overwhelming majority led to the seizure of an item listed in the warrant.⁴¹ It was the rare warrant where the seizure involved only items not listed in the warrant application.⁴²

What happens when law enforcement officers perform a warrantless probable-cause search? These searches are far more difficult to study. For instance, when an officer can perform a search based on probable cause in exigent circumstances, there is generally no record of the search if it turned out to be fruitless. However, some data relating to automobile searches have become available as a result of the recent interest in unearthing and preventing racial profiling by law enforcement. The Supreme Court has made clear that search warrants generally are not required to search an automobile.⁴³ As a result, we can analyze automobile searches as a method of determining the

37. VAN DUIZEND ET AL., *supra* note 6.

38. *Id.* at 38 tbl.21. Van Duizend noted that his results may include inflated figures for the seizure rates in certain jurisdictions because of a failure to file returns when the warrants did not lead to a recovery. If we assume that all of the unreturned warrants are failures, the success rate only falls dramatically in one jurisdiction, River City. The River City success rate falls to 46%, while Harbor City (76%), Plains City (79%), Forest City (91%), Hill City (78%), Mountain City (93%), and Border City (81%) each still have a success rate over 75%. Van Duizend’s results also have been subject to some questions because the studied warrants were not generated through purely random sample. Craig D. Uchida & Timothy S. Bynum, *Search Warrants, Motions to Suppress and “Lost Cases”: The Effects of the Exclusionary Rule in Seven Jurisdictions*, 81 J. CRIM. L. & CRIMINOLOGY 1034, 1048 (1991). Uchida and Bynum, as part of their study of the exclusionary rule, returned to the seven jurisdictions and examined all of the warrants over a given period. *Id.* at 1049–50. Their results also show very high success rates, although they provided data only about the number of warrants leading to arrests—data that are not directly comparable. *Id.* at 1052.

39. Laurence A. Benner & Charles T. Samarkos, *Searching for Narcotics in San Diego: Preliminary Findings from the San Diego Search Warrant Project*, 36 CAL. W. L. REV. 221 (2000).

40. *Id.* at 250 tbl.20.

41. *Id.* at 249–50.

42. *Id.*

43. See *United States v. Johns*, 469 U.S. 478, 485 (1985) (“A vehicle lawfully in police custody may be searched on the basis of probable cause to believe that it contains contraband”); *Michigan v. Thomas*, 458 U.S. 259, 260 (1982) (per curiam) (holding that police officers may conduct a warrantless search of an impounded vehicle when they have probable cause to believe there is contraband in the vehicle); *Texas v. White*, 423 U.S. 67, 68 (1975) (per curiam) (finding that where there is probable cause for a search, officers may conduct the search at the stationhouse, as well as at the scene); *Chambers v. Maroney*, 399 U.S. 42, 48–49 (1970) (holding that the post-seizure warrantless search of a vehicle was justified by probable cause).

degree of law enforcement success when officers believe that they have probable cause to search but do not obtain a warrant in advance.

Over the past several years, the Department of Justice has surveyed the public about their contacts with law enforcement, as part of its ongoing series, *Contacts Between Police and the Public*.⁴⁴ While a primary purpose of these surveys is to determine whether members of racial minorities are treated differently, the surveys also provide data on when law enforcement performs searches and what the results are. The surveys produce consistent results—searches recover evidence between 11.6% and 12.6% of the time.⁴⁵ Compared to the recovery rates for search warrants discussed above, these rates are extremely low.⁴⁶

Of course, one response to this data might be skepticism about reliability. Individuals asked about their contacts with law enforcement are likely to underreport the frequency with which the contact led to incriminating evidence. Alternatively, these results may conflate multiple types of searches. While the *Contacts* surveys distinguish between consensual and nonconsensual searches (and find similar outcomes),⁴⁷ they do not distinguish between nonconsensual searches where law enforcement officers rely solely on probable cause and those cases where, for instance, the search occurred as part of the inventory of a car incident to the arrest of the driver.⁴⁸

44. MATTHEW R. DUROSE ET AL., U.S. DEP'T OF JUSTICE, CONTACTS BETWEEN POLICE AND THE PUBLIC, 2005 (2007) [hereinafter 2005 NATIONAL SURVEY]; 2002 NATIONAL SURVEY, *supra* note 4; PATRICK A. LANGAN ET AL., U.S. DEP'T OF JUSTICE, CONTACTS BETWEEN POLICE AND THE PUBLIC, FINDINGS FROM THE 1999 NATIONAL SURVEY (2001) [hereinafter 1999 NATIONAL SURVEY].

45. See 2005 NATIONAL SURVEY, *supra* note 44, at 7 (showing an 11.6% recovery rate); 2002 NATIONAL SURVEY, *supra* note 4, at 13 (showing an 11.7% recovery rate); 1999 NATIONAL SURVEY, *supra* note 44, at 20 (showing a 12.6% recovery rate).

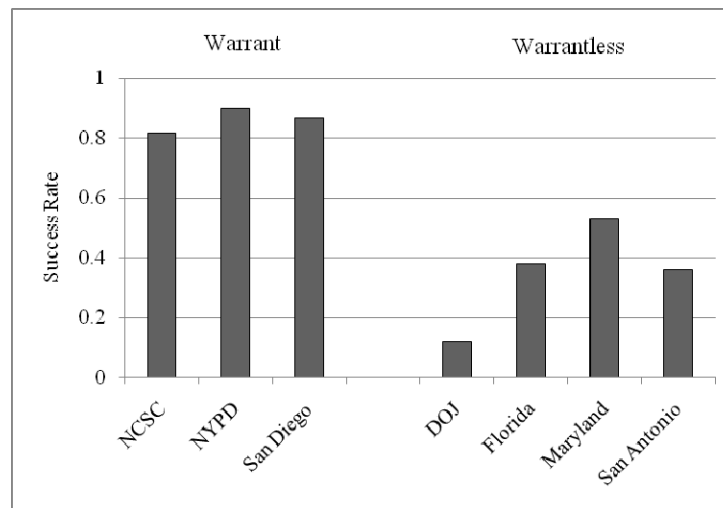
46. If the survey data are right, these results suggest that the success rates of searches based on probable cause are comparable to those based solely on reasonable suspicion. In 2006, a group of statisticians reviewed the New York City Police Department's stop-and-frisk policy for evidence of racial bias. Andrew Gelman et al., *An Analysis of the NYPD's "Stop-and-Frisk" Policy in the Context of Claims of Racial Bias*, 102 J. AM. STAT. ASS'N 813 (2007). The Gelman study found that when NYPD officers engaged in stops based on what the officer reported as "reasonable suspicion," an arrest resulted about one time in seven, or 14% of the time. *Id.* at 816.

47. See 2005 NATIONAL SURVEY, *supra* note 44, at 7 ("Consent and nonconsent searches turned up evidence of criminal wrong-doing at similar rates."); 2002 NATIONAL SURVEY, *supra* note 4, at 13 ("The likelihood of finding criminal evidence was not significantly different between the 379,500 searches without consent (12.2%) and the 458,300 searches with consent (11.4%) . . ."); 1999 NATIONAL SURVEY, *supra* note 44, at 21 ("The likelihood of finding criminal evidence was not significantly different between the 845,000 searches without consent (12.9%) and the 427,000 searches with consent (14.2%) . . .").

48. See 2005 NATIONAL SURVEY, *supra* note 44, at 7 (acknowledging that nonconsensual searches could arise in three situations—when permission had not been requested, when permission had been requested but refused, and when the vehicle search was conducted pursuant to the arrest of the driver—but failing to use this information in the tabulation of statistics); 2002 NATIONAL SURVEY, *supra* note 4, at 12–13 (enumerating how many nonconsensual searches of vehicles were conducted after the arrest of the driver and how many of the searches were probable-cause searches, but failing to incorporate this categorization when tabulating statistics); 1999 NATIONAL SURVEY, *supra* note 44, at 21 (distinguishing only between nonconsensual searches in which no permission

Fortunately, some jurisdictions have started requiring officers to record every car stop they perform.⁴⁹ Among the data collected are records of every time an officer searches a vehicle or driver and the reason the search was conducted, including whether it was based on probable cause or was justified by some other reason.⁵⁰ These data show hit rates for probable-cause searches of automobiles that are higher than those reported in the *Contacts* survey, but still far below the hit rates for searches on warrants. For instance, San Antonio hired a consulting company to review its car-stop data in 2003 and found an overall hit rate of 35.1%,⁵¹ while Maryland claimed a 52.5% hit rate from its probable-cause searches.⁵² The Florida State Highway Patrol data, discussed in the Introduction, reflect a 38.2% success rate.⁵³ Figure 1 compares these success rates to those of warranted searches, discussed above.

Figure 1: Searches With and Without Warrants



had been asked—96% of nonconsensual searches—and those in which permission had been requested but was not granted—4% of nonconsensual searches).

49. *E.g.*, MO. REV. STAT. § 590.650.2(1)–(10) (2001).

50. *E.g.*, Samuel R. Gross & Katherine Y. Barnes, *Road Work: Racial Profiling and Drug Interdiction on the Highway*, 101 MICH. L. REV. 651, 674 tbl.9 (2002) (comparing the percentage of consent searches in which drugs were discovered to the percentage of probable-cause searches in which drugs were discovered).

51. *See* JOHN C. LAMBERTH, SAN ANTONIO RACIAL PROFILING DATA ANALYSIS STUDY, FINAL REPORT FOR THE SAN ANTONIO POLICE DEPARTMENT 48 tbl.8 (2003), available at http://www.sanantonio.gov/SAPD/pdf/LambersSanAntonioRpt_2003.pdf (listing the overall hit rates for white drivers, black drivers, and Hispanic drivers).

52. *See* Gross & Barnes, *supra* note 50, at 700 tbl.14 (reporting that 47.5% of probable-cause searches yielded no drugs).

53. Florida State Highway Patrol Data, *supra* note 5.

B. The Costs of Searches

This divergence in hit rates does not immediately mean that adding them to the magistrate judge's calculation would improve search accuracy. If the facts leading to the divergent success rates are already present in the probable-cause analysis, success rates are superfluous. For instance, if the difference between search warrants and warrantless searches is simply the difference between searching a home and searching a car, we might presume that courts are already considering this fact. The core question is the cause—why do these hit rates diverge so substantially?

We do not know for sure, but the most likely answer is *cost*. Searches pursuant to a warrant are more expensive for law enforcement than those without warrants.⁵⁴ Along with a coauthor, I recently analyzed the success rates of a specialized type of search warrant—federal wiretaps.⁵⁵ Our results strongly suggest that cost greatly influences the likelihood of a wiretap's succeeding.⁵⁶

Wiretaps are a form of search subject to the constraints of the Fourth Amendment⁵⁷ and the more specific limitations of the Wiretap Act, which requires the investigating agent and the line prosecutors to seek approval from high-level officials at the Department of Justice and from a federal judge.⁵⁸ In addition, the Wiretap Act requires that the details of the wiretap as well as the outcome of the interception be reported to Congress and to the general public.⁵⁹ We performed a detailed statistical analysis on the data for all federal wiretaps from 1998 until 2004.⁶⁰

First and foremost, the results of the regression analysis suggest that the warrant requirement is not the effective binding constraint on the use of

54. As far as I am aware, Donald Dripps is the first to suggest that law enforcement is primarily constrained by cost, rather than the warrant requirement, when making search decisions. Donald Dripps, *Living with Leon*, 95 YALE L.J. 906, 926 (1986). His suggestion has received some attention in the legal literature, but it deserves more. See, e.g., William J. Stuntz, *O.J. Simpson, Bill Clinton, and the Transsubstantive Fourth Amendment*, 114 HARV. L. REV. 842, 848 (2001) (arguing that the high cost of obtaining a warrant effectively raises the substantive criminal standard applied to a search).

55. Max Minzner & Christopher M. Anderson, *Do Warrants Matter?* (Cardozo Law Sch. Legal Studies Research Paper Series, Paper No. 212, 2007), available at <http://ssrn.com/abstract=1073142>.

56. *Id.* at 2.

57. See *Katz v. United States*, 389 U.S. 347, 353 (1967) (“The government’s activities in electronically listening to and recording the petitioner’s words violated the privacy upon which he justifiably relied . . . and thus constituted a ‘search and seizure’ within the meaning of the Fourth Amendment.”); *Berger v. New York*, 388 U.S. 41, 51 (1967) (reiterating that the use of electronic devices to capture conversations is considered a search within the meaning of the Fourth Amendment).

58. 18 U.S.C. § 2516(1) (2006) (listing the officials eligible to authorize an application for electronic surveillance).

59. See *id.* § 2519 (imposing the requirements that judges and law enforcement report wiretap results to the Administrative Office of the Courts (AOC) and that the AOC report the results to Congress).

60. Minzner & Anderson, *supra* note 55, at 30 tbl.1.

wiretaps.⁶¹ If the probable-cause constraint works as the binding constraint, marginal wiretaps, i.e., the wiretaps least likely to succeed, should recover evidence at about the level predicted by the probable-cause constraint. However, the predicted success rate for virtually every wiretap in the data set exceeded even a conservative definition of probable cause.⁶²

The second and related conclusion of the work is an analysis of the factors that do limit law enforcement. The wiretap data set allowed us to analyze the factors that increase or decrease the likelihood of the wiretap's succeeding. We concluded that the effective constraint on the use of wiretaps is not external, but rather occurs within law enforcement.⁶³ Because it operates in a universe of limited resources, law enforcement carefully selects which searches to perform based on the costs and benefits of each search.⁶⁴

Three factors supporting this conclusion stood out in the research. First, the cost of the wiretap has a strong positive relationship to its likelihood of success.⁶⁵ More expensive wiretaps are more likely to succeed. Because law enforcement budgets are limited, they only engage in expensive searches when those searches are especially likely to succeed. Second, the number of searches performed by a law enforcement agency has a strong negative relationship to the success rate of its searches.⁶⁶ Law enforcement agencies (in this case, U.S. Attorney's offices) that do more searches do them less well. This result reinforces the importance of the budget constraint. When law enforcement officers have more resources available, they dip further into the

61. *See id.* at 15–16 (finding that over 80% of wiretaps have a better chance of ending in arrest compared to even the most restrictive standard of probable cause).

62. *Id.* The Supreme Court has refused to assign a number to the probability required for probable cause. *See, e.g., Maryland v. Pringle*, 540 U.S. 366, 371 (2003) (“The probable-cause standard is incapable of precise definition or quantification into percentages because it deals with probabilities and depends on the totality of the circumstances.”); *Illinois v. Gates*, 462 U.S. 213, 235 (1983) (holding that probable cause cannot be defined with a “general, numerically precise degree of certainty”). We adopted a 50% threshold as a conservative approach—most of the Court's statements suggest that probable cause lies somewhere south of 50%. In *Gerstein v. Pugh*, 420 U.S. 103 (1975), the Supreme Court noted that a probable-cause evaluation “does not require the fine resolution of conflicting evidence that a reasonable-doubt or even a preponderance standard demands.” *Id.* at 121. A preponderance standard requires simply that a fact be more likely than not. The seminal *Illinois v. Gates* opinion also suggested that the standard is actually substantially lower than 50%, indicating that the government must simply show “a fair probability that contraband or evidence of a crime will be found in a particular place.” 462 U.S. at 238. Other judges express similar views. A 1982 survey of 166 federal judges asked their view of the numeric equivalent of probable cause and found a mean level of 45.8% and a median of 50%. C.M.A. McCauliff, *Burdens of Proof: Degrees of Belief, Quanta of Evidence, or Constitutional Guarantees?*, 35 VAND. L. REV. 1293, 1327 (1982). Of course, a strict linguistic analysis of probable cause would place it at exactly 50%—more probable than not. *See Amar, supra* note 14, at 784 (“A warrant should issue if it is ‘probable’—more likely than not.”).

63. Minzner & Anderson, *supra* note 55, at 22.

64. *Id.* at 22–23.

65. *Id.* at 17.

66. *Id.* at 18.

pool of potential targets and begin to engage in more low-probability searches. Third, law enforcement is responsive to the target of the search.⁶⁷ When the targeted crime is especially serious—such as terrorism—law enforcement is more aggressive and seeks more marginal wiretaps.⁶⁸ The Benner and Samarkos study includes some results that support this conclusion as well. They found that the success rates of the narcotics wiretaps they studied varied based on the targeted drug.⁶⁹ Searches targeting marijuana were very successful while searches for methamphetamine and rock cocaine were far more likely to fail.⁷⁰

These results strongly suggest that law enforcement is aggressive when searches are cheap or the expected results valuable, but conservative when searches are expensive or the results relatively unimportant. This information is completely absent from the magistrate's current probable-cause decision making, but the regression results discussed here suggest that it may be one of the most important factors predicting the actual likelihood of recovering evidence. Adding in law enforcement's history of success will pick up this effect—cheap searches will get the additional scrutiny they deserve and expensive ones will be approved more easily.

One alternative (and perhaps better) approach would be to require law enforcement officers to report the costs of the search they intend to engage in as well as require them to provide the historic success rate of a search of that cost. This proposal is less plausible since most searches do not involve clearly defined budgetary costs; nor would the magistrate know the relative budget constraints for each law enforcement agency. Further, police are generally on salary and the costs of searches are mostly opportunity costs—the lost chances to make other arrests or pursue other investigations. We are limited to observing data that is relatively easy to collect.

I have spent most of this subpart discussing the second requirement for success rates to be useful, i.e., that the reasons for the differences in success rates be hidden from magistrate judges. What about the requirement that past success rates must be useful in predicting future success? The arguments above provide some intuitive reason to believe that the requirement is met. If budget constraints do shape success rates, those constraints are likely to persist from year to year within any law enforcement unit. We can also test this requirement against the data set of federal wiretaps. The data allow a year-by-year analysis of search success rates at the law enforcement level, which can compare the predictive power in a given federal district of prior years' wiretap success rates for future years.

Using arrests as the measure of success, Figure 2 is a scatterplot of the success rates of federal wiretaps from 1997 to 2004, with the vertical axis

67. *Id.* at 19–20.

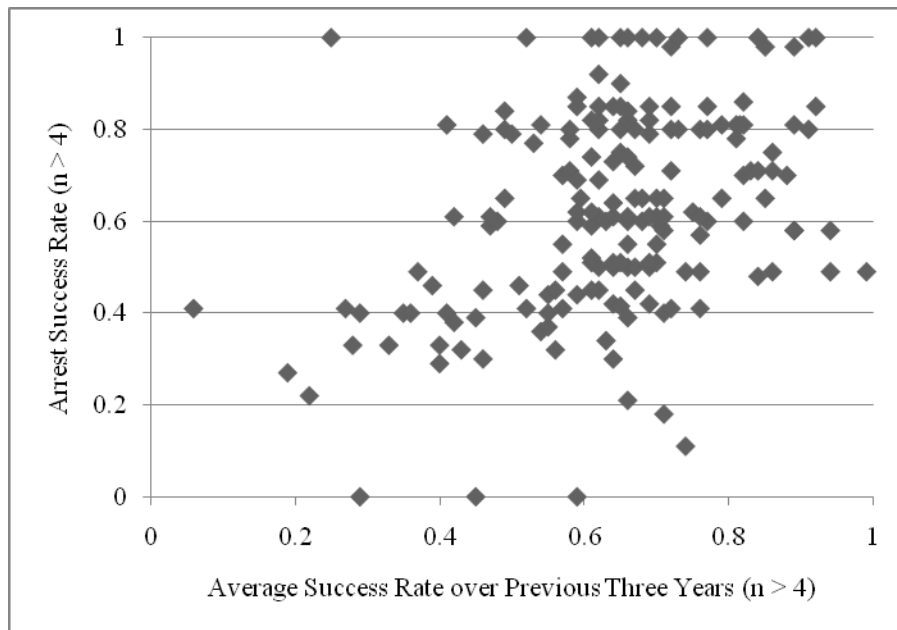
68. *Id.* at 20.

69. Benner & Samarkos, *supra* note 39, at 252.

70. *Id.* at 250.

reflecting the average success rate of the current year's wiretaps and the horizontal axis reflecting the average success rates of the previous three years.⁷¹ If there were no relationship between the previous years' success rates and those of the current year, the observations would be scattered evenly over the Figure; if the success rate of the previous three years perfectly predicted the current year's success rate, all the observations would lie exactly on the 45-degree line running from the bottom left to the upper right. The correlation coefficient between the success rates of the previous three years and that of the current year is .40.⁷² While the relationship is not perfect, it is stronger than correlations in other situations where we routinely use data to predict outcomes. For instance, it is stronger than the median correlation between LSAT scores and first-year law-school grades.⁷³

Figure 2: Wiretap Success Rates by District



71. This analysis limits the observations to situations where there are at least five observations in both the current year and the previous three years.

72. The correlation coefficient ranges from 0 to 1 and measures the linear relationship between two data sets. See THAD MIRER, *ECONOMIC STATISTICS AND ECONOMETRICS* 49 (2d ed. 1988).

73. See LAW SCH. ADMISSIONS COUNCIL, 2008–2009 LSAT AND LSDAS INFORMATION BOOK 27 (2008), available at <http://lsat.org/pdfs/2008-2009/InformationBook08web.pdf> (“During 2006, validity studies were conducted for 187 law schools. Correlations between LSAT scores and first-year law school grades ranged from .04 to .60 (median is .33).”).

IV. Hidden Facts: Types of Officers

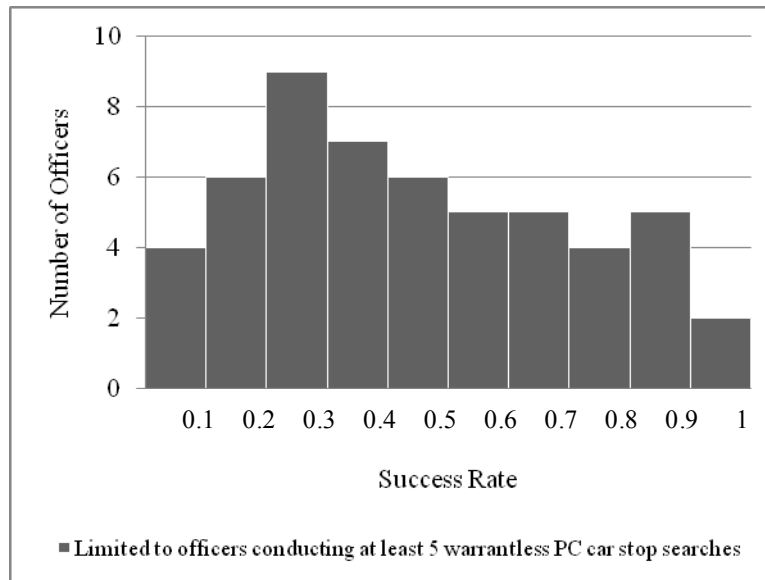
The preceding Part argued that the current search process misses a substantial amount of information about the likelihood that a given search will recover evidence because courts are not looking at success rates for the type of search for the relevant law enforcement unit. This Part argues in favor of considering another statistic—officer-level success rates. Here we begin with the first critical question—do success rates of officers vary and are they consistent over time? Is an officer's history of success in the past predictive of his probability of success in the future? I consider a data set of searches by the Florida State Highway Patrol.⁷⁴ From January 2000 until November 2001, the Florida State Highway Patrol recorded every stop made by an officer in the state and provided detailed statistical information about the stop, including the date, time, and county; demographic information about the driver; the reason for the search; and the outcome.⁷⁵ Crucially, the data also include a numeric identifier for each officer, so we can consider whether officer success is consistent over time.⁷⁶ Figure 3 is a histogram of the success rates of the officers who conducted at least five probable-cause car-stop searches during the period covered by the data set. The Figure demonstrates substantial variation across officers. Some officers (those on the right of the histogram) are fantastic when they conduct probable-cause searches—they recover evidence most of the time. Others (those on the left) are far less successful—they usually fail when they conduct a probable-cause search.

74. Florida State Highway Patrol Data, *supra* note 5. The charts on the following pages of this Article illustrate the data collected in the course of this Florida State Highway Patrol study.

75. *Id.* This same data set has been used in the racial-profiling literature. See Shamera Anwar & Hamming Fang, *An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence* 4 (Nat'l Bureau of Econ. Research, Working Paper No. 11264, 2004), available at <http://ssrn.com/abstract=705584>.

76. Florida State Highway Patrol Data, *supra* note 5.

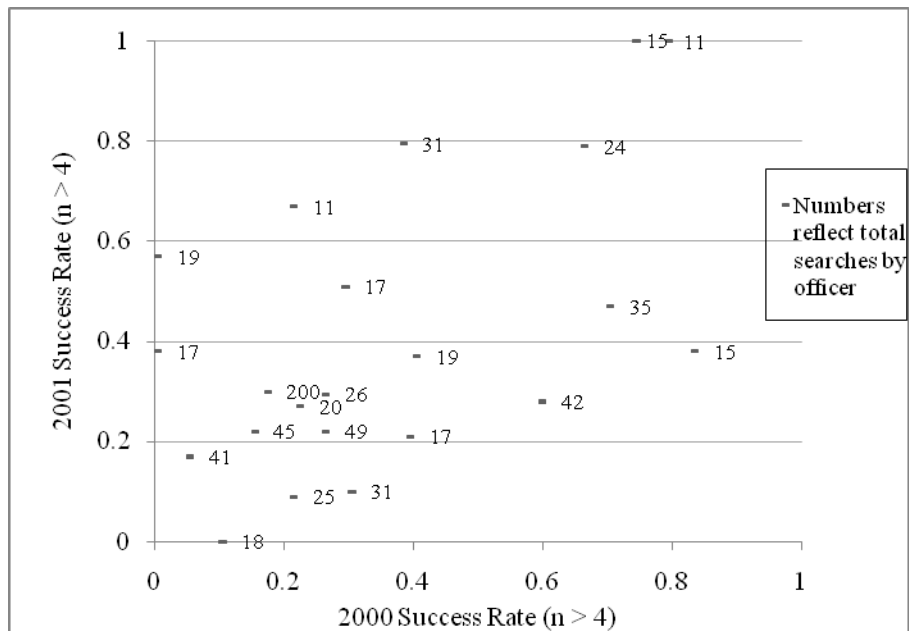
Figure 3: Success Rates of Florida State Police Officers



Moreover, we have substantial reason to believe that these differences are not simply random variation, as they tend to be consistent over time. Figure 4 is a similar scatterplot to that presented in Figure 2 for wiretaps, but it compares the success rates for individual officers in 2001 to their success rates in 2000. If the previous year's success rate perfectly predicted the current year's success rate, all of the observations would lie on the 45-degree line. Here the correlation coefficient is .53, even higher than that of the wiretap data.⁷⁷ The same officers who succeeded at high rates in 2000 also succeeded at high rates in 2001. Similarly, officers who were less successful in 2000 tended to be relatively less successful in 2001.

77. This analysis also limits the observations to situations where the officer conducted at least five searches in both 2000 and 2001.

Figure 4: Florida State Police Probable-Cause Searches



Of course, as discussed above, this correlation only matters if the reasons for differential officer success rates are currently excluded from the probable-cause analysis. The history of prior success for the searching officer will be relevant when there is important information about that officer that is not apparent from the facts in the affidavit or the suppression-hearing testimony. One easy example of important, officer-specific information is officer competence. If some officers are simply more capable at identifying incriminating information than others, they will have different probabilities of success despite appearing equally competent to a magistrate.⁷⁸

More troubling, success rates might diverge as a result of perjury. Law enforcement perjury in criminal cases is widely viewed as a serious problem and has been the subject of substantial scholarly attention.⁷⁹ For obvious

78. Along related lines, see Craig S. Lerner, *Reasonable Suspicion and Mere Hunches*, 59 VAND. L. REV. 407, 466 (2006). Professor Lerner argues persuasively in favor of taking police hunches more seriously because some officers have better hunches than others. *Id.*

79. See generally, e.g., David N. Dorfman, *Proving the Lie: Litigating Police Credibility*, 26 AM. J. CRIM. L. 455 (1998) (questioning the credibility of police testimony and examining the judicial mechanisms for minimizing police perjury); Jennifer E. Koepke, Note, *The Failure to Breach the Blue Wall of Silence: The Circling of the Wagons to Protect Police Perjury*, 39 WASHBURN L.J. 211 (2000) (discussing how a code of silence among police officers protects police offenders from being held accountable for misconduct); Andrew J. McClurg, *Good Cop, Bad Cop: Using Cognitive Dissonance Theory to Reduce Police Lying*, 32 U.C. DAVIS L. REV. 389 (1999) (exploring the reasons why police officers lie and arguing that innovative training and mentoring programs help address the problem); Christopher Slobogin, *Testifying: Police Perjury and What To*

reasons, limited reliable data exist on the scope of police perjury. Alan Dershowitz has claimed, "Almost all police lie about whether they violated the Constitution in order to convict guilty defendants."⁸⁰ This view is shared by many participants in the criminal justice process. Myron Orfield surveyed participants in the Chicago criminal justice system regarding law enforcement perjury at suppression hearings.⁸¹ Orfield's survey data suggested that judges, prosecutors, and defense attorneys all view police perjury at suppression hearings as at least relatively common.⁸² Only 8% of the respondents believed it never happened.⁸³ Notably, only 25% of prosecutors (and none of the defense attorneys surveyed) believed police never lie in court.⁸⁴ Of course, the views summarized by this survey data are based on conjecture. Measuring law enforcement perjury is an extremely difficult task and no participant in the criminal justice process currently has good data to estimate it with.⁸⁵

Despite the dearth of data, many commentators have claimed that the most common form of perjury is the post hoc fabrication of evidence during suppression hearings.⁸⁶ In the current system, the primary opportunity to uncover police perjury in the search process comes at the suppression hearing. When an officer engages in a warrantless search that leads to the seizure of evidence, the Government bears the burden of establishing an exception to the warrant requirement and the existence of probable cause.⁸⁷ Similarly, once a defendant makes an initial showing that a warrant affidavit contains a potentially false statement of deliberate falsity or reckless disregard, that defendant is entitled to a hearing to determine whether the affidavit, after excluding any false information, contained sufficient evidence to support a probable-cause finding.⁸⁸ The standard tools for determining

Do About It, 67 U. COLO. L. REV. 1037 (1996) (describing the nature and causes of police perjury and proposing ways to curtail it).

80. ALAN M. DERSHOWITZ, *THE BEST DEFENSE*, at xxi (1982).

81. Myron W. Orfield, Jr., *Deterrence, Perjury, and the Heater Factor: An Exclusionary Rule in the Chicago Criminal Courts*, 63 U. COLO. L. REV. 75 (1992).

82. *Id.* at 107.

83. *Id.*

84. *Id.* at 107 n.148.

85. See Kevin R. Reitz, *Testilying as a Problem of Crime Control: A Reply to Professor Slobogin*, 67 U. COLO. L. REV. 1061, 1062 (1996) ("We know almost nothing about the testilying 'rate,' its variations across and within police departments, its changes over time, or its etiology. . . . Compared with many other offenses, the crime of testilying has been poorly measured, and we should be suspicious of claims that its incidence is known or its causes understood.").

86. McClurg, *supra* note 79, at 398 & n.48 (collecting sources and concluding that "[c]ommentators generally agree that police falsification is most common in connection with search and seizure activity, particularly the post hoc manufacturing of probable cause").

87. *California v. Acevedo*, 500 U.S. 565, 579 (1991); see also *Mincey v. Arizona*, 437 U.S. 385, 390-91 (1978) (placing the burden upon the State to show the existence of an exceptional situation, justifying an exception to the warrant requirement).

88. *Franks v. Delaware*, 438 U.S. 154, 171-72 (1978). For a recent summary of the current case law surrounding *Franks* and the Fourth Amendment's Warrant Clause, see Steven W. Gard,

credibility are available at such a hearing: demeanor, cross-examination, and the oath.⁸⁹ While there is always reason to be skeptical of the ability of an audience to judge the veracity of a live witness,⁹⁰ there is particular reason to be skeptical in the search context. Recent research has suggested that lie detection is likely to be highly accurate when observers possess private information about the topic that is unavailable to the witness.⁹¹ In the search context, this will seldom be true. Probable-cause decisions are often based on testimony that is only observable to the arresting officer, including, for example, whether the defendant's conduct was "evasive"⁹² or "surveillance-conscious."⁹³ While courts do sometimes reject law enforcement testimony as false,⁹⁴ there is not much reason to believe that the decision to do so is accurate.⁹⁵

Academic commentators have proposed a wide variety of solutions to the problem of police perjury in suppression hearings, none of which is entirely satisfactory. Professor Dripps has proposed that all officers testifying at suppression hearings undergo polygraph tests.⁹⁶ Professors Chin and Wells suggest that fact finders be exposed to expert testimony on the "blue wall of silence" that discourages law enforcement officers from testifying against each other,⁹⁷ and Professor McClurg endorses an aggressive training

Bearing False Witness: Perjured Affidavits and the Fourth Amendment 44–45 (Cleveland-Marshall Coll. of Law, Paper No. 06-132, 2006), available at <http://ssrn.com/abstract=946476>.

89. See *Maryland v. Craig*, 497 U.S. 836, 845–46 (1990) (citing these three mechanisms as the primary processes for ensuring that testimony is reliable).

90. For a general discussion of this subject, see Max Minzner, *Detecting Lies Using Demeanor, Bias, and Context*, 29 CARDOZO L. REV. 2557 (2008).

91. *Id.* at 2571. For the underlying studies on this point, see Maria Hartwig et al., *Strategic Use of Evidence During Police Interrogations: When Training to Detect Deception Works*, 30 LAW & HUM. BEHAV. 603 (2006) (testing the effects of early and late disclosure of evidence on lie-detection accuracy, and finding that late disclosure of evidence leads to more accurate lie detection); Maria Hartwig et al., *Detecting Deception via Strategic Disclosure of Evidence*, 29 LAW & HUM. BEHAV. 469, 477 (2005) (determining that study participants were better able to detect deception in situations where evidence against a suspect was presented later in the participants' exposure to the scenario). This result is consistent with our intuition. When trying to evaluate a speaker's credibility, we are better off if we know more about the topic than the other party.

92. *E.g.*, *United States v. Humphries*, 372 F.3d 653, 657, 660 (4th Cir. 2004).

93. *E.g.*, *United States v. Gutierrez*, 891 F. Supp. 97, 104 (E.D.N.Y. 1995).

94. See Orfield, *supra* note 81, at 107 (finding that public defenders, state's attorneys, and judges estimate that judges disbelieve law enforcement at suppression hearings approximately 20% of the time).

95. This analysis is in addition to the common claim that judges are simply insufficiently aggressive in scrutinizing law-enforcement testimony, due to their bias in favor of believing officers over defendants, distaste for the exclusionary rule, or other reasons. See Dorfman, *supra* note 79, at 472 (citing Professor Morgan Cloud's five reasons why judges do not sufficiently scrutinize police witnesses during suppression hearings); Morgan Cloud, *The Dirty Little Secret*, 43 EMORY L.J. 1311, 1321, 1321–24 (1994) (elaborating five reasons why judges will "accept perjured testimony from police officers about the manner in which searches and seizures are conducted").

96. Donald A. Dripps, *Police, Plus Perjury, Equals Polygraphy*, 86 J. CRIM. L. & CRIMINOLOGY 693, 693 (1996).

97. Gabriel J. Chin & Scott C. Wells, *The "Blue Wall of Silence" as Evidence of Bias and a Motive to Lie: A New Approach to Police Perjury*, 59 U. PITT. L. REV. 233, 237 (1998).

program for new law enforcement officers modeled on Alcoholics Anonymous to inculcate antiperjury values.⁹⁸ Finally, as is true in many areas of Fourth Amendment jurisprudence, abandoning the exclusionary rule⁹⁹ and expanding tort remedies¹⁰⁰ have been suggested as solutions to the problem.

None of these solutions, though, takes advantage of the unique nature of perjury in the probable-cause context. It is reasonable to assume that police perjury works in only one direction—when law enforcement officers lie, they are attempting to increase the perceived likelihood of recovering evidence. They are attempting to convince the court that the probability of finding contraband is higher than it actually is. If this is true, hit rates provide an effective and efficient mechanism for uncovering perjury. While false testimony can increase the apparent likelihood of recovering evidence, it cannot increase the actual likelihood of finding evidence.¹⁰¹ By comparing the apparent likelihood of recovering evidence to the actual recovery rate, we can distinguish good cops from bad cops.

This concept is developed in the Appendix in more formal terms, but the intuition is straightforward. Imagine two officers, one honest and one dishonest. They both present to a magistrate warrant applications containing the same apparent likelihood of recovering evidence. However, while the apparent likelihood of finding evidence reflects the true probability of recovering evidence for the honest officer, the application by the dishonest officer inflates the actual likelihood of finding contraband. Under these circumstances, the honest officer is more likely to find evidence. In the current process, the perjury is very difficult to detect, because we only analyze the case if a recovery takes place. Once that happens, the officers look identical. If we focus on the overall hit rate, however, and pay attention to the cases where no evidence is found, the hit rates will diverge and we will be able to distinguish honest officers from dishonest officers.

Of course, the hit rates for these officers will not diverge immediately. Consider two officers, both of whom present facts that suggest an 80% likelihood of recovering evidence. However, one of these two officers is lying—the actual facts give only a 40% chance of finding contraband. When they initially join the force, neither has conducted any searches. Under these circumstances, the court has no prior data to evaluate and can only look at the history of the police force in general and the quality of the evidence in particular. As these officers begin to conduct searches, though, they will

98. McClurg, *supra* note 79, at 430–31.

99. Slobogin, *supra* note 79, at 1057–58.

100. Koepke, *supra* note 79, at 237.

101. In addition to increasing the apparent likelihood of recovering evidence (by misrepresenting probable cause), officers can increase the likelihood of apparently recovering evidence (by planting evidence). For more discussion on this subject, see *infra* subpart V(A).

succeed at differing rates, and their success rates will begin to diverge. Table 1 below reflects the 95% confidence intervals for the recoveries for honest and dishonest law enforcement officers for a given number of warrants.¹⁰²

Table 1: Confidence Intervals for Recoveries

Number of Searches	95% Confidence Interval for Recoveries for Honest Law Enforcement	95% Confidence Interval for Recoveries for Dishonest Law Enforcement
10	(6,10)	(1,7)
15	(9,15)	(2,10)
20	(12,20)	(4,12)
25	(16,24)	(5,15)

After as few as ten or fifteen searches, the court will observe a substantial difference in outcomes between honest and dishonest law enforcement. After twenty-five searches, the 95% confidence intervals for the search outcomes do not overlap at all.

Examining hit rates is certainly not a panacea for the problem of police perjury. The example above rests on a number of assumptions. First, it assumes that officers face a constant level of probable cause in their warrant applications, or that the potential levels of probable cause fall within a relatively narrow range. If there is substantial variation in the strength of the evidence on which the search decision is based, there will also be variation in the recovery rates of evidence. This noise will delay the rate at which the rates of recovery diverge but will not prevent it entirely. In the long run, the good cops will succeed at higher rates than the bad.

Second, this mechanism works well if police perjury is a phenomenon that involves some officers who lie frequently. I suggest that we compare the hit rates of each individual officer to the average success rate of his colleagues. Of course, if all officers lie a little bit sometimes, the success rates will be similar. By contrast, the method will easily identify officers who lie frequently and whose lies are substantial.

Third, the divergence in success rates will be most apparent for officers who engage in a large number of searches. We have little data on how many searches individual officers perform, but the limited data that do exist sug-

102. The analysis reflects the outcome of a given number of Bernoulli trials with parameters .4 or .8, respectively. The total number of recoveries after n trials is therefore normally distributed with parameters $N(.4n, \sqrt{.32n})$ and $N(.8n, \sqrt{.16n})$, respectively.

gest that search rates vary substantially by officer.¹⁰³ Of the 192 officers in the Florida State Police data set who completed a probable-cause search between January 2000 and September 2001, the average number of searches per officer was 6.14, but the numbers vary substantially around this mean.¹⁰⁴ Most officers (60.4%) conducted only one or two searches.¹⁰⁵ However, officers who perform a large number of searches conduct the substantial majority of searches.¹⁰⁶ Twenty-five officers conducted at least ten searches during this time period, and these officers were responsible for 767 (65%) of the 1,180 searches.¹⁰⁷ More research is needed on the number of searches actually conducted by individual officers in other jurisdictions, but these data suggest that magistrates are likely to be faced with probable-cause claims from officers with substantial track records.

V. Incorporating Success Rates: Law Enforcement Responses

How will law enforcement respond to data-collection and reporting requirements? The long-term effects of this approach may be somewhat difficult to predict, but we can make some guesses. First, and most seriously, some officers may be tempted to cheat by generating artificially inflated success rates. Subpart V(A) discusses some of the ways in which law enforcement might cheat and analyzes the possibility of detecting misbehavior. Second, law enforcement may respond with a shift in focus by both individual officers and the law enforcement entity as a whole. In particular, individual officers who succeed at high rates may get used more frequently, especially in high-priority cases. Subpart V(B) examines these effects. Finally, subpart V(C) argues that law enforcement will be forced to generate new, useful data about searches. Right now, courts make assumptions about the value of information underlying search claims, but we know little about whether these assumptions are right. Success rates will help generate that data. Additionally, this data will help identify and eliminate racial profiling.

A. Law Enforcement Responses: Data Manipulation

The most serious concern about a data-reporting requirement would be that some officers or agencies would cheat on their data collection. Of course, accurate data is essential to the success of this proposal. If data re-

103. See Kate L. Antonovics & Brian G. Knight, *A New Look at Racial Profiling: Evidence from the Boston Police Department* 9 (Nat'l Bureau of Econ. Research, Working Paper No. 10634, 2004), available at <http://ssrn.com/abstract=565845> (noting that search rates in their Boston data set varied substantially across officers).

104. Florida State Highway Patrol Data, *supra* note 5.

105. *Id.*

106. *Id.*

107. *Id.*

ported to magistrates are not accurate, using the data is a mistake. There are two ways that law enforcement could inflate success rates, both of which involve eliminating failures. First, law enforcement officers could simply not report their failures. Second, law enforcement could convert failures into apparent successes by planting evidence during searches. I discuss both of these options in turn and describe some of the mechanisms with which we could prevent such misbehavior.

Hiding failures, as opposed to converting them into successes, is almost exclusively a problem for warrantless searches. Because search warrants are sought in advance, we can easily know the total number of searches made pursuant to a warrant by each officer and by each law enforcement unit, and thus officers cannot conceal their failures. We can harness this power of precommitment in the warrantless-search context as well. If officers were required to announce the search in advance, the problem of concealing failures would disappear as well. The rise of information technology makes such an option feasible. As a prerequisite to a search, we could require a simple e-mail or phone call to a designated record keeper—either in law enforcement or, better yet, at the court—announcing the identity of the officer, the identity of the unit, and the location to be searched.¹⁰⁸

Even without an additional legal requirement of predisclosure, we currently have tools available to prevent law enforcement from concealing failures in car-stop searches. Law enforcement's self-policing will deter some misbehavior. To the extent that officers work in pairs, searches can only be concealed by a conspiracy between partners. Additionally, many law enforcement units have installed video cameras on police cars, primarily to protect against claims of police misconduct.¹⁰⁹ These cameras will usually detect a car-stop search if it is conducted. If car-stop data become common evidence in suppression hearings, defense attorneys will certainly begin to subpoena and analyze the videotapes of prior stops conducted by the officers in the case. Finally, we have data-related mechanisms to detect hidden failures. Officers who hide failures to inflate their success rates will appear to conduct fewer searches than they actually do. They will appear to be offi-

108. This approach does not fall prey to the objection usually raised in opposition to requiring a warrant prior to a car-stop search. The standard rationale behind the car-stop exception is that cars are inherently mobile; as a result, exigency is presumed. *See Pennsylvania v. Labron*, 518 U.S. 938, 940 (1996) (per curiam) (noting that the automobile exception to the Fourth Amendment's warrant requirement is based on "the automobile's 'ready mobility,' an exigency sufficient to excuse failure to obtain a search warrant once probable cause to conduct the search is clear"); *California v. Carney*, 471 U.S. 386, 390–91 (1985) (tracing the history of the exception). The time it takes to report that a search is going to be conducted is far less than the time needed to obtain a warrant. Officers would not need to find a magistrate and give the judge time to deliberate. They would simply need to inform someone that a search will happen.

109. The importance of the rise of law-enforcement videos can be seen in *Scott v. Harris*, 550 U.S. 372, 127 S. Ct. 1769 (2007), where a car-mounted video camera captured a high-speed chase and crash that injured a fleeing driver. *Id.* at 1773. When the driver sued, the video was determinative in the court's granting summary judgment. *Id.* at 1776.

cers with high success rates but comparatively fewer searches. As discussed in the next subpart, one likely law enforcement response to this proposal would be to pressure these officers to do more searches. In this way, law enforcement is also likely to detect this type of misconduct.

The problem of evidence planting is, of course, more serious. If a primary goal of analyzing success rates is to root out police perjury, a natural concern is that officers who lie would become officers who fabricate successes by framing the innocent. While we cannot be sure, there is reason to believe that at least some officers who are willing to lie in a suppression hearing would not be willing to frame an innocent person. "Police officers rarely lie to intentionally convict innocent persons. They lie to convict those whom they believe to be guilty."¹¹⁰ Because officers in a suppression hearing are invariably confronted with a guilty defendant, their obligation to testify truthfully is in conflict with their usual goal of fighting crime. By contrast, when officers frame the innocent, they are not serving any crime-reduction goal.

In the end, like any legal requirement, individual officers and law enforcement units that are committed to misbehavior can find ways to evade the rules. Evidence can be planted, videotape doctored, and e-mails altered. All of these proposed requirements, though, at least have the advantage that they would force misbehavior to occur at a more severe and detectable level than the current system. It takes far more dedication, effort, and conscious misbehavior to plant evidence or alter data records than it does to claim falsely that a search was appropriate because the defendant's conduct was "furtive."¹¹¹ For exactly those reasons, we have many more tools to detect those types of misbehavior than we do a false claim about the officer's subjective mental state.

B. Law Enforcement Responses: Shifting Priorities

Adding a success-rate requirement will begin to shift law enforcement incentives in positive ways. First and foremost, we should expect individual police officers to respond to failures and try to reduce the frequency of their searches. In the Fourth Amendment context, of course, a "failure" is a search of an innocent person. One of the primary problems, widely recognized in the academic literature, is that the exclusionary rule provides no benefit for

110. McClurg, *supra* note 79, at 412.

111. *See, e.g.*, United States v. Hernandez Leon, 379 F.3d 1024, 1028 (8th Cir. 2004) (relying on "defendants' furtive behavior" to establish probable cause); Cost v. Commonwealth, 657 S.E.2d 505, 510 (Va. 2008) ("[F]urtive gestures coupled with other indicia of criminal activity may suffice to establish probable cause."); *see also* Sibron v. New York, 392 U.S. 40, 66–67 (1968) ("[D]eliberately furtive actions and flight at the approach of strangers or law officers are strong indicia of *mens rea*, and when coupled with specific knowledge on the part of the officer relating the suspect to the evidence of crime, they are proper factors to be considered in the decision to make an arrest.").

searches of the innocent.¹¹² Aside from the rare lawsuit under Section 1983, law enforcement suffer no consequences from searching those who do not possess contraband. By contrast, a focus on success rates would force law enforcement to care a great deal about whether innocent people are searched and how to respond to those failures.

What shape will this response take? One probable outcome is a shift in officer assignment. Officers who have high search success rates will be encouraged to conduct more searches and officers who have low rates will be retrained or reassigned. This is certainly a positive outcome. As argued above, success rates can be a proxy—at least in some cases—for officer competence and honesty. We want searches to be conducted by the most competent and honest officers.

Second, law enforcement units will likely respond by shifting officers with high success rates onto cases that they value more highly. Law enforcement is unlikely to engage in low-probability searches if officers are relatively uninterested in the target of the search, but law enforcement is more likely to be more aggressive for high-priority searches. Officers would be unwilling to risk their success rates on nonviolent drug cases, but they would be more likely to roll the dice if the target of the search is a potential terrorist. In this way, a success-rate requirement could make the Fourth Amendment substantive rather than transsubstantive. Currently, the Fourth Amendment is blind to the type of crime underlying the search.¹¹³ Nonviolent drug cases are the same as those where violence is alleged; terrorism is treated the same as burglary. This approach has substantial costs, which have been outlined elsewhere,¹¹⁴ and an approach based on success rates will help move away from this transsubstantive model.¹¹⁵

112. See, e.g., William J. Stuntz, *Warrants and Fourth Amendment Remedies*, 77 VA. L. REV. 881, 911 (1991) (“Only the guilty are directly helped by the [exclusionary] rule, because the rule leads to suppression of evidence in criminal cases, and the only reason to seek suppression is if the evidence tends to establish the defendant’s guilt.”). This objection or observation has been repeated very frequently. E.g., Shenequa L. Grey, *Revisiting the Application of the Exclusionary Rule to the Good Faith Exceptions in Light of Hudson v. Michigan*, 42 U.S.F. L. REV. 621, 633 (2008); Alicia M. Hilton, *Alternatives to the Exclusionary Rule After Hudson v. Michigan: Preventing and Remediating Police Misconduct*, 53 VILL. L. REV. 47, 47 (2008); Carol S. Steiker, *Second Thoughts About First Principles*, 107 HARV. L. REV. 820, 848 (1994); see also Amar, *supra* note 14, at 799 (lamenting that the Fourth Amendment exclusionary rule has become associated with criminals getting off on minor technicalities).

113. See Stuntz, *supra* note 54, at 847 (“Like almost everything else in the law of criminal procedure, the Fourth Amendment treats one crime just like another.”).

114. Professor Stuntz persuasively outlines three significant costs: the standard is too high for relatively more serious cases, too low for relatively less serious cases, and inflexible in light of the changing nature of the criminal law. *Id.* at 853–55.

115. A plausible critique of the imposition of a success-rate requirement is that law enforcement might overcompensate and focus on easy cases where the evidence will obviously be recovered. From a Fourth Amendment standpoint, that outcome is perfectly acceptable. If law enforcement acts only on information that guarantees the recovery of evidence, the probable-cause standard is met. Of course, that outcome is undesirable from a policy perspective, but there is reason to believe that political control over the executive should eliminate that problem.

C. Law Enforcement Responses: Data Generation

A third effect of a success-rate requirement will be data generation. First, law enforcement will be forced to learn the correlates of criminal activity. The current system does not give any participant an incentive to learn how strongly various types of evidence actually predict the likelihood of recovering contraband. Judges currently make assumptions about the value of different evidence types, but we have no real information about what types of evidence actually predict whether contraband is likely to be found. A success-rate requirement would generate this type of information. As an example, below I will discuss this effect in terms of judicial assumptions about informant testimony and stale evidence.

Second, hit rates will help improve our understanding of racial profiling. Racial profiling has received a tremendous amount of attention in the legal literature.¹¹⁶ Under the standard approach to searches, racial profiling is virtually impossible to detect. In the search warrant context, officers are extremely unlikely to list the race of the target as a factor supporting probable cause. Similarly, officers attempting to justify a search after the fact generally are unlikely to try to support that search by pointing to the race of the target.¹¹⁷ By contrast, as discussed below, an outcome-based approach provides real promise in detecting racial profiling.

1. *Generating Data on Informants and Staleness.*—Courts often make strong assumptions about the strength or weakness of certain evidence types in the search context. For instance, the Supreme Court has drawn a clear line

116. See, e.g., Albert W. Alschuler, *Racial Profiling and the Constitution*, 2002 U. CHI. LEGAL F. 163 (examining the constitutionality of racial profiling under both the Fourth Amendment and the Equal Protection Clause); R. Richard Banks, *Racial Profiling and Antiterrorism Efforts*, 89 CORNELL L. REV. 1201 (2004) (arguing that because normative disagreements underlie racial profiling, an outright prohibition would simply relocate the problem from suspect descriptions to profile boundaries); Bernard E. Harcourt, *Rethinking Racial Profiling: A Critique of the Economics, Civil Liberties, and Constitutional Literature, and of Criminal Profiling More Generally*, 71 U. CHI. L. REV. 1275 (2004) (seeking to clarify the economic, civil liberties, and constitutional scholarship on the issue of racial profiling); Russell L. Jones, *A More Perfect Nation: Ending Racial Profiling*, 41 VAL. U. L. REV. 621 (2006) (detailing the dangers and increased use of racial profiling in America, and providing a judicial remedy for the problem); Andrew E. Taslitz, *Racial Profiling, Terrorism, and Time*, 109 PENN. ST. L. REV. 1181 (2005) (proposing that the 1860s should be privileged when interpreting the Fourth Amendment, and arguing that this will eschew the problem of racial profiling); Floyd D. Weatherspoon, *Racial Profiling of African-American Males: Stopped, Searched, and Stripped of Constitutional Protection*, 38 J. MARSHALL L. REV. 439 (2004) (lamenting the use of racial profiling by law-enforcement agencies as a technique to selectively enforce laws).

117. There is some support in older cases for allowing race to be used in certain search contexts. See *United States v. Martinez-Fuerte*, 428 U.S. 543, 562 (1976); *United States v. Brigoni-Ponce*, 422 U.S. 873, 886 (1975). Both cases held that in the context of border searches, the Mexican appearance of the defendants could be considered as a factor leading to additional law-enforcement scrutiny. On the whole, though, courts have been very skeptical of the use of race in other criminal law contexts. See Gross & Barnes, *supra* note 50, at 733 (collecting examples of such cases).

between anonymous informants and informants whose identity is known to law enforcement. Anonymous informants are presumed to be less reliable than those who provide their identities, who, by definition, are accountable for their information.¹¹⁸ Similarly, citizen informants are often assumed to be more reliable than those who have been involved in criminal activity.¹¹⁹ There is little or no empirical support for these assumptions, so while they may be true, they may also be false. One might imagine, for instance, that known citizen informants are less likely to be aware of the external indicia of criminal activity while anonymous criminal informants are far less likely to make a mistake. Requiring the collection and reporting of success rates would give us the opportunity to test the assumptions against reality and learn whether courts' fondness for known, civilian informants is well-founded.¹²⁰

Similarly, a success-rate requirement will add to our understanding of allegedly stale evidence. Search cases frequently involve litigation of the issue of whether the information contained in the warrant application has grown stale. In addition to showing that the information in the warrant application suggests that the location once contained contraband, "the proof must be of facts so closely related to the time of the issue of the warrant as to justify a finding of probable cause at that time."¹²¹ The decision of whether the information underlying a particular warrant has grown stale is extremely fact-specific.¹²² Courts have primarily focused on two factors: the time elapsed between the information and the warrant application, and the underlying criminal activity.¹²³ For instance, courts generally have a generous reading of the staleness doctrine in cases involving child pornography; there

118. See *Florida v. J.L.*, 529 U.S. 266, 270 (2000) (distinguishing an anonymous tip from "a tip from a known informant whose reputation can be assessed and who can be held responsible if her allegations turn out to be fabricated"); *Adams v. Williams*, 407 U.S. 143, 146, 146-47 (1972) (explaining that a known informant "is a stronger case than obtains in the case of an anonymous telephone tip").

119. See *United States v. Elmore*, 482 F.3d 172, 180 (2d Cir. 2007) ("The veracity of identified private citizen informants (as opposed to paid or professional criminal informants) is generally presumed in the absence of special circumstances suggesting that they should not be trusted."); *United States v. Angulo-Lopez*, 791 F.2d 1394, 1397 (9th Cir. 1986) ("Citizen informants, while not carrying the same presumption of reliability as police officers, nevertheless require less evidence to establish their veracity than criminal informants." (citations omitted)); *United States v. Gagnon*, 635 F.2d 766, 768 (10th Cir. 1980) ("We have long subscribed to the rule that an affidavit need not set forth facts of a named person's prior history as a reliable informant when the informant is a citizen/neighbor eyewitness with no apparent ulterior motive for providing false information.").

120. The data generated, of course, will not answer the question perfectly. The decision to seek a warrant, as well as the corroboration requirements, will act as intervening variables that make any analysis of the data complicated. However, some data will certainly be more helpful than no data.

121. *Sgro v. United States*, 287 U.S. 206, 210 (1932).

122. *Id.* at 210-11.

123. See, e.g., *United States v. Spikes*, 158 F.3d 913, 923 (6th Cir. 1998) (explaining that the court assesses staleness not only by calendar days, but also by the character of the crime, among other things); *United States v. Ortiz*, 143 F.3d 728, 732 (2d Cir. 1998) (stating that the court examines the age of the facts and the nature of the conduct).

is a consensus view that child-pornography suspects tend not to dispose of incriminating material.¹²⁴ Similarly, allegations of ongoing or sustained criminal conduct can support a warrant application where an allegation relating to a single incident would be inadequate.¹²⁵

There is no immediate reason to decide that these assumptions are wrong, but there is little reason to conclude that they are correct. We have little data demonstrating how long drug dealers retain evidence compared to child pornographers, and our current system does not provide any participant with an incentive to collect that data. A system with a greater focus on law enforcement success rates than on the specific information alleged in the warrant application *would* reach that outcome because law enforcement would be in a position where they would need to learn the rate at which evidence becomes stale or else suffer the consequences of a decline in their success rate.

2. *Generating Data on Racial Profiling.*—Racial profiling has been the focus of the work of economists who have attempted to develop tests for determining its existence in the real world. This work generally builds on a model initially introduced by Professors Knowles, Persico, and Todd (KPT) in 2001.¹²⁶ They criticize traditional methods for identifying discrimination, namely, a regression analysis that controls for other variables thought to be predictive of criminal activities to determine whether race retains any explanatory power.¹²⁷ If researchers cannot observe every variable that might correlate with criminality, they argue, the problem of omitted-variable bias might lead to the incorrect conclusion that race has no impact on the search decision. Additionally, the regression approach has no capacity to distinguish between prejudice-based discrimination and what the economics literature generally refers to as “statistical discrimination.”¹²⁸ Statistical dis-

124. See *United States v. Lacy*, 119 F.3d 742, 745–46 (9th Cir. 1997) (holding that a search warrant based on information ten-months old was valid where an agent offered expert testimony that child-pornography subjects retain material); *United States v. Harvey*, 2 F.3d 1318, 1322–23 (3d Cir. 1993) (rejecting a staleness challenge in part because those who collect child pornography tend to retain material); *United States v. Ricciardelli*, 998 F.2d 8, 12 n.4 (1st Cir. 1993) (“[H]istory teaches that [pornography] collectors prefer not to dispose of their dross, typically retaining obscene materials for years.”).

125. Compare *United States v. Wagner*, 989 F.2d 69, 75 (2d Cir. 1993) (holding that a six-week delay between a one-time drug purchase and the search warrant made the information stale), with *United States v. Wright*, 343 F.3d 849, 864–65 (6th Cir. 2003) (finding a warrant not stale despite a three-year delay given the continuing nature of the criminal activity).

126. See John Knowles et al., *Racial Bias in Motor Vehicle Searches: Theory and Evidence*, 109 J. POL. ECON. 203 (2001).

127. *Id.* at 204–05, 227.

128. *Id.* at 205. Kenneth Arrow is generally credited with developing the concept of “statistical discrimination” to describe discrimination that is not based purely on racial animus. Kenneth J. Arrow, *The Theory of Discrimination*, in *DISCRIMINATION IN LABOR MARKETS* 3, 3 (Orley Ashenfelter & Albert Rees eds., 1973).

crimination refers to race-conscious targeting that occurs because it improves the likelihood of recovering evidence rather than because of racial animus.¹²⁹ Statistical discrimination and animus-based discrimination appear equivalent in regression analyses.¹³⁰

The KPT approach rests on the core assumptions that potential drug dealers take into account the possibility of being searched when they make the decision to participate in criminal activity, and individuals more likely to be searched are less likely to carry narcotics.¹³¹ In turn, unbiased law enforcement allocates search decisions with the aim of increasing the likelihood of recovering evidence, while biased law enforcement considers both the likelihood of recovering evidence and the race of the target.¹³² Biased officers are assumed to prefer discrimination.¹³³ This model implies that, at equilibrium, searches of members of targeted minorities should be less successful than those of majority group members.¹³⁴ All other things being equal, if law enforcement officers are indifferent between searching African-American and Caucasian motorists and only make search decisions based on the likelihood of recovering evidence, drivers will respond and the probability of recovering contraband will equalize across racial groups.¹³⁵ On the other hand, if officers prefer to search African-Americans, those searches will be less successful on average.¹³⁶ KPT applied this model to Maryland search data and concluded that because the likelihood of recovering evidence is equal regardless of whether African-American or Caucasian drivers are searched, animus-based discrimination is unlikely.¹³⁷

The KPT model has been highly influential in the economics literature, and also highly controversial. Perhaps the most serious challenge is that the KPT model overlooks officer-level variations; if officers differ in significant ways, the KPT model fails. First, if some officers like searching more than others, the equality of success rates will not support the conclusion of no discrimination.¹³⁸ Similarly, if different officers have conflicting prejudices,

129. Knowles et al., *supra* note 126, at 205.

130. *See id.* (explaining that regression analyses “are not informative about the underlying motivation,” i.e., whether race is being used as a predictor of criminality or out of preferencing).

131. *Id.* at 206.

132. *See id.* at 205 (explaining how maximizing successful searches by considering race is merely statistical discrimination and unbiased, while racial preferencing is biased law enforcement).

133. *See id.* (“[O]fficers are *racially prejudiced* if . . . they have a preference for searching motorists of a particular race.”).

134. *See id.* at 228 (finding that while success rates of searches are equal across whites and blacks, the lower rate of success with Hispanics indicates a possible bias against Hispanics).

135. *Id.* at 206.

136. *See supra* note 134 and accompanying text.

137. *Id.* at 228. KPT did find that Hispanic drivers were less likely to possess contraband when searched, which raised the possibility of discrimination, but they concluded that the sample size was too small. *Id.*

138. *See Antonovics & Knight, supra* note 103, at 19 (stating that the predictions of the KPT model do not hold up if officers vary in their search costs).

an equality of average recovery rates also loses its significance.¹³⁹ For instance, assume that Caucasian officers are biased against African-American drivers and vice versa. In that situation, everyone suffers discrimination, but equal hit rates will apply across the groups.

Whether these objections limit the KPT model is unclear.¹⁴⁰ What they do suggest, though, is that the best hope for uncovering racial profiling is more data, especially more data at the officer level. This is hardly a new suggestion.¹⁴¹ Proposals made in the past have generally focused on a mandate-based approach.¹⁴² Jurisdictions have begun to require officers to maintain data on the race of the target of the search and the race of the officer, among other factors.¹⁴³ But the simple requirement of maintaining the data does not answer the question of how that data should be used. Currently, no party possesses the incentive to analyze and use data in any specific litigation. Bringing this data into the search process overtly will eliminate that gap.

VI. Success Rates and the Supreme Court

The previous Parts have looked at the value of hit rates when information is hidden from the eyes of courts. This Part considers the potential doctrinal support for their inclusion in the probable-cause analysis. In its line of cases analyzing random roadblock searches of automobiles, the Supreme Court initially considered success rates to be relevant in determining whether such an approach passed constitutional muster. However, in the capstone case, *Indianapolis v. Edmond*,¹⁴⁴ the Supreme Court stepped back from this methodology and instead focused on the purpose of the roadblock rather than its success rate. As a result, this line of cases provides little support for incorporating hit rates into the probable-cause analysis.

However, courts have permitted, and sometimes have required, law enforcement to provide information about success rates of component facts presented in a probable-cause claim. Law enforcement may use success rates to demonstrate that apparently incriminating information is reliable.

139. See Anwar & Fang, *supra* note 75, at 4 (reasoning that, under the KPT model, if minority police officers have a prejudice towards white drivers and white police officers have a prejudice towards minority drivers, then the police force may appear unprejudiced in the aggregate).

140. Two of the authors of the original study respond to some of these objections in Nicola G. Persico & Petra Todd, *Using Hit Rate Tests to Test for Racial Bias in Law Enforcement: Vehicle Searches in Wichita* 7–9 (Nat'l Bureau of Econ. Research, Working Paper No. 10947, 2004), available at <http://ssrn.com/abstract=629589>.

141. Gross & Barnes, *supra* note 50, at 656 (“Racial profiling is impossible to detect or prove without detailed information on police conduct The essential step is to require that the information be recorded and kept in the first place.”).

142. See *id.* at 729 (discussing requirements in Maryland).

143. *Id.*

144. 531 U.S. 32 (2000).

Specifically, below I consider informants and canine sniffs, two circumstances where courts have used prior success rates to determine whether probable cause exists. If courts are willing to consider success rates for the facts underlying the probable-cause claim, there is reason to do the same for the officers actually making the assertion.

A. *Incorporating Success Rates: The Checkpoint Cases*

There is a glimmer of doctrinal support for the use of hit rates in Fourth Amendment cases. For a time, success rates seemed likely to play a key role in the acceptability of random stops of motorists, usually in the context of highway checkpoints.¹⁴⁵ Perhaps surprisingly, though, the success rate affected only the level of constitutional protection available and not the strength of the evidence provided.

The Supreme Court began its line of checkpoint cases in *United States v. Martinez-Fuerte*.¹⁴⁶ The Court considered the constitutionality of United States Border Patrol checkpoints just north of the United States–Mexico border. The Court concluded that while the highway roadblocks constitute a “seizure” under the Fourth Amendment, they are neither subject to the warrant requirement nor is any individualized suspicion required.¹⁴⁷ Consistent with the Court’s related case law,¹⁴⁸ the Court applied a balancing test, comparing the effectiveness of the search to the level of privacy intrusion.¹⁴⁹ The Court noted that “[t]he record . . . provides a rather complete picture of the effectiveness”¹⁵⁰ of the checkpoint. The record showed that 171 vehicles of the 146,000 passing through the checkpoint contained illegal aliens, a hit rate of .12%.¹⁵¹ The Court relied on this record in declining to require a warrant or individualized suspicion before such a seizure can take place.¹⁵²

Although it reached the opposite conclusion, the Court in *Delaware v. Prouse*¹⁵³ took a similar approach by analyzing success rates as part of the effectiveness side of the balancing test in the reasonableness analysis. Acting without individualized suspicion or a warrant, the officer in *Prouse* stopped a driver to perform a random spot check for license and

145. For an excellent, recent analysis of these cases and the value of success rates, see Rosenthal, *supra* note 33, at 671–74.

146. 428 U.S. 543 (1976).

147. *Id.* at 565. While the stop itself was not a search, any subsequent search of the vehicle would be subject to the warrant requirement. *Id.* at 567.

148. *See, e.g.*, *Brown v. Texas*, 443 U.S. 47, 50 (1979) (stating that the reasonableness of a seizure depends “on a balance between the public interest and the individual’s right to personal security free from arbitrary interference by law officers” (quoting *United States v. Brignoni-Ponce*, 422 U.S. 873, 878 (1975))).

149. *Martinez-Fuerte*, 428 U.S. at 554–60.

150. *Id.* at 554.

151. *Id.*

152. *Id.* at 562.

153. 440 U.S. 648 (1979).

registration.¹⁵⁴ Here, however, the Court rejected this procedure after applying the reasonableness balancing test.¹⁵⁵ The *Prouse* Court prohibited these types of stops and mentioned as part of the analysis that they produced evidence at an insufficient rate.¹⁵⁶

The Court's strongest endorsement of the role of success rates came in *Michigan Department of State Police v. Sitz*,¹⁵⁷ which appeared to cement a role for success rates in the effectiveness analysis.¹⁵⁸ *Sitz* involved sobriety checkpoints designed to identify drunk drivers.¹⁵⁹ In concluding the checkpoint was effective, the Court noted that 1.6% of the drivers passing through the checkpoint were arrested for alcohol impairment, which it viewed as comparing favorably to the 0.12% success rate in *Martinez-Fuerte* and the complete absence of evidence in *Prouse*.¹⁶⁰

The Supreme Court's most recent jurisprudence has stepped back from this central focus on success rates. In *Indianapolis v. Edmond*, the Court considered a challenge to a drug interdiction checkpoint operated by the City of Indianapolis.¹⁶¹ Justice O'Connor's opinion led with the fact that the "overall 'hit rate' of the [roadblock] program was . . . approximately nine percent,"¹⁶² but the decision did not turn on this fact. Instead, the Court focused on the "primary purpose" of the roadblocks.¹⁶³ Because detecting ordinary criminal activity was the primary purpose of the roadblocks, the Court refused "to suspend the usual requirement of individualized suspicion."¹⁶⁴ The Court did not overrule *Sitz* or *Martinez-Fuerte*; it simply classified the primary purposes of sobriety and border security as something other than general crime control.¹⁶⁵ For these types of searches, hit rates might still matter. The Court explained, "The constitutionality of such checkpoint programs still depends on a balancing of the competing interests at stake and *the effectiveness of the program*."¹⁶⁶ However, because the searches discussed here are certainly criminal in nature, *Edmond* cuts back substantially on the importance of hit rates.

154. *Id.* at 650.

155. *Id.* at 663.

156. *Id.* at 660. "In terms of actually discovering unlicensed drivers or deterring them from driving, the spot check does not appear sufficiently productive to qualify as a reasonable law enforcement practice." *Id.*

157. 496 U.S. 444 (1990).

158. *See id.* at 455 (holding that a sobriety checkpoint was constitutional based on empirical data showing the arrest rate for alcohol impairment at the checkpoint).

159. *Id.* at 447.

160. *Id.* at 454–55.

161. 531 U.S. 32, 34 (2000).

162. *Id.* at 35.

163. *Id.* at 40.

164. *Id.* at 44.

165. *Id.* at 47.

166. *Id.* (emphasis added).

B. Incorporating Success Rates of Component Parts

In contrast to the checkpoint cases, courts have been willing to look at the history of success in other contexts. Specifically, when courts need to determine the value of particular components of the probable-cause analysis, they do allow law enforcement to present evidence of prior successes. First, when law enforcement makes a probable-cause claim based on informant information, it must present reasons why the court should trust the informant's claim. Law enforcement frequently meets this requirement by showing that information from the informant produced successful results in the past. For instance, in *McCray v. Illinois*,¹⁶⁷ an arresting officer testified that "the informant had supplied him with information about narcotics activities 'fifteen, sixteen times at least,' that the information had proved to be accurate and had resulted in numerous arrests and convictions."¹⁶⁸ Demonstrating that an informant has previously provided evidence leading to convictions—as in *McCray*—generally suffices to demonstrate probable cause,¹⁶⁹ but convictions are not always necessary. The fact that the informant's information has led to the seizure of contraband can also demonstrate an adequate track record.¹⁷⁰

Courts have taken a similar approach with canines. Narcotics-detecting dogs have become a common law enforcement tool for attempting to establish probable cause to search. An alert by a trained narcotics-detecting dog is usually considered sufficient to establish probable cause once the magistrate is convinced that the dog is "reliable."¹⁷¹ Establishing reliability takes a vari-

167. 386 U.S. 300 (1967).

168. *Id.* at 303. The other arresting officer provided similar testimony. *Id.* at 304.

169. *See, e.g.*, *United States v. Taylor*, 471 F.3d 832, 840 (7th Cir. 2006) (noting that the informant's evidence had led to "multiple arrests and convictions"); *United States v. Lucca*, 377 F.3d 927, 933 (8th Cir. 2004) (accepting as probative that the informant "had a track record of supplying reliable information" that resulted in convictions); *United States v. Bynum*, 293 F.3d 192, 197 (4th Cir. 2002) ("[A] proven, reliable informant is entitled to far more credence than an unknown, anonymous tipster."). *See generally* 2 WAYNE R. LAFAVE, SEARCH AND SEIZURE: A TREATISE ON THE FOURTH AMENDMENT § 3.3(b), at 115 n.60 (4th ed. 2004) (collecting cases on this point).

170. *See, e.g.*, *United States v. Gabrio*, 295 F.3d 880, 883 (8th Cir. 2002) ("Reliability may be found on the basis that past tips have led to seizures of contraband or other evidence."); *United States v. Warren*, 42 F.3d 647, 652 (D.C. Cir. 1994) (relying on the fact that the informant's information had previously led to "seizure of large amounts of cocaine [and] about eight guns"). *See generally* LAFAVE, *supra* note 169, at 116–17 nn.71–74 (collecting cases involving various types of contraband, including narcotics, stolen property, counterfeit money, and the bodies of homicide victims).

171. *See United States v. Sundby*, 186 F.3d 873, 876 (8th Cir. 1999) ("A dog's positive indication alone is enough to establish probable cause for the presence of a controlled substance if the dog is reliable."); *United States v. Owen*, 167 F.3d 739, 749 (1st Cir. 1999) ("The existence of probable cause based on an alert by a drug dog depends upon the dog's reliability."); *United States v. Diaz*, 25 F.3d 392, 394 (6th Cir. 1994) ("For a positive dog reaction to support a determination of probable cause, the training and reliability of the dog must be established."); *United States v. Lingenfelter*, 997 F.2d 632, 639 (9th Cir. 1993) ("A canine sniff alone can supply the probable cause necessary for issuing a search warrant if the application for the warrant establishes the dog's

ety of forms; courts have generally held that a simple allegation that the dog is trained and certified is sufficient to establish probable cause and that law enforcement need not provide detailed statistical data on the canine's history.¹⁷² However, courts will often consider the canine's history of success when reliability is challenged,¹⁷³ and law enforcement frequently provides data on the canine's success history as a standard method of establishing reliability.¹⁷⁴

The Supreme Court has taken an approach that is at least implicitly even more reliant on success rates when considering the question of whether exposing items to narcotics-detecting dogs constitutes a search for Fourth Amendment purposes. In *United States v. Place*,¹⁷⁵ the Supreme Court held that exposing seized luggage to a narcotics dog was not a search,¹⁷⁶ the Court extended the *Place* doctrine to automobile stops in *Illinois v. Caballes*.¹⁷⁷ The linked holdings of *Place* and *Caballes* rely heavily on the fact that drug-detecting dogs are highly accurate. In *Place*, the Court concluded that a dog sniff of luggage is sui generis, and does not invade a privacy interest protected by the Fourth Amendment:

A "canine sniff" by a well-trained narcotics detection dog, however, does not require opening the luggage. It does not expose noncontraband items that otherwise would remain hidden from public

reliability." *But see* *United States v. Florez*, 871 F. Supp. 1411, 1419–20 (D.N.M. 1994) (requiring review of the alerting dog's entire individual history of success in every canine case).

172. *See Sundby*, 186 F.3d at 876 ("To establish the dog's reliability, the affidavit need only state the dog has been trained and certified to detect drugs. An affidavit need not give a detailed account of the dog's track record or education."); *United States v. Kennedy*, 131 F.3d 1371, 1377 (10th Cir. 1997) (refusing to require law enforcement "to include a complete history of a drug dog's reliability beyond the statement that the dog has been trained and certified to detect drugs"); *United States v. Berry*, 90 F.3d 148, 153 (6th Cir. 1996) (holding an affidavit sufficient where the dog is described as "trained," even without a discussion of the dog's success history); *United States v. Daniel*, 982 F.2d 146, 151 n.7 (5th Cir. 1993) (rejecting the defendant's argument that the author of an affidavit must prove the dog's past reliability before a warrant is issued).

173. *See United States v. Donnelly*, 475 F.3d 946, 955 (8th Cir. 2007) (holding that because the canine's "performance record raises questions about his reliability, further inquiry was necessary"); *Kennedy*, 131 F.3d at 1377 (indicating that a court may consider accuracy rates when the dog's reliability is challenged). *But cf.* *United States v. Boxley*, 373 F.3d 759, 762 (6th Cir. 2004) (holding that when the dog's alert is substantive evidence in a case, "after it is shown that the dog is certified, all other evidence relating to his accuracy goes only to the credibility of the testimony, not the dog's qualifications").

174. *See Donnelly*, 475 F.3d at 955 (finding probable cause based on an alert from a dog with a false-positive rate of 46%); *United States v. Koon Chung Wu*, No. 06-4172, 2007 WL 412169, at *6 (4th Cir. Feb. 2, 2007) (40% false-positive rate); *United States v. Limares*, 269 F.3d 794, 798 (7th Cir. 2001) (38% false-positive rate); *Kennedy*, 131 F.3d at 1378 ("70%–80% success rate"); *United States v. Scarborough*, 128 F.3d 1373, 1378 (10th Cir. 1997) (8% false-positive rate). For an excellent summary (and critique) of the recent case law on canine reliability, see Richard E. Myers II, *Detector Dogs and Probable Cause*, 14 GEO. MASON L. REV. 1, 20–22 (2006).

175. 462 U.S. 696 (1983).

176. *Id.* at 707.

177. 543 U.S. 405, 409 (2005).

view, as does, for example, an officer's rummaging through the contents of the luggage. Thus, the manner in which information is obtained through this investigative technique is much less intrusive than a typical search. Moreover, the sniff discloses only the presence or absence of narcotics, a contraband item. Thus, despite the fact that the sniff tells the authorities something about the contents of the luggage, the information obtained is limited. This limited disclosure also ensures that the owner of the property is not subjected to the embarrassment and inconvenience entailed in less discriminate and more intrusive investigative methods.¹⁷⁸

The Court extended this reasoning in *Caballes*. It concluded that exposing an automobile to a narcotics-dog sniff was not a search and relied on the sui generis language from *Place*.¹⁷⁹ Despite a strong objection from Justice Souter in dissent,¹⁸⁰ the Court has concluded that drug-sniffing dogs only alert to the presence of narcotics, in which an individual has no protected privacy interest.

The Court's holding in *Caballes* has been subjected to substantial criticism,¹⁸¹ particularly in light of the rule that a dog alert is sufficient to

178. *Place*, 462 U.S. at 707.

179. *Caballes*, 543 U.S. at 409.

180. *Id.* at 411 (Souter, J. dissenting) ("The infallible dog, however, is a creature of legal fiction. . . . [T]heir supposed infallibility is belied by judicial opinions describing well-trained animals sniffing and alerting with less than perfect accuracy . . .").

181. In addition to doctrinal critiques, there is a basic Bayesian problem with the notion that a dog alert is sufficient for probable cause. See Robert C. Bird, *An Examination of the Training and Reliability of the Narcotics Detection Dog*, 85 KY. L.J. 405, 427 (1997) (explaining that the number of false-positive alerts is likely to greatly exceed the number of accurate alerts even given an error rate (false positive and false negative) of just two percent); Myers, *supra* note 174, at 15 (demonstrating that, according to Bayes Theorem, the "gross number of searches that result from the error rate [may] be higher than the gross number of searches that result from correct alerts"). The fact that a dog with a 5% false-positive rate (sometimes described as a 95% accuracy rate) has alerted on a particular location does not mean that location will contain contraband 95% of the time. Consider a different version of Bayes Theorem, expressed in terms of odds ratios:

$$\frac{P(C | E)}{P(-C | E)} = \frac{P(C)}{P(-C)} * \frac{P(E | C)}{P(E | -C)}$$

We assume here that the false-positive rate is 5%, or .05. We also need the true-positive rate, $P(E | C)$, which we will assume to be 90%, or .90. This figure implies a false-negative rate of 10%: When narcotics are present, the dog fails to alert 10% of the time. So we have .90 divided by .05, which produces a likelihood ratio of 18.

$$\frac{P(C | E)}{P(-C | E)} = 18 * \frac{P(C)}{P(-C)}$$

This calculation has a very straightforward and intuitive interpretation. Whatever the original odds are of recovering narcotics from a target, the addition of a dog alert makes it eighteen times more likely that narcotics will be found. Whether a dog alert means that a location is sufficiently likely to contain contraband to exceed the probable-cause standard, though, depends crucially on the prior odds:

$$\frac{P(C)}{P(-C)}$$

establish probable cause on its own.¹⁸² For instance, it makes little sense to allow law enforcement to present evidence of prior successes without considering the number of attempts by the informant or the canine.¹⁸³ But *McCray*, *Caballes*, and *Place* give support for the idea that history can matter. Looking at the past success rates of the components of a probable-cause claim is the equivalent of a baseball manager looking at a hitter's height and weight when trying to decide whether to draft him.¹⁸⁴ They probably bear some relationship to his performance during a game, but better information is available.¹⁸⁵

VII. Objections to Overtly Probabilistic Measures

While the use of success rates in the Fourth Amendment context has received very little attention, scholars have raised objections to the use of overtly probabilistic measures in other legal contexts. These objections have generally fallen into two forms: capacity objections and normative objections.¹⁸⁶ The 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.¹⁸⁷ Normative objections turn on the appropriate nature of legal rules. For instance, Charles Nesson has claimed that the use of statistics reduces the acceptance of jury verdicts

Assume we believe that the probable-cause standard lies at the 50% threshold and we permit searches whenever the probability of recovering contraband given the dog sniff exceeds the probability of not recovering contraband. This will be true whenever the prior odds of guilt are greater than one-nineteenth, or when the pre-dog-sniff probability that the location contains contraband is at least 5%. Consider the facts of *Place* and *Caballes*. It may be true that at least 5% of the individuals who provide false information to airlines are transporting narcotics. See *Place*, 462 U.S. at 699 (determining that an airport *Terry* stop in order to subject luggage to a dog sniff was acceptable). By contrast, it is difficult to believe that at least 5% of the vehicles that are going slightly over the speed limit on Interstate 80 contain drugs. See *Caballes*, 543 U.S. at 412 (Souter, J., dissenting) (arguing that the fallibility of drug dogs turns sniffs that are incident to mere traffic stops into full searches subject to the Fourth Amendment requirement of probable cause).

182. See *supra* note 171.

183. See Charles E. Moylan, Jr., *Hearsay and Probable Cause: An Aguilar and Spinelli Primer*, 25 MERCER L. REV. 741, 759 (1974) (“If the furnishing of good information in the past contributes to a belief in an informant’s credibility, the furnishing of bad information in the past would certainly derogate therefrom. The policeman who works with an informant knows of his full batting average, not just of his successes.”).

184. Cf. MICHAEL LEWIS, *MONEYBALL: THE ART OF WINNING AT AN UNFAIR GAME* 16 (2003) (describing how statistical odds of Major League success were against a high-school baseball player, though he had a body “that looked as if it had been created to wear a baseball uniform”).

185. *Id.*

186. I draw this categorization of objections from Professor Tracey L. Meares’ article, *Three Objections to the Use of Empiricism in Criminal Law and Procedure—and Three Answers*, 2002 U. ILL. L. REV. 851.

187. *Id.* at 854.

and that their use is inherently immoral.¹⁸⁸ Similarly, other authors argue that the use of statistics is inherently depersonalizing.¹⁸⁹ These objections have received frequent and vigorous rebuttals in the academic literature.¹⁹⁰

In this Part, I argue that even if these objections carry weight in other legal contexts, the search process is fundamentally different for three reasons. First, while the trial process is designed to make a truth statement about a past event, the warrant process is explicitly probabilistic. The judge is instructed to determine whether there is a sufficient probability that the identified location contains contraband. Second, unlike the trial process, the search process has a well-defined reference class—previous searches by the same law enforcement entity. Finally, the search process provides an audit function. Unlike the trial process, we learn whether the initial determination was right.

A. *Capacity Objections to the Use of Statistical Evidence*

The capacity objection is perhaps the oldest challenge to the use of overtly probabilistic evidence.¹⁹¹ Judges and lawyers are not trained in the

188. See Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357, 1378 (1985) (pointing out that there is not an exact correlation between statistical probability and acceptability: “[A] probable verdict may not be acceptable, and an acceptable verdict may not be probable.”).

189. See Kevin M. Clermont, *Procedure’s Magic Number Three: Psychological Bases for Standards of Decisions*, 72 CORNELL L. REV. 1115, 1147 (1987) (asserting that quantification arguably leads to “dehumanization of the legal process”).

190. See, e.g., Ronald J. Bacigal, *Making the Right Gamble: The Odds on Probable Cause*, 74 MISS. L.J. 279, 298 (2004) (“Once we acknowledge that innocents may be searched and seized, any differentiation between ‘individualized’ case-specific evidence and ‘statistical’ evidence is largely illusory.”); see also 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, 248 (1990) (arguing that overtly probabilistic evidence is “no less probative of legally material facts than other types of evidence”); Meares, *supra* note 186, at 866 (responding to the argument that the field of criminal theory should be thought of as a humanist inquiry rather than a social science); Daniel Shaviro, *Statistical-Probability Evidence and the Appearance of Justice*, 103 HARV. L. REV. 530, 531 (1989) (rejecting the holding in *Smith v. Rapid Transit*, 58 N.E.2d 754 (1945), that courts should refuse to base verdicts in favor of plaintiffs or prosecutors entirely on statistical-probability evidence).

191. See, e.g., Samuel M. Fahr & Ralph H. Ojemann, *The Use of Social and Behavioral Science Knowledge in Law*, 48 IOWA L. REV. 59, 74–75 (1962) (positing that lawyers’ distrust of empirical social-science data stems from the wide disparities and lack of accountability in the research process); Phillip R. Lochner, *Some Limits on the Application of Social Science Research in the Legal Process*, 1973 L. & SOC. ORD. 815, 824–27 (explaining that social-science research is limited in its application to the legal process because of lawyers’ unfamiliarity with the methodology of social science and the potential uncertainties in social-science studies); David M. O’Brien, *The Seduction of the Judiciary: Social Science and the Courts*, 64 JUDICATURE 8, 13, 12–13 (1980) (opining that judicial decision making based on empirical social-science data presents “more obstacles than opportunities”); Michael Rustad & Thomas Koenig, *The Supreme Court and Junk Social Science: Selective Distortion in Amicus Briefs*, 72 N.C. L. REV. 91, 93–95, 152 (1993) (contending that the Supreme Court is ill-equipped to determine how much weight should be given to statistical arguments made to the Court through amicus curiae briefs, which frequently present empirical data in a distorted, self-interested manner).

use of statistical evidence and when it is used, the results have been mixed. Most prominently, Chief Justice Warren cited Kenneth Clark's famous doll study in footnote 11 of *Brown v. Board of Education*.¹⁹² Clark's research methodology and results have subsequently been challenged by later research.¹⁹³ More significantly, the argument goes, the citation of the doll study represents an even more fundamental error in reasoning. Since the *Brown* decision is right even if the doll study was wrong, the social-science evidence is simply irrelevant to the decision.

More recently, cognitive-psychology research has attempted to investigate the mechanisms by which we generate conclusions from aggregations of facts. Most of this work has focused on the trial process and jurors' attempts to reach conclusions about guilt and innocence (or the presence or absence of legal liability). The standard model of jury decision making involves a Bayesian-style approach—jurors begin a case with a prior probability of guilt or liability which they then update based on the facts presented to them during the course of the trial.¹⁹⁴ These facts, when combined with the juror's prior conceptions, produce a posterior expectation of guilt or liability.¹⁹⁵ If this posterior exceeds the relevant truth threshold—either a preponderance standard in civil cases or proof beyond a reasonable doubt in criminal cases—the jurors vote in favor of the party with the burden of proof.

This standard model is now being called into significant doubt in a variety of experiments. The most frequently cited countertheory travels under the names “narrative theory,” the “story model,” or “holistic reasoning.”¹⁹⁶ This theory suggests that legal decision makers do not consider pieces of evidence individually but instead incorporate them holistically into larger

192. 347 U.S. 483, 495 n.11 (1954).

193. See, e.g., ROY L. BROOKS, INTEGRATION OR SEPARATION?: A STRATEGY FOR RACIAL EQUALITY 18 (1996) (calling Clark's study “methodologically flawed” for failing to distinguish between personal and group identity); Phyllis A. Katz & Sue Rosenberg Zalk, *Doll Preferences: An Index of Racial Attitudes*, 66 J. EDUC. PSYCHOL. 663, 663–64 (1974) (asserting that doll-preference tasks, though generally accepted by many scholars, raise a number of interpretive and methodological problems); Mark G. Yudof, *School Desegregation: Legal Realism, Reasoned Elaboration, and Social Science Research in the Supreme Court*, LAW & CONTEMP. PROBS., Autumn 1978, at 57, 62 (referring to federal courts' disinclination to follow the social-science evidence from *Brown* as “fortunate”).

194. See generally Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. CHI. L. REV. 511, 559 (2004) (describing Bayes Theorem); Stephen E. Fienberg & Mark J. Schervish, *The Relevance of Bayesian Inference for the Presentation of Statistical Evidence and for Legal Decisionmaking*, 66 B.U. L. REV. 771 (1986) (discussing the normative application of Bayes Theorem to the law).

195. See, e.g., Simon, *supra* note 194, at 559 (“[Under Bayes Theorem,] verdicts are determined by comparing the posterior likelihood produced by the Bayesian computation with the numerical assigned to the respective standard of proof.”).

196. 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, 1087–88 (2007) (discussing the use of various names for the theory); Simon, *supra* note 194, at 561 (explaining the theory and calling it “the story model”).

narratives.¹⁹⁷ Nancy Pennington and Reid Hastie's work is the backbone of this research. They suggest that the mechanism at work involves the jurors' constructing narratives out of multiple inputs, including the evidence presented at trial, the jurors' backgrounds, and the jurors' preexisting expectations about narratives.¹⁹⁸ Ronald Allen has not only taken the view that the holistic approach describes how jurors actually behave, but he argues that the approach is a more appropriate method for deciding cases than a Bayesian method.¹⁹⁹

The second aspect of the capacity objection turns on data availability. Good data-based decision making requires good data; in particular, the use of a base rate requires specifying an appropriate reference class.²⁰⁰ In the search-warrant context, should we consider the success rate of this officer, this police department, or search warrants by any law enforcement agency? Does the location to be searched matter? Are homes and cars fundamentally different from wiretaps? This is not a merely theoretical objection—the specificity of the reference class matters in the real world. The most comprehensive review of base rates in legal decision making found that when “the reference classes for proffered statistics are closely linked to a focal issue or a defendant, courts are more likely to find the base rates to be relevant.”²⁰¹ Reference classes matter for the public at large as well. Experimental research shows that even statistically untrained observers are more responsive to base-rate evidence when the reference class matches the case presented to them more specifically.²⁰²

These capacity objections carry less weight in the search context than in the trial context. First, the statistical evidence here is exclusively presented to judges, rather than jurors, and specifically to magistrate judges who are

197. See Simon, *supra* note 194, at 561 (summarizing Pennington and Hastie's story model).

198. Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 *CARDOZO L. REV.* 519, 522 (1991).

199. Ronald J. Allen, *Factual Ambiguity and a Theory of Evidence*, 88 *NW. U. L. REV.* 604, 604 (1994). Along similar lines, Kevin Jon Heller has recently brought research on the Wells Effect into the legal literature. See Kevin Jon Heller, *The Cognitive Psychology of Circumstantial Evidence*, 105 *MICH. L. REV.* 241, 258 (2006). The Wells Effect suggests that jurors give more weight to direct, but potentially flawed, testimony than purely circumstantial evidence, even if the probabilistic weight of the evidence is identical. *Id.* at 256–58. Wells suggests that this effect results from an “ease of simulation.” See Keith E. Niedermeier et al., *Jurors' Use of Naked Statistical Evidence*, 76 *J. PERSONALITY & SOC. PSYCHOL.* 533, 538 (1999) (explaining that jurors give weight to evidence partially based on how easy it is for them to imagine specific circumstances after considering different types of evidence).

200. See L.J. Cohen, *Are There Any A Priori Constraints on the Study of Rationality*, 4 *BEHAV. & BRAIN SCI.* 359, 365 (1981) (providing two examples in which comparison to an inappropriate reference class would lead to “absurd” conclusions).

201. Jonathan J. Koehler, *When Do Courts Think Base Rate Statistics Are Relevant?*, 42 *JURIMETRICS* 373, 393 (2002).

202. Jonathon J. Koehler, *The Base Rate Fallacy Reconsidered*, 19 *BEHAV. & BRAIN SCI.* 1, 10 (1996); Amos Tversky & Daniel Kahneman, *Causal Schemas in Judgments Under Uncertainty*, 1 *PROGRESS SOC. PSYCHOL.* 49, 61 (1980).

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. Judges already make the complicated calculation of evaluating the weight to be given to a law enforcement officer's belief that a subject's activity was "furtive."²⁰³ Analyzing a base rate is easier—not harder—than that decision.

Furthermore, the data in the search-warrant context are simpler and more standardized. The easy answer to the reference-class objection, regardless of the situation, is that the best available base rate should be used. If the best available rate is inadequately specific, it should be given little weight, while a highly specific base rate should be given substantial weight.²⁰⁴ More fundamentally, the relevant base rate is always the one most closely tied to the phenomenon to be measured. In the search-warrant context, we should select the base rates that capture the significant, otherwise uncaptured information discussed above—the positive and negative information that we know shapes the decision to seek a warrant.

Two specific rates stand out as especially pertinent to the probable-cause inquiry. The first is the success rate of the individual officer presenting the warrant. At its core, a search-warrant application is a claim by the presenting officer that there is probable cause to believe the target contains relevant evidence. That officer's prior success rate, then, is the relevant rate to consider. One claim above is that officers often possess information at the individual level that is difficult or impossible to present in a warrant application—information about the reliability of any informants, the officer's own personal skill and effort in the investigation so far, and any hidden racial animus shaping the search decision. Since we are interested in individual-level effects, we should use an individual-level statistic.

The second major claim above is that law enforcement priorities and resource constraints shape the success rates of searches. Law enforcement is more restrained when budgets bind more tightly, and it is more aggressive when money is freely available. Additionally, law enforcement officials seek more marginal warrants when they care more about the underlying offense, but they self-impose more restraint for less important crimes. As a result, the relevant reference class should be the law enforcement unit that is subject to the budget constraint and is setting the policy priorities. This class needs to be evaluated differently for different governmental entities.

203. See *supra* note 111 and accompanying text.

204. See Koehler & Shaviro, *supra* note 190, at 261 (noting that, while there is no level of specificity required for a base rate to be relevant, base rates derived from relatively unspecific reference classes may be associated with greater second-order uncertainty).

B. Normative Objections to the Use of Statistical Evidence

Normative objections to the use of statistical evidence come in two forms. The first form is an acceptance objection involving a claim about the goal of the adjudicative process. The most straightforward goal of the trial process is to reach the correct outcome; trials should result in convictions if and only if the defendants are guilty. Academics raising the acceptance objection claim that a secondary goal of the verdict process—generating publicly acceptable outcomes—precludes the use of explicitly probabilistic evidence even when it reaches the most accurate outcome.²⁰⁵ The second form of objection turns on individuation. Lawrence Tribe has claimed that the rejection of purely probabilistic evidence acts as a prophylactic rule standing in the way of society's treating individual defendants as expendable.²⁰⁶ In essence, this argument turns on the claim that we need to treat individuals as individuals rather than as members of a reference class and, in particular, individuals cannot ethically be judged based on attributes outside their control.²⁰⁷ This normative commitment to individuation is essentially the position adopted by the Supreme Court in *McCleskey v. Kemp*.²⁰⁸ In *McCleskey*, the defendant raised a challenge to the Georgia capital-sentencing scheme by relying on the now-famous Baldus study,²⁰⁹ which showed that black murderers of white victims were more likely to be sentenced to death than white murderers of black victims.²¹⁰ The Court rejected the defendant's Eighth and Fourteenth Amendment challenges, in large part because they would apply to any capital case in which "the victim was white and the defendant is black."²¹¹

Nesson analyzes these questions in the context of the famous Blue Bus example.²¹² Based on the facts in *Smith v. Rapid Transit*,²¹³ the example hy-

205. Nesson, *supra* note 188, at 1377–78.

206. See Lawrence H. Tribe, *An Ounce of Detention: Preventative Justice in the World of John Mitchell*, 56 VA. L. REV. 371, 387 (1970) ("[T]he very enterprise of formulating a tolerable ratio of false convictions to false acquittals puts an explicit price on an innocent man's liberty and defeats the concept of a human person as an entity with claims that cannot be extinguished, however great the payoff to society.").

207. See Lea Brilmayer & Lewis Kornhauser, *Review: Quantitative Methods and Legal Decisions*, 46 U. CHI. L. REV. 116, 149–50 (1978) (highlighting the argument that individuals should not be culpable for their statuses, a characteristic beyond their control); Barbara D. Underwood, *Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgment*, 88 YALE L.J. 1408, 1427 (1978) ("The claim is that in some situations the individual applicant is entitled to a decision that not only contributes to a useful aggregate, but also reflects consideration of the distinctive aspects of his individual case.").

208. 481 U.S. 279, 311 (1987) ("[I]t is the jury's function to make the difficult and uniquely human judgments that defy codification . . .").

209. *Id.* at 293.

210. DAVID C. BALDUS, GEORGE WOODWORTH & CHARLES A. PULASKI, JR., *EQUAL JUSTICE AND THE DEATH PENALTY* 400–01 (1990).

211. *McCleskey*, 481 U.S. at 293, 293–94.

212. Nesson, *supra* note 188, at 1378–83.

pothesizes a late-night car accident in which a driver is run off the road by a vehicle he can tell is a bus, but cannot identify any further.²¹⁴ The driver brings suit against the Blue Bus Company for his injuries arising in the accident.²¹⁵ His only evidence tying the defendant to the collision is the naked statistical base rate—the Blue Bus Company operates 80% of the buses that travel on that road.²¹⁶ While it is more likely than not that the defendant injured the plaintiff, courts will generally grant directed verdicts for defendants on these facts.²¹⁷

Nesson justifies the Blue Bus result by pointing to the need for the acceptance of verdicts by the public, current litigants, and future litigants as true statements about past events.²¹⁸ He sees a verdict as not merely a statement about evidence presented at trial, but a conclusion about what actually transpired.²¹⁹ Juries, acting as a black box, provide closure for the public by establishing an official statement about what happened in the past.²²⁰ Since a verdict based on purely statistical evidence can never present more than a probabilistic statement, it cannot generate the necessary public acceptance to support a verdict.²²¹ In Nesson's eyes, not only does acceptance of verdicts by the public as a whole turn on the use of nonstatistical evidence, but future litigants will not alter their conduct based on purely statistical evidence.²²² The deterrent value of the law will be lost if verdicts arise from such evidence.²²³

The argument that public acceptance of verdicts requires nonprobabilistic evidence does not apply in the search context because the goal of probable-cause analysis is fundamentally different than that of the trial process. First, trials are *retrospective*. They attempt to make a truth

213. 58 N.E.2d 754 (Mass. 1945). Compare *id.* with *Howard v. Wal-Mart Stores, Inc.*, 160 F.3d 358, 360 (7th Cir. 1998) (distinguishing a case in which a customer of a store injured herself by slipping on soap when there was no evidence that an employee spilled the soap from *Smith* by noting that, unlike in *Smith*, there was no reason to suspect the plaintiff was holding back unfavorable evidence or to expect the plaintiff to conduct a thorough investigation).

214. Nesson, *supra* note 188, at 1378.

215. *Id.* at 1379.

216. *Id.*

217. *Id.* at 1379 n.70. Nesson offers a similar example from criminal law. Imagine that twenty-four prisoners conspire to commit a murder in an enclosed yard while a twenty-fifth prisoner hides and is not involved. If the beyond-a-reasonable-doubt standard is met with a 95% level of certainty, all twenty-five prisoners can be convicted on that evidence alone, even though we are certain one defendant is innocent. Charles R. Nesson, *Reasonable Doubt and Permissive Inferences: The Value of Complexity*, 92 HARV. L. REV. 1187, 1192–93 (1979).

218. See Nesson, *supra* note 188, at 1368 (“Many of the procedures of our legal system are best understood as ways to promote public acceptance of verdicts.”).

219. *Id.* at 1363.

220. *Id.* at 1362.

221. *Id.*

222. *Id.*

223. *Id.* at 1368.

statement about a past event—determining whether a particular individual did or did not commit a particular crime. Probable-cause analysis, on the other hand, is *prospective*—it evaluates the likelihood of evidence’s being found in a particular location. It is a claim about the future, not the past. Moreover, it is a claim about a location, not a person.²²⁴ The warrant application does not claim that a particular individual is guilty, but instead that a particular location contains evidence. As a result, the depersonalization claim carries far less force here than in the trial context.

Second, as noted above, the probable-cause determination is explicitly probabilistic. We do not expect it to serve the same validating function as the trial process. These two features combine to give us something better than the black-box determination by the trial process. We have an *audit*—we can know with certainty whether the prior probabilistic prediction was accurate or inaccurate. Judges and the public can learn whether the search recovered evidence. That process, rather than abstract concerns about acceptance and individuation, provides the necessary mechanism for public acceptance.²²⁵

VIII. Conclusion

Success rates will help the probable-cause decision when magistrates do not learn all of the important facts about the probable-cause claim. The divergence in the success rates of searches with and without warrants strongly suggests that magistrates are not now fully informed. Adding in success rates will mean that magistrates would begin to apply additional scrutiny to law enforcement probable-cause claims for searches that generally fail. Magistrates would become more skeptical of car-stop searches but become more lenient when law enforcement seeks a warrant. There is every reason to think this is a positive development. The Supreme Court has often expressed a desire to create incentives for law enforcement to seek warrants rather than engage in warrantless searches.²²⁶ Focusing on success rates does

224. *See, e.g.*, *United States v. Harris*, 403 U.S. 573, 584 (1971) (“[T]he issue in warrant proceedings is not guilt beyond reasonable doubt but probable cause for believing the occurrence of a crime and the secreting of evidence in specific premises.”).

225. Under the current system, judges in fact can learn whether the warrants they issue recover evidence. The Federal Rules require agents to file returns listing an inventory of the material seized in a search warrant. FED. R. CRIM. P. 41(f). The fact that information is available does not mean that it is used: Based on my experiences as a federal prosecutor, returns are seldom scrutinized by judges or lawyers.

226. *See, e.g.*, *Ornelas v. United States*, 517 U.S. 690, 699 (1996) (“The Fourth Amendment demonstrates a ‘strong preference for searches conducted pursuant to a warrant’” (quoting *Illinois v. Gates*, 462 U.S. 213, 236 (1983))); *Murray v. United States*, 487 U.S. 533, 539–40 (1988) (defending the independent-source rule against a challenge that it would encourage law-enforcement officers to search premises without a warrant by arguing that warrantless entry would be “foolish” because the entry would create a far higher burden at trial to get any evidence discovered as a result of the search admitted).

this in a very direct manner—it provides additional scrutiny exactly where it is needed.

IX. Mathematical Appendix

The following analysis develops the intuitive concepts outlined in Part IV. Specifically, I assume that some facts affecting the probability of recovery in a potential search are publicly known and can be observed by the court and law enforcement. Others are not. For instance, while officers are required to disclose information about informants, they may have additional, private information suggesting that the informant is not as accurate as the data suggest. Alternatively, law enforcement sometimes engages in perjury. This Appendix demonstrates more formally that when courts have the opportunity to engage in repeated observations of law enforcement, eventually they will be able to determine which officers make accurate probable-cause claims and which do not. The model only requires that the inaccurate claims are concentrated and substantial, i.e., that the same officers repeatedly make inaccurate claims, and that when these inaccurate claims happen, they substantially increase the apparent likelihood of recovering evidence.

Assume that law enforcement knows the true likelihood of recovering evidence based on the information it possesses. Assume further that police officers come in two types, Good (G) and Bad (B). Good law enforcement officers do not add false information (or fail to correct misleading information) while Bad law enforcement officers enhance their presentations in their warrant requests. Specifically, represent the probability of Bad law enforcement's recovering evidence in a series of n search claims as a series of independent, identically distributed random variables $X_1, X_2 \dots X_n$ drawn on the interval $[0,1]$ pursuant to some distribution function $F(x)$ and a density function $f(x)$ with mean $x \in (0,1)$ and variance σ_1^2 . Similarly, represent the probability of Good law enforcement's recovering evidence in a series of n search claims as a series of independent, identically distributed random variables $Y_1, Y_2 \dots Y_n$ drawn on the interval $[0,1]$ pursuant to some distribution function $G(x)$ and a density function $g(x)$ with mean $\hat{x} \in (0,1)$ and variance σ_2^2 . Assume $\hat{x} > x$. We can see \hat{x} as the probability of recovering evidence based solely on the publicly available information about the case, x is the actual probability of recovering evidence, and $t = \hat{x} - x$ represents the potential additional value (dishonestly) added to the probable-cause calculation by Bad law enforcement. For Good law enforcement, $\hat{x} = x$, and the actual probability of recovery is equal to the publicly available information. In contrast, for Bad law enforcement, $\hat{x} - t = x$, so the actual probability of recovering evidence is less than the publicly available information would suggest. The total number of recoveries for Good and Bad

officers after n searches is simply a random variable reflecting the sum of the probabilities of recovery in each individual search, i.e.:

$$\sum_{i=1}^n Y_i \text{ and } \sum_{i=1}^n X_i$$

According to the Central Limit Theorem, these sums will converge to normal distributions with means $n\hat{x}$ and $nx = n(\hat{x} - t)$ and, respectively, variances of:

$$\frac{\sigma_2^2}{n} \text{ and } \frac{\sigma_1^2}{n}$$

Our ability to detect the divergence in the resulting histories of success for Good and Bad officers depends on the size of t , the size of n , and the variances σ_1^2 and σ_2^2 .

As an example, consider the discussion in the main text, where we assume that each draw for Good and Bad law enforcement represents a Bernoulli trial with the respective parameters \hat{x} and $x = \hat{x} - t$. Assume that officer Good has a probability of success of $1 - \gamma$, while Bad has a probability of success of γ . After $n - 1$ searches, let $\mu_n(r_{n-1}; BAD, \hat{x})$ have its natural meaning as the probability of observing r recoveries in $n - 1$ searches given the observed probability \hat{x} , and that the officer is type Bad, and let similar constructions have similar meanings. The court is attempting to determine, given the evidence presented to it, the probability that the officer is of the type Bad, i.e., $\mu_n(BAD; r_{n-1}, \hat{x})$. To do so, the court takes a Bayesian approach, which, for ease of presentation, we consider through the odds ratio:

$$\frac{\mu_n(BAD; \hat{x}, r_{n-1})}{\mu_n(GOOD; \hat{x}, r_{n-1})} = \frac{\mu_n(BAD)\mu_n(r_{n-1}; BAD, \hat{x})}{\mu_n(GOOD)\mu_n(r_{n-1}; GOOD, \hat{x})} = \frac{\gamma\mu_n(r_{n-1}; BAD, \hat{x})}{(1-\gamma)\mu_n(r_{n-1}; GOOD, \hat{x})}$$

$\mu_n(r_{n-1}; BAD, \hat{x})$ and $\mu_n(r_{n-1}; GOOD, \hat{x})$ are simply the probabilities of seeing r_{n-1} recoveries in $n - 1$ Bernoulli trials with parameters $\hat{x} - t$ and \hat{x} , respectively. That is:

$$\mu_n(r_{n-1}; GOOD, \hat{x}) = {}_{(n-1)}C_r \hat{x}^r (1-\hat{x})^{n-r}, \quad \text{and} \\ \mu_n(r_{n-1}; BAD, \hat{x}) = {}_{(n-1)}C_r (\hat{x} - t)^r (1 - (\hat{x} - t))^{n-r}$$

So we have:

$$\begin{aligned} \frac{\mu_n(BAD; \hat{x}, r_{n-1})}{\mu_n(GOOD; \hat{x}, r_{n-1})} &= \frac{\gamma_{(n-1)} C_r (\hat{x} - t)^r (1 - (\hat{x} - t))^{n-r}}{(1 - \gamma)_{(n-1)} C_r \hat{x}^r (1 - \hat{x})^{n-r}} \\ &= \frac{\gamma (\hat{x} - t)^r (1 - (\hat{x} - t))^{n-r}}{(1 - \gamma) \hat{x}^r (1 - \hat{x})^{n-r}} \\ &= \frac{\gamma}{(1 - \gamma)} * \left[\frac{\hat{x} - t}{\hat{x}} \right]^r * \left[\frac{1 - (\hat{x} - t)}{1 - \hat{x}} \right]^{n-r} \end{aligned}$$

The value of this term, of course, depends on the number of recoveries r , the number of observations n , and both \hat{x} and t , the amount by which the apparent probable cause is inflated. In practice, of course, courts will not be able to observe all of this data. Courts will see the history of individual officers (r and n) and the history of law enforcement as a whole. Assuming that the substantial majority of searches are performed by Good officers, the success rate of law enforcement as a whole will serve as a measure of \hat{x} . Courts will be forced to guess at γ and t .

To clarify the analysis, consider the same numeric example provided in the main text. Consider $\hat{x} = .8$ and $t = .4$, i.e., the face of the officer's prior warrant applications appear to present an 80% probability of recovering evidence while the private information in the possession of Bad law enforcement cuts the actual probability of recovering evidence to 40%. We can then evaluate the Bayes Factor,

$$\left[\frac{(\hat{x} - t)}{\hat{x}} \right]^r * \left[\frac{1 - (\hat{x} - t)}{1 - \hat{x}} \right]^{n-r} = \left[\frac{(.8 - .4)}{.8} \right]^r * \left[\frac{1 - (.8 - .4)}{1 - .8} \right]^{n-r} = .5^r * 3^{n-r}$$

for various numbers of successes and failures. Assume we have twenty observations for a given officer. For ten or fewer successes ($r \leq 10$), the Bayes Factor is greater than 57, providing very strong evidence that the applicant is Bad. Similarly, for fourteen or more successes ($r \geq 14$), the Bayes Factor is less than .045, providing very strong evidence that the applicant is Good. Only for observations of 11, 12, and 13 successes are the Bayes Factors relatively unclear, taking on values of 9.6, 1.6, and .27, respectively. With relatively few observations and only moderate levels of misinformation contained in the application, the success rates of Good and Bad law enforcement diverge rapidly.

Alternatively, consider the distribution of successes in graphical form, in Figure 5 below. The following Figure represents the probability of seeing a given number of successes after only ten searches for Good and Bad law enforcement respectively. If the Court observes exactly six successes, the data is difficult to interpret, but that result takes place only about 10% of the time; for all other observations, the data strongly identifies the officer's type.

As noted in the main text, then, this model is only effective under certain circumstances. Perjury must be concentrated among a few officers, and those officers must tell substantial lies when they do testify falsely. If perjury is dispersed widely (that is, officers draw from both the distributions $F(x)$ and $G(x)$) or only involves small falsehoods (t is small), this model will fail to detect it.

Figure 5: Success Rates of Hypothetical Good and Bad Law Enforcement

