2024

The Automated Fourth Amendment

Maneka Sinha

THE AUTOMATED FOURTH AMENDMENT

Maneka Sinha[*]

ABSTRACT

Courts routinely defer to police officer judgments in reasonable suspicion and probable cause determinations. Increasingly, though, police officers outsource these threshold judgments to new forms of technology that purport to predict and detect crime and identify those responsible. These policing technologies automate core police determinations about whether crime is occurring and who is responsible.

Criminal procedure doctrine has failed to insist on some level of scrutiny of—or skepticism about—the reliability of this technology. Through an original study analyzing numerous state and federal court opinions, this Article exposes the implications of law enforcement's reliance on these practices given the weighty interests that hang in the balance. After revealing the infirmity of current case law, this Article argues for a doctrinal shift to require assessment of policing technology reliability as part of Fourth Amendment reasonableness determinations and offers a framework that would allow courts to do so. Such a shift may prevent further erosion of privacy rights, particularly for Black, Latine, and other marginalized communities subjected to rampant Fourth Amendment abuses. Recognizing that even a necessary doctrinal shift cannot resolve every concern related to ever-growing police reliance on automated technologies to justify seizures and searches, this Article also goes beyond a

[*] Associate Professor of Law, University of Maryland Francis King Carey School of Law. Thank you to the Bellow Scholar community, Chaz Arnett, Peter Danchin, Gabriel Diaz, Andrew G. Ferguson, David Gray, Richard Gutierrez, Daniel Harawa, Vida B. Johnson, Seema Kakade, Katie Kronick, Amanda Levendowski, Will Moon, Ngozi Okidegbe, Aadhithi Padmanabhan, Michael Pinard, Emily Prokesch, Amanda Pustilnik, Natalie Ram, Nathan Robertson, Vincent M. Southerland, Kim Taylor-Thompson, Tanya Thomas, Anthony C. Thompson, Kate Weisburd, Rebecca Wexler, and participants of the AAPI-MENA Women's Workshop, Decarceration Roundtables, Mid-Atlantic Clinicians' Writing Workshop, ABA-AALS Criminal Justice Section Academic Roundtables, Mid-Atlantic Regional Clinical Conference, Richmond Junior Faculty Forum, and Temple Law Faculty Colloquium Series for helpful conversations and feedback. I am grateful to the research assistants who made the analysis in this Article possible: Paige Bellamy, Maria-Gracia Beltran, Andrea Chaknis, J. Sam Chase, Madalyn Cohee, Erin Ellington, Chase Griffin, Lindsay Hemminger, Faith Kim, Ana Martinez, Cheryl Milevsky, Emilee Shannon, and Marina Sullivan. Thanks also to Susan McCarty, Managing Research Fellow at Maryland Carey Law. Finally, I thank the editors of the *Emory Law Journal*, especially Akarshna Premanand and Celia Golod, for their improvements to this Article. All errors are my own.

focus on doctrine to recommend targeted policy interventions where Fourth Amendment intrusions do not result in criminal litigation.

TABLE OF CONTENTS

INTRODUCTION

In 2022, police in Louisiana were investigating a series of designer purse thefts from across the state.[1] In search of a lead, they turned to facial recognition software hoping to identify a suspect.[2] After police uploaded surveillance footage, the software program spit out a match: Randal Reid.[3] A judge signed an arrest warrant, and police officers arrested Reid in Georgia, three states away from Louisiana.[4]

The software was wrong.[5] Reid had never been to Louisiana.[6] Yet, he faced serious charges and spent almost a week in jail before authorities released him.[7] If there were not significant physical differences between him and the suspect, he may have faced conviction and spent much longer in prison.[8]

Around 2020, a man was walking alongside a building in Chicago just as police received an alert from ShotSpotter,[9] an automated gunshot detection system that purports to detect and locate gunfire in near-real time.[10] The alert notified them that a shot may have been fired from the side of the same building.[11]

The software may have been wrong in this instance too. The man was not acting suspiciously; his proximity to the purported location of the gunfire served as the only link to a possible crime.[12] Police officers conducted an investigatory stop anyway.[13] The officers observed what they would later describe as "bulges"

[1] John Simerman, *JSPO Used Facial Recognition Technology to Arrest a Man. The Tech Was Wrong*, NEW ORLEANS ADVOC. (Jan. 2, 2023), https://www.nola.com/news/crime_police/jpso-used-facial-recognition-to-arrest-a-man-it-was-wrong/article_0818361a-8886-11ed-8119-93b98ecccc8d.html.

[2] Thomas Germain, *Innocent Black Man Jailed After Facial Recognition Got It Wrong, His Lawyer Says*, GIZMODO (Jan. 3, 2023), https://gizmodo.com/facial-recognition-randall-reid-black-man-error-jail-1849944231.

[3] *See id.*

[4] Simerman, *supra* note 1.

[5] Germain, *supra* note 2.

[6] *See* Simerman, *supra* note 1.

[7] *Id.*

[8] *Id.*

[9] *See* CITY OF CHI.: OFFICE OF THE INSPECTOR GEN., THE CHICAGO POLICE DEPARTMENT'S USE OF SHOTSPOTTER TECHNOLOGY 18 (Aug. 2021) [hereinafter CHICAGO OIG REPORT], https://igchicago.org/wp-content/uploads/2021/08/Chicago-Police-Departments-Use-of-ShotSpotter-Technology.pdf.

[10] *See Save Lives and Find Critical Evidence with Proven Gunshot Detection*, SOUNDTHINKING, https://www.soundthinking.com/law-enforcement/leading-gunshot-detection-system/ (last visited Jan. 14, 2024).

[11] CHICAGO OIG REPORT, *supra* note 9, at 18.

[12] *See id.*

[13] *Id.*

in the man's pockets and conducted a frisk.[14] They found drugs and paraphernalia—but nothing connected to a shooting—and arrested the man.[15]

The same year, sixteen-year-old Bobby Jones had just moved with his family to a nice neighborhood in Pasco County, Florida.[16] After being expelled from school in another county for smoking marijuana and fighting, Bobby was looking for a fresh start.[17] Bobby and his family did not know that the local police department had deployed what it described as an "intelligence-led" software program that police claimed could predict which kids in the neighborhood were likely to "fall into a life of crime."[18] The program used factors like school disciplinary history, grades, and past abuse to categorize kids' likelihood of engaging in criminal conduct.[19]

Unbeknownst to Bobby and his parents, Bobby had made the list.[20] Not long after Bobby and his family settled into their new home, police officers showed up unannounced, without a warrant, and without suspicion that Bobby was involved in any crime.[21] Officers searched their home and found empty baggies police claimed contained trace amounts of marijuana.[22] Bobby had been attending his new school for barely a week when police arrested him on drug charges.[23] Florida authorities held him in juvenile detention for three weeks before a judge dismissed all charges against him.[24]

[14] *Id.*

[15] *Id.*

[16] Olivia Solon & Cyrus Farivar, *Predictive Policing Strategies for Children Face Pushback*, NBC NEWS (June 6, 2021), https://www.nbcnews.com/tech/tech-news/predictive-policing-strategies-children-face-pushback-n1269674.

[17] *Id.*

[18] *Id.*; Neil Bedi & Kathleen McGrory, *Pasco's Sheriff Uses Grades and Abuse Histories to Label Schoolchildren Potential Criminals. The Kids and Their Parents Don't Know*, TAMPA BAY TIMES (Nov. 19, 2020), https://projects.tampabay.com/projects/2020/investigations/police-pasco-sheriff-targeted/school-data/ (quoting PASCO SHERIFF'S OFFICE, INTELLIGENCE-LED POLICING MANUAL (2018), https://www.documentcloud.org/documents/20412738-ilp_manual012918).

[19] Solon & Farivar, *supra* note 16; Bedi & McGrory, *supra* note 18.

[20] Solon & Farivar, *supra* note 16.

[21] *See id.*

[22] The Pasco County Sheriff's Office says Bobby's father consented to the search. He disputes this. Solon & Farivar, *supra* note 16. While police alleged that the substance in the baggies field tested positive for marijuana, available information does not indicate whether laboratory testing confirmed this. *See id.* Field tests, typically used at a crime scene before laboratory tests can be conducted, cannot confirm the presence of marijuana, and can result in false positive identifications. Jenny Roberts, *The Innocence Movement and Misdemeanors*, 98 B.U. L. REV. 779, 795 (2018).

[23] Solon & Farivar, *supra* note 16.

[24] *Id.*

Once again, police were wrong: the baggies contained no measurable amount of marijuana.[25]

***

These cases are not anomalies. Police reliance on technologies like those described above bears responsibility for frequent wrongful arrests and prosecutions.[26] The Fourth Amendment permits police officers to conduct seizures or searches of citizens under certain circumstances if supported by reasonable suspicion[27] or probable cause.[28] But, in order to justify an intrusion, police officers must, under either standard, be able to point to specific articulable facts that establish particularized suspicion both that (1) criminal activity is occurring (or has occurred or is imminent)[29] and (2) the person subjected to the intrusion is responsible.[30] Traditionally, police officers have developed suspicion through conventional forms of investigation, including their own direct observation or as a result of third-party information they can corroborate.[31]

[25] *Id.*

[26] *See, e.g.*, T.J. Benedict, *The Computer Got It Wrong: Facial Recognition Technology and Establishing Probable Cause to Arrest*, 79 WASH. & LEE L. REV. 849, 851 (2022) (discussing three wrongful arrests resulting from misidentification errors by facial recognition technology ("FRT") and noting that there "are likely many more unreported wrongful arrests, given FRT's prominence in American policing"); David Gray, *Bertillonage in an Age of Surveillance: Fourth Amendment Regulation of Facial Recognition Technologies*, 24 SMU SCI. & TECH. L. REV. 3, 4 (2021) (describing misidentifications by facial recognition software).

[27] Terry v. Ohio, 392 U.S. 1, 27 (1968).

[28] U.S. CONST. amend. IV.

[29] *See* United States v. Hensley, 469 U.S. 221, 227 (1985) (extending rule permitting stops based on reasonable suspicion for crimes occurring or about to occur to completed crimes); Brinegar v. United States, 338 U.S. 160, 175–76 (1949) ("Probable cause exists where 'the facts and circumstances within [officers'] knowledge and of which they had reasonably trustworthy information [are] sufficient in themselves to warrant a man of reasonable caution in the belief that' an offense has been or is being committed" (quoting Carroll v. United States, 267 U.S. 132, 162 (1925))).

[30] United States v. Cortez, 449 U.S. 411, 417–18 (1981) (finding that "detaining officers must have a particularized and objective basis for suspecting the particular person stopped of criminal activity"); *see also* Beck v. Ohio, 379 U.S. 89, 96–97 (1964) (holding that police did not have probable cause to arrest a suspect merely on the basis of his past criminal record and without a specific showing of reason to believe the suspect had been engaged in criminal behavior).

[31] *See* ANDREW GUTHRIE FERGUSON, THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT 54 (2017) (explaining that police traditionally attempted to discover crime by "patroll[ing] the streets looking for criminal activities"); WAYNE R. LAFAVE, SEARCH AND SEIZURE: A TREATISE ON THE FOURTH AMENDMENT § 9.5(i) (Thompson West, 6th ed. 2021) (summarizing information commonly used by police to justify a search or seizure as including information provided by third-parties and information acquired "first-hand").

Increasingly, however, police outsource critical decisions about whether a crime is occurring and who is responsible to technological tools.[32] Predictive analytics, like crime-mapping and social media analysis software, claim to tell police where crime will happen, when it will happen, what crime will occur, and who will be involved.[33] Sensory enhancing technologies, like automated gunshot detection systems, purport to enable police to identify criminal activity and perpetrators that humans cannot see or hear.[34] Biometric technologies, such as voice and facial recognition software, purport to enable police to identify suspects by analyzing human characteristics like facial features or fingerprints.[35] Such technology dictates whom police officers stop-and-frisk, search, and arrest.[36] In other words, policing technologies *automate* suspicion.[37]

This is precisely what happened in each of the three cases described in the opening vignettes. In each case, police officers failed to make one or both of the judgments required to justify a search or seizure using conventional policing techniques. Police did not interview Randal Reid, place him under observation, or speak to witnesses about his whereabouts during the purse thefts.[38] Police in Chicago had no reason to think the man walking alongside the building had committed any crime before receiving the ShotSpotter alert.[39] Nobody had seen or reported him acting suspiciously and police did not observe him engaging in

[32] *See infra* Part I.

[33] *See* FERGUSON, *supra* note 31, at 35 (“Person-based predictive policing involves the use of data to identify and investigate potential suspects or victims . . . .”); Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 U. PA. L. REV. 871, 873 (2016) (“[S]oon a computer may spit out a person’s name, address, and social security number along with the probability that the person is engaged in a certain criminal activity, with no further explanation.”); Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*, 10 HARV. L. & POL’Y REV. 15, 16 (2016) [hereinafter Joh, *New Surveillance Discretion*] (explaining that law enforcement is “experimenting with predictive policing software to identify geographic places where crime is likely to take place”).

[34] *See, e.g.*, Maneka Sinha, *The Dangers of Automated Gunshot Detection*, 5 U. PA. J.L. & INNOV. 63, 67 (2023) [hereinafter Sinha, *Automated Gunshot Detection*] (describing how automated gunshot detection system company ShotSpotter claims that its software allows police officers to respond to gunfire that would not have been detected by individual reports). The company has since changed its name to SoundThinking. *Shotspotter Changes Corporate Name To SoundThinking And Launches Safetysmart Platform For Safer Neighborhoods*, SOUNDTHINKING, https://www.soundthinking.com/press-releases/shotspotter-changes-corporate-name-to-soundthinking-and-launches-safetysmart-platform-for-safer-neighborhoods/ (last visited Jan. 14, 2024).

[35] *See, e.g.*, Gray, *supra* note 26, at 12–13.

[36] *See infra* Part I.

[37] *See infra* Part I.

[38] *See supra* notes 2–3 and accompanying text; Joh, *New Surveillance Discretion*, *supra* note 33, at 15 (explaining that traditional police investigation includes “observation, questioning, and information conveyed by witnesses, victims, or other third parties”).

[39] *See* CHICAGO OIG REPORT, *supra* note 9, at 18.

criminal conduct.[40] And police in Florida had no reason at all to think that Bobby Jones (or anyone) had done anything criminal; they had not observed suspicious activity at Bobby's home, nor had they received any information suggesting that crime was occurring there.[41]

In each case, a policing technology made the threshold judgments necessary for a search or seizure by automating criminal suspicion.[42] Facial recognition software pinpointed Randal Reid.[43] When authorities issued a warrant, no one seriously questioned whether the program had identified the right person or attempted to verify its accuracy.[44] ShotSpotter gave Chicago police suspicion that a gun crime may have occurred, and by directing officers to a specific location, made anyone present in the vicinity a potential suspect.[45] Police did not question ShotSpotter's accuracy even though they found no evidence of a shooting.[46] Instead, they used it to stop a man with no apparent connection to the ShotSpotter alert.[47] Police only went to Bobby Jones's home because their computer program categorized Bobby as a someone who might commit an undefined crime at an undetermined time in the future.[48]

Fourth Amendment doctrine has not kept pace with such technology-driven policing. A hallmark of Fourth Amendment law, as pronounced consistently by the Supreme Court, is that the information provided to or observed by law enforcement to justify a search or seizure must be reliable.[49] The reliability requirement serves as a buffer against unreasonable searches or seizures; wholly

[40] *See id.*

[41] *See* Solon & Farivar, *supra* note 16.

[42] Policing technologies refer here to any hardware, software, or combination tools used to support, assist, or enhance traditional human police functions including predicting, discovering, identifying, or stopping crime and perpetrators of crime. Policing technologies represent only one subset of a broader array of carceral technologies used throughout criminal and criminal-adjacent processes. Carceral technologies include technologies used in prisons, for surveillance, and at the border, in addition to technologies used to police. *See Why We Build, Work, & Fight for Community Defense Against Carceral Systems & Their Tech*, CARCERAL TECH RESISTANCE NETWORK (Mar. 30, 2020), https://www.carceral.tech/why ("[C]arceral technologies are tech that are bound up in the control, coercion, capture, and exile of entire categories of people.").

[43] Simerman, *supra* note 1.

[44] *See id.* (explaining that the detective "took the algorithm at face value" when securing the warrant for Randal Reid's arrest); Germain, *supra* note 2.

[45] CHICAGO OIG REPORT, *supra* note 9, at 18.

[46] *Cf. id.* (explaining that police conducted an investigatory stop and frisk despite finding no evidence that a shooting occurred).

[47] *See id.*

[48] Solon & Farivar, *supra* note 16.

[49] Illinois v. Gates, 462 U.S. 213, 230 (1983); Alabama v. White, 496 U.S. 325, 328–29 (1990); Florida v. J.L., 529 U.S. 266, 270–72 (2000); Navarette v. California, 572 U.S. 393, 397 (2014); *cf.* Florida v. Harris, 568 U.S. 237, 246–47 (2013).

unreliable facts cannot establish either reasonable suspicion or probable cause.[50] In turn, the Court has outlined frameworks for assessing the reliability of some information used to support reasonable suspicion and probable cause.[51]

The Supreme Court has not announced how courts should address the reliability of policing technologies that officers increasingly rely upon to support a search or seizure. As a result, courts routinely uphold searches and seizures driven by technologies like those described above.[52]

The flaws in Randal Reid's and Bobby Jones's cases were detected early in their prosecutions.[53] Others will not be as lucky. Many individuals subjected to Fourth Amendment intrusions justified by policing technology will be prosecuted and will have the opportunity to challenge the legality of their searches and seizures. How do courts assess reliability when police rely on technology to justify searches and seizures? This Article explores these cases and begins to answer that question.

This Article examines how reliance on policing technologies confounds Fourth Amendment jurisprudence and exposes the consequences of courts' failure to address reliability of such technologies meaningfully in reasonableness determinations. Through an empirical analysis of numerous state and federal opinions where courts addressed the reliability of policing technologies used to justify a search or seizure or were asked to do so, this Article builds on and extends the rich literature examining the problems that flow from police reliance on technological assistance.

Much of the post-*Carpenter v. United States*[54] Fourth Amendment and technology scholarship examines how *Carpenter* influences how a "search" is

[50] *E.g.*, *J.L.*, 529 U.S. at 272 ("[R]easonable suspicion . . . requires that a tip be reliable in its assertion of illegality . . . .").

[51] *E.g.*, *White*, 496 U.S. at 332 (holding that information provided through an anonymous tip will be considered reliable if corroborated); *Harris*, 568 U.S. at 246–47 (finding that a drug dog alert is sufficiently reliable to establish probable cause if the dog satisfactorily completed a certification or training program); *see also Gates*, 462 U.S. at 230–31 (explaining that probable cause is assessed under a "totality-of-the-circumstances" test which may be "illuminate[d]" by an "informant's 'veracity,' 'reliability,' and 'basis of knowledge'").

[52] *See, e.g.*, Wisconsin v. Nimmer, 975 N.W.2d 598, 599–600 (Wis. 2022); United States v. Rickmon, 952 F.3d 876, 878 (7th Cir. 2020).

[53] *See* Simerman, *supra* note 1 (explaining that Reid was released from DeKalb County jail and his warrant was rescinded after seven days in detention); Solon & Farivar, *supra* note 16 (explaining that Bobby Jones's charges were dropped). Because the details of the Chicago man's stop and arrest were reported anonymously, the outcome of his case is unknown. *See* CHICAGO OIG REPORT, *supra* note 9, at 18.

[54] 138 S. Ct. 2206 (2018).

defined.[55] This Article, however, is among the first to investigate a different, underexplored intersection of Fourth Amendment procedure and policing technology. Instead of focusing on what is and is not a search, it explores how courts do—and should—analyze reliability of information used to establish reasonable suspicion or probable cause when policing technology is the source of such information. In doing so, it shifts focus away from *Carpenter*-driven questions about what constitutes a search onto how policing technology complicates reasonableness analysis and how doctrine and policy can adapt in response.

The empirical analysis provides a snapshot of how some courts analyze these issues. It reveals that courts frequently fail to subject policing technology to the same sort of critical examination applied to traditional information, such as first-hand police observations or information provided by third parties. Courts often simply decline to assess reliability, even when the accused asks them to do so. Despite the Supreme Court's repeated pronouncement that information supporting probable cause or reasonable suspicion needs to be reliable, those courts that do address reliability of policing technologies in assessing the legality of seizures and searches often do so in a cursory manner, without conducting a substantive reliability evaluation.

The Supreme Court's deference to officer judgment and its laxity in preserving constitutional safeguards has caused the original protections of the Fourth Amendment to wane over time.[56] Lower courts' failure to scrutinize policing technologies in Fourth Amendment challenges exacerbates this trend. It sends a clear message that police use of such technology can continue without real scrutiny or consequence. As the three opening vignettes reveal, policing technologies that automate suspicion allow officers to avoid meaningful

[55] For a very small slice of such scholarship, see generally Paul Ohm, *The Many Revolutions of* Carpenter, 32 HARV. J.L. & TECH. 357, 360 (2019); Matthew Tokson, *The Next Wave of Fourth Amendment Challenges After* Carpenter, 59 WASHBURN L.J. 1, 18 (2020); Andrew Guthrie Ferguson, *Persistent Surveillance*, 74 ALA. L. REV. 1, 23 (2022); and Matthew Tokson, *The* Carpenter *Test as a Transformation of Fourth Amendment Law*, 2023 U. ILL. L. REV. 507, 508 (2023).

[56] Anna Lvovsky, *Rethinking Police Expertise*, 131 YALE L.J. 475, 488 (2021); Eric J. Miller, *Detective Fiction: Race, Authority, and the Fourth Amendment*, 44 ARIZ. ST. L.J. 213, 223–24 (2012) ("[T]he Court . . . removes both judicial and public scrutiny through deference to some inarticulable police 'sixth sense' about crime."); Julian A. Cook, III, *Suspicionless Policing*, 89 GEO. WASH. L. REV. 1568, 1574–75 (2021) (arguing that Supreme Court precedent "has communicated to police departments from coast to coast that they enjoy vast investigative authority, enormous discretion, and will often suffer little in terms of consequence when constitutional safeguards are violated"); *see also* David A. Harris, *Frisking Every Suspect: The Withering of* Terry, 28 U.C. DAVIS L. REV. 1, 4–6 (1994) (explaining that "lower courts have stretched the law governing frisks").

assessments of whether particularized suspicion exists. The consequences are highest when the technology that propels the Fourth Amendment intrusion is inaccurate: police reliance can result in unwarranted stops, searches, arrests, detentions, prosecutions, and incarcerations.

Under the guise of seeming neutrality, policing technologies are widely used to worsen the over-policing and hyper-surveillance of communities that have long been targets of both.[57] By embedding the biases they are built on, machines themselves exhibit bias.[58] Technology covers more ground, is not limited by human capacity, can process exponentially more information than police officers, and can accomplish all of this faster than humans.[59] These features create more opportunities for harmful interactions between police and civilians, exacerbating feelings of distrust, alienation, and being targeted.[60]

Evidence suppression hearings—hearings held prior to a criminal trial at which judges determine the legality of searches or seizures that resulted in the discovery of evidence against the accused[61]—are often the only phase of a criminal prosecution at which the reliability of such technologies can be examined. Although an established, if not always effective,[62] framework for assessing reliability of technological evidence exists at the trial stage,[63] the

[57] *See, e.g.*, WENDY LEE, JUMANA MUSA & MICHAEL PINARD, GARBAGE IN, GOSPEL OUT 22 (2021), https://www.nacdl.org/getattachment/eb6a04b2-4887-4a46-a708-dbdaade82125/garbage-in-gospel-out-how-data-driven-policing-technologies-entrench-historic-racism-and-tech-wash-bias-in-the-criminal-legal-system-11162021.pdf (describing how policing technologies further disparities, but "tech-wash" these effects with an appearance of objectivity); Sinha, *Automated Gunshot Detection*, *supra* note 34, at 107.

[58] *See* Itiel E. Dror, *Cognitive and Human Factors in Expert Decision Making: Six Fallacies and the Eight Sources of Bias*, 92 ANALYTICAL CHEMISTRY 7998, 7999 (2020) (describing how technology developed by humans manifests bias and can introduce new biases); Eldar Haber, *Racial Recognition*, 43 CARDOZO L. REV. 71, 90–91 (2021) (explaining that machines replicate the biases of their developers and datasets).

[59] *See* Christopher Slobogin, *Panvasive Surveillance, Political Process Theory, and the Nondelegation Doctrine*, 102 GEO. L.J. 1721, 1723 (2014) (describing government and law enforcement mass surveillance tools that "routinely and randomly reach across huge numbers of people, most of whom are innocent of any wrongdoing" as "panvasive").

[60] *See* Tracey L. Meares, *Programming Errors: Understanding the Constitutionality of Stop-and-Frisk as a Program, Not an Incident*, 82 U. CHI. L. REV. 159, 164–65 (2015) (explaining how, although the constitutionality of a stop-and-frisk is assessed on an incident-by-incident basis, many of those stopped, who are overwhelmingly young men of color, experience the stops as a "program to police them as a group"); LEE ET AL., *supra* note 57, at 62 ("These tools broaden the net that hyper-criminalizes poor communities of color . . . .").

[61] *See* LAFAVE, *supra* note 31, § 3.1(d) (summarizing the purpose of suppression hearings).

[62] *See generally* Maneka Sinha, *Junk Science at Sentencing*, 89 GEO. WASH. L. REV. 52 (2021) [hereinafter Sinha, *Junk Science*]; Maneka Sinha, *Radically Reimagining Forensic Evidence*, 73 ALA. L. REV. 879 (2022) [hereinafter Sinha, *Radically Reimagining*].

[63] *See* FED. R. EVID. 702; Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 597 (1993); Kumho Tire Co. v. Carmichael, 526 U.S. 137, 158 (1999).

overwhelming majority of cases never proceed to a trial.[64] Even in those that do, by the time a trial begins, the legality of the search or seizure in question has already been decided, and thus evidence relating to police use of technology in searches and seizures need not be presented at the trial itself.[65] Meaning, if judges do not evaluate policing technology reliability at suppression hearings, it may go entirely untested through the life of a case.[66]

To examine these problems, this Article unfolds in four Parts. Parts I and II provide background for understanding how the Supreme Court's Fourth Amendment reliability jurisprudence is ill-suited for assessing the reliability of policing technologies increasingly used to justify searches and seizures. Part I describes commonly used policing technologies that automate criminal suspicion, how they automate criminal suspicion, and what is currently known about their reliability. Part II begins by setting out how the Supreme Court has thus far addressed reliability in deciding Fourth Amendment questions and how these approaches are inadequate for evaluating policing technology reliability. Part II then compares the Court's Fourth Amendment reliability frameworks to how reliability of technological evidence is assessed at the trial stage.

Through an empirical analysis of numerous lower court cases in which courts addressed or were confronted with the issue, Part III examines the need for a doctrinal shift toward a clear framework for addressing reliability of policing technologies in Fourth Amendment determinations. The results of the analysis demonstrate that lacking a clear directive for how to do so, courts employ one or a combination of several commonly used, but insufficient, approaches for addressing reliability of technology in evaluating the lawfulness of searches and seizures.

Part IV examines the implications of these results, beginning with doctrinal implications. It first reveals that when asked to address the reliability of policing technology used to justify a search or seizure, courts frequently fail to comply with current Fourth Amendment law. It also exposes a significant gap in criminal procedure doctrine: even when courts do attempt to address reliability of policing technology, the existing frameworks for doing so prove insufficient when applied to technology.

[64] Beth Schwartzapfel, Abbie VanSickle & Annaliese Griffin, *The Truth About Trials*, MARSHALL PROJECT (Nov. 4, 2020), https://www.themarshallproject.org/2020/11/04/the-truth-about-trials.

[65] *See infra* Part II.

[66] *See infra* Part II.

Part IV then proposes a new normative framework for addressing reliability of technology in Fourth Amendment reasonableness determinations. Under current doctrine, while judges are the arbiters of whether reliable information exists to support a search or seizure, courts often find information developed by officers or their agents sufficient. Officers are unequipped, however, to assess the reliability of technology, yet have an incentive to convince judges that the tools they rely on are trustworthy. For these reasons, this Article recommends that courts assess policing technology in reasonableness determinations by evaluating external, independent, and disinterested evidence of reliability. "External" refers here to evidence separate from the outputs of the technology itself. "Independent" evidence is that developed by actors who are not parties to the case or their affiliates, such as law enforcement, prosecution, or defense entities. "Disinterested" connotes neutrality and lack of bias in favor of or against a party or its affiliates.[67] Studies describing a technology's reliability produced by researchers unaffiliated with the case or interested parties might constitute such external, independent, and disinterested evidence. Where no such evidence exists, this proposal recommends that courts utilize the reliability test applicable at the trial stage to evaluate policing technology reliability.

Courts could implement this procedure in several ways. They could ask the parties to supply such evidence and determine if it meets the described criteria. Courts could also look for such evidence on their own and give the parties an opportunity to bolster or rebut it through arguments or by providing their own evidence.

Because only a slim minority of Fourth Amendment events that occur on the street result in litigation, Part IV also addresses the policy implications that flow from this analysis. It examines how police reliance on technology to justify searches and seizures and courts' failure to question such reliance exacerbate already pronounced erosion of Fourth Amendment protections. Recognizing that even a necessary doctrinal shift cannot resolve every concern related to the ever-increasing reliance on policing technologies, Part IV suggests additional targeted policy interventions that may alleviate harms that result from use of policing technologies even where Fourth Amendment intrusions do not result in criminal litigation.

[67] *See* Sinha, *Radically Reimagining*, *supra* note 62, at 893 (explaining that actors perceived as neutral may take actions favoring one party or another).

I. AUTOMATED SUSPICION

Understanding the implications of judicial failure to evaluate reliability of policing technologies in Fourth Amendment reasonableness determinations requires some context. Using illustrative examples, this Part offers a definition of automated suspicion distinct from those provided by other scholars to explain how policing technology can automate suspicion of crime and of persons. Additionally, it considers the reliability of commonly used policing technology.

A. Conceptualizing Automated Suspicion

Automation generally refers to the systematization of a process by application of a technology that eliminates or minimizes human input.[68] As used here, policing technology that automates suspicion has three key features. First, policing technologies automate suspicion when they replace or usurp, entirely or near-entirely, the traditional role of law enforcement agents in developing criminal suspicion.[69] They *automate* the suspicion required to conduct a search or seizure because the officer relies entirely or near-entirely on the technology's output to decide whom to search or seize.

Second, technology replaces or usurps the traditional officer in a way that cannot be meaningfully checked by human police officers.[70] This might be because humans cannot fully understand or replicate a technology's process. Such technologies are black boxes. A human officer relying on such a technology does not understand how the technology reaches its conclusions, cannot employ the same method as the technology to check its work,[71] and would not be qualified to do so even if the means were available. Reciprocally, because how such technology works is not readily knowable, an accused person who was subjected to a police intrusion based on the output of an automated suspicion policing technology cannot check its work either.[72]

This black box nature of automated suspicion policing technology is important to emphasize because there are other policing technologies that assist officers to develop suspicion that police can understand and can vet

[68] *Automation*, MERRIAM-WEBSTER, https://www.merriam-webster.com/dictionary/automation (last visited Jan. 14, 2024).

[69] *See infra* Part I.B.1.

[70] *See infra* Part I.B.1.

[71] Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 6 (2014) ("[A black box] converts inputs to outputs without revealing how it does so.").

[72] *See id.*

independently,[73] meaning the technology alone does not necessarily dictate whom or which crimes police decide to target. Law enforcement investigation and judgment may still be substantially required to develop a suspect or solve a crime.[74]

Alternatively, the technology's method is understood and could potentially be replicated by human analysis, but doing so would require exponentially more time and resources to perform manually. Database searches, which electronically pore through massive volumes of data for some target information (say, whether an individual has an open arrest warrant), are an example.

Third, humans can play a role in automated processes.[75] Typically, humans cannot employ the same analytical process as the machine and do not base their decisions on the same facts and features as the machine. Humans can attempt to check a machine's results, if not its process. Such human checks are usually superficial.

Facial recognition software is helpful to elucidate these three features. Consider the scenario in which a target photograph of a potential suspect is scanned by a facial recognition software like in Randal Reid's case. The software generates a suspect by comparing the target photograph to an existing database of photographs, without any involvement (besides running the software) of a human law enforcement agent. It does so in a way that humans may understand in an abstract way but cannot replicate. Algorithms perform an "automated step" of generating a "mathematical representation" of the uploaded image to compare to templates in the database.[76] Law enforcement agents do not know how the software chooses features to compare or how it conducts the comparison, and humans cannot reproduce the comparative method.[77]

Finally, while an officer will typically play a role in the process by "checking" that the facial recognition-generated suspect actually appears to be the person in the target photograph, the check is only surface-level. A human

[73] *See* Brandon L. Garrett & Cynthia Rudin, *The Right to a Glass Box: Rethinking the Use of Artificial Intelligence in Criminal* Justice, 101 CORNELL L. REV. (forthcoming 2024) (manuscript at 21–22) (defining interpretable artificial intelligence models as those that utilize decision-making that humans can understand); *id.* at 24–25 (describing some risk assessment tools as an example of a "glass box," or interpretable technology).

[74] *Cf. id.* at 31–32 (explaining how when decision-makers understand a model's processes, they can evaluate its output and make independent judgments based on its decisions).

[75] *See infra* Part I.B.1.

[76] CLARE GARVIE, A FORENSIC WITHOUT THE SCIENCE: FACE RECOGNITION IN U.S. CRIMINAL INVESTIGATIONS 11 (Geo. L. Ctr. on Priv. & Tech., 2022).

[77] *See id.* at 12, 22.

"check" will not employ the same process as the software and will not render a conclusion using the same information as the software. Common sense clarifies the truth of this: in Reid's case, in which the software's decision led to a wrongful arrest, police officers did attempt to "check" the software's work by comparing the photographs at issue.[78] But the check did not reveal the program's error; if it had, the wrong person may not have been arrested.[79]

Confirmation bias, or the phenomenon through which people interpret information in ways that confirm existing beliefs, is further reason to understand that so-called human "checks" do little to de-automate the conclusions of automated suspicion technologies.[80] Knowing the software's determination of the target suspect's identity can bias a human officer tasked with checking the accuracy of the software's conclusion.[81]

Automated Biometric Identifications Systems (ABIS) or Automated Fingerprint Identifications Systems (AFIS) are similarly clarifying. ABIS and AFIS databases house massive quantities of biometric data including millions of digital fingerprints.[82] When a print from an unknown source is recovered, it can be uploaded to an ABIS/AFIS database, which will analyze features of the print and search for possible matches.[83] The database then produces a list of potential suspects.[84] Typically a fingerprint examiner will then independently compare the features of the evidentiary print to those of the candidate prints, although a law enforcement agent could simply treat the listed candidates as suspects without an examiner's evaluation.[85] Even when a manual analysis and

[78] Simerman, *supra* note 1.

[79] *See id.*

[80] *See* Saul M. Kassin, Itiel E. Dror & Jeff Kukucka, *The Forensic Confirmation Bias: Problems, Perspectives, and Proposed Solutions*, 2 J. APPLIED RSCH. IN MEMORY & COGNITION 42, 44–48 (2013) (describing varieties of and summarizing research on confirmation bias).

[81] *See id.* at 44.

[82] Aaron Boyd, *Pentagon Will Move Primary Biometrics Systems to Amazon Cloud*, NEXTGOV (Oct. 20, 2020), https://www.nextgov.com/it-modernization/2020/10/pentagon-will-move-primary-biometrics-systems-amazon-cloud/169392/; Erin Murphy, *Databases, Doctrine & Constitutional Criminal Procedure*, 37 FORDHAM URB. L.J. 803, 808 (2010).

[83] Meghan J. Ryan, *Escaping the Fingerprint Crisis: A Blueprint for Essential Research*, 2020 U. ILL. L. REV. 763, 770 (2020). Sometimes a human examiner will evaluate a print to determine if it is "of value"—i.e., if it is of sufficient quality for entry into a database and for analysis and comparison to another print before a fingerprint left at a crime scene can be uploaded for a database search. *See* Heidi Eldridge, Marco DeDonno, Julien Furrer & Christophe Champod, *Examining and Expanding the Friction Ridge Value Decision*, 314 FORENSIC SCI. INT'L 1, 1–2 (2020) (explaining that, as a first step, fingerprint analysts make determinations of whether a latent print is suitable, and that suitability encompasses suitability for comparison and database entry).

[84] Ryan, *supra* note 83, at 770.

[85] *See* LATENT PRINT AFIS INTEROPERABILITY WORKING GRP., NAT'L INST. OF STANDARDS & TECH., DRAFT GLOSSARY OF AFIS TERMS 5 (final glossary published May 1, 2012) (describing the process of

comparison of features is conducted by a human examiner, the examiner cannot employ the same method used by the software.[86] The examiner may not even compare the same features as the software[87] and the examiner does not know the software's algorithm or how that algorithm conducts its comparison. Nor can the examiner compare the evidentiary print to *all* of the prints in the database; they only compare it to some subset of prints on the candidate list.

The definition used here builds on and challenges prior scholarly conceptions of automated suspicion. Scholars have previously used intersecting and overlapping definitions to explain how policing technologies automate suspicion. Professor Elizabeth Joh has focused on how police can use alerts produced by "big data tools," which apply computer analytics to extremely large sets of digitized data to "identify suspicious persons and activities on a massive scale."[88] The late Professor Michael Rich coined the term "automated suspicion algorithm," or "ASA," to describe machine learning programs used to predict individual criminality.[89] Rich defined ASAs as algorithms that attempt to detect patterns in data to automatically predict individual criminality.[90]

Policing technologies can automate suspicion even if they do not fall within the parameters outlined by Joh and Rich. Technology need not be based on "big data tools" like machine learning processes, which automatically improve their performance over time with experience,[91] to automate suspicion. Typically, machine learning technologies are trained to perform a certain function using a large data set.[92] The program improves its performance the more times it

identifying latent fingerprints, which includes comparison of the latent print with prints in the AFIS database by a human examiner, followed by an evaluation to determine whether the impressions were made by the same source); *see also* BUREAU OF JUST. ASSISTANCE, U.S. DEP'T OF JUST., PLANNING FOR AUTOMATED FINGERPRINT IDENTIFICATION SYSTEMS (AFIS) IMPLEMENTATION 5 (June 1988), https://www.ojp.gov/pdffiles1/Digitization/115419NCJRS.pdf (explaining that when analyzing a latent fingerprint using the AFIS database, after a list of suspects has been generated, the investigating agency will "compare the fingerprint cards of the suspects with the latent crime scene print").

[86] H. Swofford, C. Champod, A. Koertner, H. Eldridge & M. Salyards, *A Method for Measuring the Quality of Friction Skin Impression Evidence: Method Development and Validation*, 320 FORENSIC SCI. INT'L 1, 3 (2021).

[87] *Id.* (noting that AFIS programs frequently do not use the same features that human analysts consider in fingerprint comparisons).

[88] Joh, *New Surveillance Discretion*, *supra* note 33, at 16, 16 n.7.

[89] Rich, *supra* note 33, at 876.

[90] *Id.*

[91] Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 89 (2014).

[92] *See* Rich, *supra* note 33, at 880.

completes the task.[93] Probabilistic genotyping software (PGS) systems—sophisticated programs that leverage mathematical and biological models to conduct complex DNA mixture analysis and statistical calculations that humans cannot perform—are not all machine learning systems.[94] Yet, by pointing police to a suspect, and doing so in a way that cannot be checked or replicated even by trained human DNA analysts due to the complexity of the process, they nevertheless automate suspicion of a suspect.

Policing technologies also need not be predictive to automate suspicion.[95] Technology that seeks to identify perpetrators of past crimes, like facial recognition technologies used in cases like Randal Reid's, also automate suspicion.

Finally, as used here, technology that automates suspicion does not focus only on the individual.[96] A policing technology may automate suspicion even if it does not generate suspicion of *who* may have committed a crime if it generates suspicion of *what* crime or type of crime has occurred or may occur. In other words, policing technologies can automate suspicion of a crime, a person, or both.

B. How Policing Technologies Automate Criminal Suspicion

Policing technologies can generate suspicion of a crime, of a person responsible for a crime, or of both.[97] How policing technologies perform these

[93] *Id.* For a more comprehensive description of machine learning, see generally David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653, 655 (2017) (clarifying that machine learning consists of several distinct steps).

[94] The primary developer of STRmix, the leading PGS system used worldwide, has stated that the system is not based on machine learning. Letter from John Buckleton, Developer, STRmix, to Jill Presser, Justice, Superior Ct. of Ont., & Kate Robertson, Rsch. Fellow, Univ. of Toronto (Aug. 9, 2021), https://johnbuckleton.files.wordpress.com/2021/08/ai-case-study-ii.pdf.

[95] *See supra* note 34–35 and accompanying text.

[96] *See infra* Part I.B.1.

[97] Policing technologies can express conclusions in categorical or probabilistic terms. Brandon Garrett, Gregory Mitchell & Nicholas Scurich, *Comparing Categorical and Probabilistic Fingerprint Evidence*, 63 J. FORENSIC SCI. 1712, 1712 (2018). Categorical conclusions either identify a specific crime or suspect, or do not, whereas probabilistic conclusions predict the likelihood of crime occurring or an individual's involvement in crime. *Cf. id.* (explaining, in the context of fingerprint analysis, that categorical opinions indicate whether prints "do or do not originate from the same source" whereas "probabilistic conclusions . . . estimat[e] the probability of a match"). The distinction is not always meaningful in practice as some technologies appear to do both. PGS systems used to interpret complex DNA profiles, for example, are designed to express conclusions probabilistically. Jeanna Neefe Matthews, Graham Northup, Isabella Grasso, Stephen Lorenz, Marzieh Babaeianjelodar, Hunter Bashaw, Sumona Mondal, Abigail Matthews, Mariama Njie & Jessica Goldthwaite, *When Trusted Black Boxes Don't Agree: Incentivizing Iterative Improvement and Accountability in Critical*

functions is outlined next, along with examples of technologies that fall within each category.

1. *Automating Suspicion of Crime*

A variety of policing technologies either predict the occurrence of a specific crime or indicate that a crime has already occurred. Examples of each are examined next.

Predictive analytics. A growing number of police departments use computer models to predict future crime.[98] Predictive policing tools can generally be broken down into two categories—tools that predict the location crime might occur, and tools that predict who is likely to be involved in a crime.[99]

Tools that predict where crime is likely to occur typically analyze large historical crime data sets, like when and where certain types of crime occurred in the past, to predict where and what types of crime might occur in the future.[100] As Professor Andrew Ferguson has explained, predictive tools analyze "event-based" information, like arrest data or calls for service, as well as "place-based" information, like addresses of known criminal suspects or places where violence is frequent.[101] Analyses of event or place-based data can be further weighted by additional factors ranging from type of crime to unique features of a specific locale.[102] Algorithms analyze patterns in the data to predict crime risk within small geographic areas of just a few hundred feet squared, giving policing

Software Systems, PROCEEDINGS OF THE AAAI/ACM CONF. ON AI, ETHICS & SOC'Y 102, 103 (2020). Nevertheless, PGS conclusions are often inappropriately conveyed in categorical terms. *See* William C. Thompson, *Uncertainty in Probabilistic Genotyping of Low Template DNA: A Case Study Comparing STRMix™ and TrueAllele™*, 68 J. FORENSIC SCI. 1049, 1059 (2023) (criticizing one PGS developer, Cybergenetics, for presenting probabilistic conclusions using the term "match" which connotes a source identification).

[98] *See* Tim Lau, *Predictive Policing Explained,* BRENNAN CTR. FOR JUST. (Apr. 1, 2020), https://www.brennancenter.org/our-work/research-reports/predictive-policing-explained (listing developments in the adoption of predictive technology by police departments).

[99] *See* Elizabeth E. Joh, *Policing by Numbers: Big Data and the Fourth Amendment*, 89 WASH. L. REV. 35, 42–46 (2014) [hereinafter Joh, *Policing by Numbers*] (describing predictive policing programs used to pinpoint locations where crime might occur); Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. 1109, 1137 (2017) (explaining that "new predictive technologies are being created to target individuals predicted to be involved in criminal activity").

[100] *See* Andrew Guthrie Ferguson, *Predictive Policing and Reasonable Suspicion*, 62 EMORY L.J. 259, 266–67 (2012) (explaining how historical crime data is used by police departments to guide decisions about how to best allocate resources to target crime).

[101] *Id.* at 266.

[102] *Id.* at 266–67.

suspicion of criminal activity.[103] The data analysis models used by these algorithms are not replicable—or even understood—by human analysis.[104]

Automated gunshot detection systems. Automated gunshot detection systems like ShotSpotter use a combination of hardware (microphones attached to city structures) and software (computer programs that listen for loud impulsive noises, separate them from other sounds, and classify them as either gunfire or non-gunfire) in an attempt to alert police to the sound and location of gunfire in near real-time.[105] Once a human analyst confirms the software's characterization of a sound as gunfire, an alert containing the believed location of gunshots, along with other information, is pushed to police.[106] Police officers respond to ShotSpotter alerts with the belief that a gun crime has occurred.[107] The alert thus gives police suspicion that gun crime has occurred even when officers have not heard gunshots on their own or have not received reports of gunfire from others.[108]

2. Automating Suspicion of Persons

Examples of how policing technologies also automate suspicion of those responsible for a past crime or those who may commit a crime in the future are discussed next.

Probabilistic genotyping software (PGS). Sophisticated PGS systems are commonly used by forensic DNA laboratories to interpret complex crime scene DNA samples and associate a suspect with a crime scene DNA sample.[109] Crime scene DNA samples are often far too complex for manual interpretation by human analysts using conventional techniques.[110] PGS software, however,

[103] *Id.*

[104] *See id.* at 267 (describing how LAPD's predictive policing system analyzes historical property crime data to identify areas of "probable criminal activity").

[105] *See* Sinha, *Automated Gunshot Detection*, *supra* note 34, at 74–75 (explaining the basics of ShotSpotter gunshot detection technology). Human analysts attempt to confirm the algorithm's classification before an alert is sent to police. *Id.*

[106] *ShotSpotter Frequently Asked Questions*, SHOTSPOTTER (2018), https://www.ShotSpotter.com/system/content-uploads/SST_FAQ_January_2018.pdf.

[107] *Cf.* Sinha, *Automated Gunshot Detection*, *supra* note 34, at 104.

[108] *See id.* at 88 ("Police respond to ShotSpotter alerts primed to believe anyone nearby is—and treat anyone nearby as—a potential armed suspect, making encounters very high stakes.").

[109] Dan E. Krane & M. Katherine Philpott, *Using Laboratory Validation to Identify and Establish Limits to the Reliability of Probabilistic Genotyping Systems*, *in* HANDBOOK OF DNA PROFILING 297, 298 (Hirak Ranjan Dash, Pankaj Srivastava & J.A. Lorente eds., 2022).

[110] NAT'L INST. OF STANDARDS & TECH., DNA MIXTURE INTERPRETATION: A NIST SCIENTIFIC FOUNDATION REVIEW 23, 30–31 (2021), https://nvlpubs.nist.gov/nistpubs/ir/2021/NIST.IR.8351-draft.pdf. This

utilizes mathematical and biological modeling to "untangle" complex mixtures and assign statistical weight to the results, using processes that humans cannot replicate.[111] While PGS software does not say who committed a crime or whose DNA is in a particular profile, it is used to associate a specific suspect with a crime.[112] PGS software, thus, automates suspicion of suspects in cases where DNA evidence is available.

Facial recognition software. Facial recognition software programs automate the process of comparing unknown facial images to images of known persons in databases.[113] Facial recognition software can be used to verify a person's claimed or suspected identity; to identify an unknown person's face, as in Randal Reid's case;[114] to perform short or long-term face surveillance; or for comparison using live or stored video feeds.[115] Face surveillance, in turn, can be used for facial tracking, for example, by searching available video feeds to track a suspect's travel.[116]

In each scenario, software directs police to a suspect using complex algorithms that turn an image into a mathematical template for comparison to other templates within the software database.[117] Even though facial recognition software typically outputs a list of "candidates" the software characterizes as most similar to the suspect face, as described earlier, it automates suspicion by

is because crime scene samples may comprise DNA contributions from multiple sources, contain relatively low quantities of DNA, have high degrees of shared DNA across contributors, and suffer from quality issues. *Id.*

[111] Krane & Philpott, *supra* note 109, at 298; Eli Siems, Katherine J. Strandburg & Nicholas Vincent, *Trade Secrecy and Innovation in Forensic Technology*, 73 HASTINGS L.J. 773, 783 (2022); *see also* Natalie Ram, *Innovating Criminal Justice*, 112 NW. U. L. REV. 659, 675–76 (2018) (explaining that, where manual DNA analysis methods fall short, probabilistic genotyping software can use mathematical modeling to explain the behavior of complicated DNA samples).

[112] It is a common misunderstanding that DNA analysis calculates the probability that a particular person contributed DNA to a sample. It does not. Modern PGS-based DNA analysis produces a statistic, known as a likelihood ratio, that compares the probabilities of two hypotheses given the DNA evidence observed. It does not indicate the probability that a particular individual contributed to a DNA sample. *See* Bess Stiffelman, *No Longer the Gold Standard: Probabilistic Genotyping Is Changing the Nature of DNA Evidence in Criminal Trials*, 24 BERKELEY J. CRIM. L. 110, 118–20 (2019) (defining likelihood ratios).

[113] CLARE GARVIE, ALVARO M. BEDOYA & JONATHAN FRANKLE, GEO. L. CTR. ON PRIV. & TECH., THE PERPETUAL LINE-UP: UNREGULATED POLICE FACE RECOGNITION IN AMERICA 9 (2016), https://www.perpetuallineup.org/ (explaining that "[f]ace recognition is the automated process of comparing two images of faces to determine whether they represent the same individual"); GARVIE, *supra* note 76, at 4.

[114] *See* Germain, *supra* note 2.

[115] Andrew Guthrie Ferguson, *Facial Recognition and the Fourth Amendment*, 105 MINN. L. REV. 1105, 1113, 1116 (2021); *see* Germain, *supra* note 2.

[116] Ferguson, *Facial Recognition and the Fourth Amendment*, *supra* note 115, at 1116.

[117] GARVIE, *supra* note 76, at 11.

significantly narrowing the pool of potential suspects using algorithmic processes that the average officer neither understands nor can replicate.[118]

Social media and social network analysis. Law enforcement also uses algorithms to collect and analyze social media data to automate suspicion of individuals likely to engage in future criminal activity.[119] Such tools claim to be able to predict threats by searching for and analyzing patterns in social media usage or connections between existing suspects and people in their network.[120] These tools can also be used to analyze an individual's connections within a social network; those with a high degree of connections may be perceived as influential and become targets for police investigation.[121]

Automated license plate readers (ALPRs). ALPRs use high-speed cameras mounted on police vehicles or stationary structures to automatically take pictures of thousands of license plates a minute.[122] Photographs of the license plate and car, along with information about when and where a license plate was observed, are stored in databases.[123] ALPRs automate suspicion in several ways. Law enforcement can search ALPR databases in cases where a vehicle or license plate (or partial plate) has been identified.[124] Alternatively, ALPR systems can also compare plates scanned in real-time against a list of plates believed to be associated with crimes.[125] Importantly, ALPRs can also identify vehicles by location. Police can search ALPR databases for all vehicles that passed through a particular location in a given time frame.[126]

Police officers are, of course, capable of taking pictures of license plates and comparing them to the plates of suspect vehicles manually, but not at the rate,

[118] *Id.* at 11–12.

[119] Ferguson, *Policing Predictive Policing*, *supra* note 99, at 1137–38; Joh, *Policing by Numbers*, *supra* note 99, at 46–48; *see also* Chaz Arnett, *Black Lives Monitored*, 69 UCLA L. REV. 1384, 1400 (2023) (describing one such company's claim that ninety-seven percent of the alerts it provides are fully automated and have no human involvement).

[120] Joh, *New Surveillance Discretion*, *supra* note 33, at 24–25.

[121] Joh, *Policing by Numbers*, *supra* note 99, at 46–47. Professor Chaz Arnett has described how law enforcement deployed social media analysis tools to monitor Black Lives Matter protestors in the wake of the murder of George Floyd, perpetuating historical surveillance of Black lives. Arnett, *supra* note 119, at 1399.

[122] Amanda Levendowski, *Trademarks as Surveillance Transparency*, 36 BERKELEY TECH. L.J. 439, 457–58 (2021); Rachel Levinson-Waldman, *Hiding in Plain Sight: A Fourth Amendment Framework for Analyzing Government Surveillance in Public*, 66 EMORY L.J. 527, 545 (2017).

[123] Levinson-Waldman, *supra* note 122, at 545.

[124] *See id.*; Levendowski, *supra* note 122, at 458.

[125] Joh, *New Surveillance Discretion*, *supra* note 33, at 22. ALPRs can also be used to conduct real-time tracking of vehicles as they travel. Levendowski, *supra* note 122, at 458.

[126] *See* Levinson-Waldman, *supra* note 122, at 544–45.

volume, or geographic magnitude of ALPR systems.[127] Nor can they track vehicles by license plate without ALPRs or another tracking device.[128] Where police know or suspect that a vehicle involved in a particular criminal activity passed through a specific location, ALPRs automate suspicion of a criminal actor simply by identifying vehicles that passed through that location during a specified time window.

Automated gunshot detection. ShotSpotter seeks not only to identify sounds of gunfire, but by cross-referencing the time the sound reaches each sensor that detects it, also attempts to determine the origin of such sounds.[129] In doing so, ShotSpotter elucidates how policing technology can automate suspicion directly by identifying a specific crime or suspect, or indirectly, by indicating criteria for identifying a suspect that narrow the suspect pool to a degree that only one or very few suspects will satisfy such criteria. When police officers arrive at the location of a ShotSpotter alert, anyone in the alert's proximity is perceived as a suspect, regardless of whether they are acting suspiciously or not.[130] ShotSpotter not only automates suspicion of gun crimes, but also of persons in the vicinity of alerts.

Geofencing and tower dumps. Some policing technologies seek to identify individuals who entered a specified geographic area during a particular timeframe. All individuals who fit the parameters become viable suspects, even if they have no other connection to criminal activity. "Geofences" are one example. Law enforcement increasingly seeks warrants for location data collected by Google for devices that enter a pre-defined geographic area—or, "geofence"—within a certain span of time.[131] Geofencing thus generates suspicion of all persons who enter the geofence during the specified time span, even where police have no other information to connect them to an alleged

[127] *See* Joh, *New Surveillance Discretion*, *supra* note 33, at 22 (noting that ALPRs can "read up to fifty license plates per second, and typically record the date, time, and GPS location of every scanned plate").

[128] *See* Levendowski, *supra* note 122, at 458 (describing the challenge of manually tracking vehicles without ALPRs).

[129] *See Save Lives and Find Critical Evidence with Proven Gunshot Detection*, *supra* note 10.

[130] Sinha, *Automated Gunshot Detection*, *supra* note 34, at 105.

[131] *See* Matthew Guariglia, *Geofence Warrants and Reverse Keyword Warrants Are So Invasive, Even Big Tech Wants to Ban Them*, ELEC. FRONTIER FOUND. (May 13, 2022), https://www.eff.org/deeplinks/2022/05/geofence-warrants-and-reverse-keyword-warrants-are-so-invasive-even-big-tech-wants; Note, *Geofence Warrants and the Fourth Amendment*, 134 HARV. L. REV. 2508, 2512 (2021) (explaining that geofence warrants seek data collected from Android users and anyone who visits a Google application on their phone in a given location); *see also* United States v. Chatrie, 590 F. Supp. 3d 901, 905 (E.D. Va. 2022) (providing background information on geofence warrants).

crime, like actions, conduct, or identifying characteristics.[132] Problematically, geofences can span large geographic areas and significant periods of time, automating suspicion of hundreds or thousands of persons.[133]

Police can also collect the cell phone numbers and other information related to all devices that connected to a particular cell tower during a specified time range via "tower dumps."[134] Tower dumps reveal sensitive personal information and, like geofences, make anyone who happened to be connected to a particular cell tower at a certain time an investigative target, regardless of their conduct or connection to a suspected crime.[135]

GPS and other locational tracking. Locational suspicion can also be generated through GPS and other tracking methods. Physical tracking devices,[136] phones,[137] Apple AirTags,[138] cell site location information,[139] and a variety of other devices and techniques are used to conduct locational tracking, typically using cellphone towers and satellites.[140]

GPS tracking particularly affects individuals subjected to pretrial supervision while pending trial, as well as those on probation or parole following a conviction, who are frequently required to wear a GPS ankle monitor as a condition of their release.[141] GPS ankle monitors use cellphone towers and

[132] *See Geofence Warrants and the Fourth Amendment*, *supra* note 131, at 2508–09 (describing how innocent individuals unconnected to target criminal activity are frequently swept up in geofence searches).

[133] *Id.* at 2509–10 (explaining that while geofence warrants do not often lead police to catch perpetrators of crime, they allow police to access the data of thousands of innocent individuals); *e.g.*, *Chatrie*, 590 F. Supp. 3d at 918 (describing geofence encompassing 17.5 acres). Much of the litigation over the lawfulness of geofence search warrants relates to their breadth and the lack of particularized suspicion of any particular person within the geofence. *See, e.g.*, *Chatrie*, 590 F. Supp. 3d at 925, 934 (finding a geofence warrant invalid for lack of particularized suspicion).

[134] Carpenter v. United States, 138 S. Ct. 2206, 2220 (2018).

[135] *See* Patrick Toomey & Brett Max Kaufman, *The Notice Paradox: Secret Surveillance, Criminal Defendants, & the Right to Notice*, 54 SANTA CLARA L. REV. 843, 890 (2014) (explaining that tower dumps reveal "sensitive data about innocent people quite literally in the wrong place at the wrong time").

[136] United States v. Jones, 565 U.S. 400, 400 (2012).

[137] Marc McAllister, *GPS and Cell Phone Tracking: A Constitutional and Empirical Analysis*, 82 U. CIN. L. REV. 207, 224 (2013).

[138] Thomas Brewster, *The DEA Quietly Turned Apple's AirTag into a Surveillance Tool*, FORBES (Mar. 23, 2023, 11:45 AM), https://www.forbes.com/sites/thomasbrewster/2023/03/23/apple-airtag-becomes-dea-surveillance-device/?sh=3d08a0a13d3d.

[139] *Carpenter*, 138 S. Ct. at 2220.

[140] Kate Weisburd, *Punitive Surveillance*, 108 VA. L. REV. 147, 155–56 (2022).

[141] Chaz Arnett, *From Decarceration to E-Carceration*, 41 CARDOZO L. REV. 641, 672 (2019); *see also* Kate Weisburd, *Sentenced to Surveillance: Fourth Amendment Limits on Electronic Monitoring*, 98 N.C. L. REV. 717, 726 (2020) [hereinafter Weisburd, *Sentenced to Surveillance*] (describing how electronic monitoring is widely used to track the locations of people involved with the criminal legal system).

satellites to constantly monitor the whereabouts of people on pre-trial release, probationers, and parolees—a feat police cannot accomplish using conventional techniques.[142]

Like ShotSpotter, geofences, and tower dumps, GPS devices isolate suspects based on their proximity to criminal activity, rather than based on suspicious conduct or identifying characteristics.[143] When a crime occurs, police may target individuals wearing ankle monitors in the vicinity of a crime, even if there is no other reason to suspect their involvement.[144]

Field tests. Police regularly use handheld devices during traffic stops to develop suspicion for a wide variety of offenses.[145] Portable breath tests, which seek to estimate blood alcohol level, and field narcotics tests—chemical tests that change color based on the presence of certain compounds—have long been in use for presumptive detection of possible drugs and blood alcohol level.[146] Expensive, newer field tests, including some that purport to allow contactless testing for hundreds of substances without the use of chemical kits, are increasingly being used nationwide.[147]

While presumptive tests provide suspicion for searches and seizures that lead to arrests and prosecutions, they cannot confirm the presence of an illicit substance and sometimes produce erroneous results.[148] Many field test results are considered too unreliable for admission in court.[149] Even so, they continue to be used widely, and are often the only type of testing conducted in a case.[150]

[142] *See* Weisburd, *Punitive Surveillance*, *supra* note 140, at 154–155.

[143] *See* Weisburd, *Sentenced to Surveillance*, *supra* note 141, at 729.

[144] *See id.* at 729, 770.

[145] *See, e.g.*, Michael D. Blanchard & Gabriel J. Chin, *Identifying the Enemy in the War on Drugs: A Critique of the Developing Rule Permitting Visual Identification of Indescript White Powder in Narcotics Prosecutions*, 47 AM. U. L. Rev. 557, 583 (1998) (explaining that field tests for controlled substances are often conducted during traffic stops).

[146] Roberts, *supra* note 22, at 797; Stacy Cowley & Jessica Silver-Greenberg, *These Machines Can Put You in Jail. Don't Trust Them.*, N.Y. TIMES (Nov. 3, 2019), https://www.nytimes.com/2019/11/03/business/drunk-driving-breathalyzer.htm.

[147] Jack Evans, *Pushing Fentanyl Fear, Pinellas Sheriff Gets Thousands for Drug Tests*, TAMPA BAY TIMES (Mar. 28, 2023), https://www.tampabay.com/news/pinellas/2023/03/28/fentanyl-pinealls-county-bob-gualtieri-debunked-misinformation-drug-testing/; *TruNarc™ Handheld Narcotics Analyzer*, THERMOFISHER SCI., https://www.thermofisher.com/order/catalog/product/TRUNARC (last visited Jan. 14, 2024).

[148] *See* Roberts, *supra* note 22, at 797.

[149] *Id.*; *see also* State v. Shuler, 858 N.E.2d 1254, 1257 (Ohio App. 2006) (finding portable breath tests unreliable).

[150] Roberts, *supra* note 22, at 797–98.

Law enforcement agencies across the country also employ Rapid DNA test kits, which purport to allow on-site DNA testing in just ninety minutes, without the need for typical laboratory equipment or analytical training.[151] Use of Rapid DNA test kits may exacerbate known concerns with traditional DNA testing, particularly because testing is not conducted in controlled laboratory environments, increasing the chance of contamination.[152]

Field testing using such devices requires no expertise.[153] Officers do not understand the scientific processes used to produce results, lack technical understanding of device limitations, and have used field tests for purposes that lack reliability.[154]

C. Reliability of Policing Technology that Automates Suspicion

Legal reliability refers to the trustworthiness of evidence.[155] Ensuring that automated suspicion technology is trustworthy matters not only because of the Fourth Amendment requirement that searches and seizures be supported by reliable information,[156] but also for practical, commonsense reasons. Premising a search or seizure on technology that is inaccurate in identifying criminal activity or suspicious actors can have a domino effect of harmful consequences. Most obviously, individual searches and seizures may not be properly justified, but in addition, whole communities may be subjected to surveillance and search and seizure tactics based on inaccurate information.

Fears of police use of unreliable technology are not theoretical; there are many documented instances of policing technology getting it wrong. Predictive

[151] *See* Maura Dolan, *'Rapid DNA' Promises Breakthroughs in Solving Crimes. So Why Does It Face a Backlash?*, L.A. TIMES (Sept. 25, 2019), https://www.latimes.com/california/story/2019-09-24/rapid-dna-forensics-crime-police.

[152] Joseph Goldstein, Note, *Guilty Until Proven Innocent: The Failure of DNA Evidence*, 12 DREXEL L. REV. 597, 622 (2020).

[153] *See* Heather Murphy, *Coming Soon to a Police Station Near You: The DNA 'Magic Box'*, N.Y. TIMES (Jan. 21, 2019), https://www.nytimes.com/2019/01/21/science/dna-crime-gene-technology.html (describing ease of use of Rapid DNA devices).

[154] *See* Vera Eidelman & Jay Stanley, *Rapid DNA Machines in Police Departments Need Regulation*, ACLU (Oct. 2, 2019), https://www.aclu.org/news/privacy-technology/rapid-dna-machines-police-departments-need (describing contamination in and police misuse of Rapid DNA kits).

[155] Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 590 n.9 (1993); Hal S. Stern, Maria Cuellar & David Kaye, *Reliability and Validity of Forensic Science Evidence*, SIGNIFICANCE, Apr. 2019, at 21, 22 (explaining that under the "traditional legal and colloquial definition[,]" "reliability is used to denote something trustworthy."). "[T]he law uses 'reliability' to mean that which can be relied on as accurate or truthful." *Stern* et al., *supra*, at 22.

[156] Steven Grossman, *Whither Reasonable Suspicion: The Supreme Court's Functional Abandonment of the Reasonableness Requirement for Fourth Amendment Seizures*, 53 AM. CRIM. L. REV. 349, 354–55 (2016).

policing tools are known to make prediction errors[157] and the companies that sell these tools are aware of errors in their databases.[158] ShotSpotter has been known to misclassify non-gunfire as gunfire and fail to correctly identify true gunfire.[159] PGS software has erroneously identified suspect DNA profiles.[160] Facial recognition software has received perhaps the most attention among policing technologies for its flaws, with frequent reports of errors and wrongful arrests dotting media headlines.[161]

Errors like these arise for a number of reasons. Most fundamentally, errors occur when a technology has not been sufficiently tested to ensure its reliability, is pushed past its reliable limits, or is used for purposes without established reliability.[162] Many policing technologies fall into the category of having undergone insufficient testing, particularly independent testing,[163] to establish their validity.[164] For example, some researchers have concluded that insufficient studies have been conducted to vet predictive policing programs.[165] Much of the

[157] *See* LEE ET AL., *supra* note 57, at 47 (describing scenarios in which predictive policing tools make skewed predictions).

[158] *Id.*

[159] Sinha, *Automated Gunshot Detection*, *supra* note 34, at 83 (summarizing known errors made by ShotSpotter).

[160] *See* Stiffelman, *supra* note 112, at 124 (describing case in which two PGS systems generated conflicting results when analyzing the same DNA profile).

[161] *See, e.g.*, Khari Johnson, *How Wrongful Arrests Based on AI Derailed 3 Men's Lives*, WIRED (Mar. 7, 2022), https://www.wired.com/story/wrongful-arrests-ai-derailed-3-mens-lives/; *see also* Gray, *supra* note 26, at 4 (describing wrongful arrest of Robert Williams based on erroneous facial recognition identification).

[162] *See* Brief of 42 Scholars of Forensic Science as *Amici Curiae* in Support of Defendant-Appellee at 2, United States v. Gissantaner, 990 F.3d 457 (6th Cir. 2021) (No. 19-2305), 2020 WL 3316839 [hereinafter Brief of 42 Scholars] (arguing that PGS software is unreliable when "stretched beyond its capacity, or when applied by a lab that failed to properly establish its limits").

[163] Independent testing refers to testing conducted by those who have no interest in the outcome of testing. Independent is used here to mean free from conflict of interest, financial or otherwise, not merely to refer to a third party. *See, e.g.*, PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS 81 (2016) [hereinafter PCAST REPORT], https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf (explaining that "[validation] studies should be performed by or should include independent research groups not connected with the developers of the methods and with no stake in the outcome").

[164] *See* LEE ET AL., *supra* note 57, at 5, 45 (defining "data-driven policing" as "tools that analyze data to determine where, how, and who to police"; and concluding that "data-driven policing technologies . . . lack scientific validity"); GARVIE, *supra* note 76, at 16 (finding that "face recognition as U.S. law enforcement uses it today currently lacks the scientific validity required to consider it a reliable forensic technique").

[165] *See, e.g.*, DAVID ROBINSON & LOGAN KOEPKE, UPTURN, STUCK IN A PATTERN 7 (2016), https://www.upturn.org/static/reports/2016/stuck-in-a-pattern/files/Upturn_-_Stuck_In_a_Pattern_v.1.01.pdf (stating that their "research surfaced few rigorous analyses of predictive policing systems' claims of efficacy, accuracy, or crime reduction"); *see also* LEE ET AL., *supra* note 57, at 50 ("[I]ndependent empirical studies have yet to be conducted on Palantir Technologies' highly secretive data-driven policing systems.").

research on ShotSpotter lacks independence and rigor, and at least some data suggests the system is not scientifically valid.[166] A comprehensive review of DNA mixture interpretation by the National Institute of Standards and Technology (NIST) found that currently available data is insufficient to even assess the reliability of PGS systems.[167] And forensic labs have been criticized for failing to adequately validate PGS software under laboratory conditions.[168]

Although facial recognition software continues to be researched, Clare Garvie, a noted facial recognition expert, has explained that "no study has comprehensively examined the reliability of face recognition as actually used by a representative sample of U.S. law enforcement officers, taking into account the full range of possible variabilities."[169] Other data suggest a lack of reliability,[170] particularly in classifying faces of individuals representing certain demographics. One landmark study found significant errors among some programs in classifying female versus male faces and specifically in classifying dark-skinned female faces.[171] Another study conducted by NIST found higher rates of false positives (incorrect identifications) in classifying Asian and Black faces than for white faces and higher rates of false positives in classifying Black female faces.[172]

Data sets that algorithmic tools rely on are an additional—and significant—source of error. Although they may seem objective and unbiased, data sets are developed using discretionary and subjective human inputs. Humans decide which analytical models to employ, which features models should consider, and

[166] Sinha, *Automated Gunshot Detection*, *supra* note 34, at 80–81 (summarizing concerns with ShotSpotter's reliability testing including researchers' conclusion that "little meaningful evidence of ShotSpotter's accuracy currently exists").

[167] *See* NAT'L INST. OF STANDARDS & TECH., *supra* note 110, at 75.

[168] *See* Brief of 42 Scholars, *supra* note 162, at 2.

[169] *See* GARVIE, *supra* note 76, at 16.

[170] *See* GARVIE ET AL., *supra* note 113, at 46 ("Compared to fingerprinting, state-of-the-art face recognition is far less reliable and well-tested.").

[171] Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 PROC. MACH. LEARNING RSCH. 1, 10–11 (2018), http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf (studying three facial recognition systems and finding that high false positives in classifying female faces versus male faces and error rates of up to 34.7% in classifying dark-skinned female faces).

[172] PATRICK GROTHER, MEI NGAN & KAYEE HANAOKA., NAT'L INST. OF STANDARDS & TECH., FACE RECOGNITION VENDOR TEST (FRVT): PART 3: DEMOGRAPHIC EFFECTS 2–3 (2019), https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf. Interestingly, the study found that algorithms developed in China had low false positive rates when classifying Asian faces. *Id.* at 2.

how to analyze data.[173] The data in crime databases also represents the discretionary decisions that police and prosecutors make about how to respond to alleged crimes, including who to stop, arrest, and charge.[174] As a result of such discretionary decision-making and recording or input errors, data sets include incomplete and erroneous crime data.[175]

No data set perfectly represents crime in any jurisdiction. Databases are both under-inclusive and over-inclusive. They are under-inclusive because not all crime is reported and recorded.[176] Police do not respond to all crimes, and do not record all crimes they respond to.[177] Databases are over-inclusive because they can include: (1) "crimes" that never happened or were affirmatively falsified; and (2) "crimes" based on planted evidence, race-based arrests, or other factors that either do not represent crime at all, or paint an inaccurate picture of crime.[178]

Database errors are replicated by predictive tools; training a predictive model on faulty data will yield faulty predictions.[179] Crucially, when crime data used to make predictions reflects biased policing practices, predictive tools will replicate such biases when used by law enforcement.[180]

[173] LEE ET AL., *supra* note 57, at 46 (noting that "the basic building blocks of a predictive software program involve many human discretionary decisions"); GARVIE, *supra* note 76, at 14 (explaining that facial recognition "involv[es] significant human judgment"); Rich, *supra* note 33, at 885 (explaining that prediction errors can result from the factors chosen for a model to consider by humans).

[174] *See* Elizabeth E. Joh, *Feeding the Machine: Policing, Crime Data, & Algorithms*, 26 WM. & MARY BILL RTS. J. 287, 296–97 (2017) [hereinafter Joh, *Feeding the Machine*] (describing how crime data is the product of many processes, including police officer discretion). In turn, scholars have argued that predictive policing tools do not predict future crime or criminality at all; rather, they predict how police will *respond* to crime. Kristian Lum & William Isaac, *To Predict and Serve?*, SIGNIFICANCE, Oct. 2016, at 14, 16. *See generally* ROBINSON & KOEPKE, *supra* note 165, at 5–6.

[175] LEE ET AL., *supra* note 57, at 46; *see also* Rich, *supra* note 33, at 884 (describing "noise" in training data or information in data sets that is incorrect).

[176] Ferguson, *supra* note 100, at 317.

[177] William Isaac & Kristian Lum, *Setting the Record Straight*, IN JUSTICE TODAY (Jan. 3, 2018), https://medium.com/in-justice-today/setting-the-record-straight-on-predictive-policing-and-race-fe588b457ca2.

[178] Rashida Richardson, Jason M. Schultz & Kate Crawford, *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice*, 94 N.Y.U. L. REV. ONLINE 15, 18, 29, 31–32 (2019).

[179] Lum & Isaac, *supra* note 174, at 15; LEE ET AL., *supra* note 57, at 46; Rich, *supra* note 33, at 885; Joh, *Feeding the Machine*, *supra* note 174, at 290.

[180] *See* Dhruv Mehrotra, Surya Mattu, Annie Gilbertson & Aaron Sankin, *How We Determined Predictive Policing Software Disproportionately Targeted Low-Income, Black, and Latino Neighborhoods*, GIZMODO (Dec. 2, 2021), https://gizmodo.com/how-we-determined-predictive-policing-software-dispropo-1848139456 (finding, in a study with limitations stemming from a lack of complete data, that there was not "a strong correlation between arrests and predictions[,]" but that "PredPol's algorithm . . . disproportionately targeted vulnerable populations" and "that its predictions disproportionately targeted neighborhoods with proportionately more Black and Latino residents"); Lum & Isaac, *supra* note 174, at 19 (finding that "predictive policing of drug

Establishing the reliability of policing technologies can reduce erroneous outputs and the harms that result from such errors. Reliability is measured by assessing scientific validity,[181] or how well a technology performs its intended function.[182] Scientific validity requires not only that a technology perform as intended in some circumstances, but that it perform as intended under the specific conditions it is being used in, or, that it performs as intended *as applied*.[183]

Scientific validity, in turn, is established through empirical testing in which testing administrators know the correct result, or ground truth.[184] Testing must be conducted under conditions representative of those that a technology will be used in.[185] This serves to establish the limits of a technology's reliability by informing users when a technology will no longer perform predictably.[186] Validity assessments are complicated by a lack of transparency surrounding many policing technologies. Governments, law enforcement agencies, and companies that produce policing technologies aggressively resist releasing information about their products, preventing independent researchers from assessing validity.[187]

crimes results in increasingly disproportionate policing of historically over-policed communities"); Lum & Isaac, *supra* note 174, at 18 ("[R]ather than correcting for the apparent biases in the police data, the [PredPol predictive policing] model reinforces these biases.").

[181] Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 590 n.9 (1993) ("In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity*." (emphasis in original)).

[182] Stern et al., *supra* note 155, at 22–23; David H. Kaye & David A. Freedman, *Reference Guide on Statistics*, *in* FEDERAL JUDICIAL CENTER, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 228 (3d ed.) ("A valid measuring instrument measures what it is supposed to.").

[183] *See* IEEE COMPUTER SOCIETY, IEEE STANDARD 1012-2016: STANDARD FOR SYSTEM, SOFTWARE, AND HARDWARE VERIFICATION AND VALIDATION 15 (2017) [hereinafter IEEE Standard 1012] ("[Verification & validation] processes determine whether the development products of a given activity conform to the requirements of that activity and whether the product satisfies its intended use and user needs."); *Daubert*, 509 U.S. at 591 ("[S]cientific validity for one purpose is not necessarily scientific validity for other, unrelated purposes.").

[184] PCAST REPORT, *supra* note 163, at 33, 47; Stern et al., *supra* note 155, at 22 ("[V]alidity . . . is best assessed by examining reported results in . . . cases in which the researcher knows the correct answer.").

[185] *See supra* note 183 and accompanying text; GARVIE, *supra* note 76, at 15 ("For a forensic technique to be considered scientifically valid, it must be subjected to empirical testing, under conditions representative of its operational use."); Stern et al., *supra* note 155, at 22 ("[V]alidity . . . is best assessed by examining reported results in representative (or more challenging) cases . . . .").

[186] Krane & Philpott, *supra* note 109, at 298.

[187] Jonathan Manes, *Secrecy & Evasion in Police Surveillance Technology*, 34 BERKELEY TECH. L.J. 503, 507 (2019); *see also* Ram, *supra* note 111, at 666–82 (describing efforts of private developers, police departments, and prosecutors to prevent the release of information about use of policing technologies).

II. INVESTIGATIVE RELIABILITY VERSUS TRIAL STAGE RELIABILITY

There are two primary stages of a prosecution at which a judge may consider the reliability of policing technology.[188] The first is the suppression hearing, at which a judge evaluates the legality of searches and seizures under the Fourth Amendment.[189] The second is prior to its admission at a trial.[190] At a trial, a party may seek to admit evidence relating to police use of technology on the issue of an accused's guilt or innocence. This Part considers the different standards that govern how judges are to assess the reliability of technological evidence at each stage.

A. Investigative Reliability

The Fourth Amendment requires that searches and seizures be supported by reasonable suspicion or probable cause, depending on the circumstance and type of intrusion.[191] Both reasonable suspicion and probable cause require police to possess particularized suspicion that: (1) criminal activity is occurring (or has occurred or is imminent) and (2) the person subjected to the intrusion is responsible.[192]

[188] This is not to suggest that policing technology reliability cannot be raised or considered in other criminal litigation. It might, for example, be challenged in a hearing on probable cause for detention. Gerstein v. Pugh, 420 U.S. 103, 103 (1975). Such challenges are beyond the scope of this Article.

[189] *See* LAFAVE, *supra* note 31, § 3.1(d) (summarizing the purpose of suppression hearings and how they are conducted). Judges' consideration of whether evidence was obtained in violation of constitutional rights at a suppression hearing is not limited to the Fourth Amendment. *See* MARK S. RHODES, ORFIELD'S CRIMINAL PROCEDURE UNDER THE FEDERAL RULES (Clark Boardman Callaghan ed., 2d ed. 2022). At a suppression hearing, judges may also determine, for example, whether government actors obtained statements from an accused person in violation of *Miranda v. Arizona*, 384 U.S. 436 (1966), or involuntarily. Dickerson v. United States, 530 U.S. 428, 432 (2000). Non-Fourth Amendment suppression considerations are beyond the scope of this Article.

[190] *See* FED. R. EVID. 702 (requiring that expert evidence admitted at trial be reliable). The reliability of scientific, technical, and specialized evidence is not typically addressed at a third prosecution stage, the sentencing hearing. Sinha, *Junk Science, supra* note 62, at 87, 92. As this author has argued previously, however, because of the importance of sentencing hearings in the modern criminal legal system, greater scrutiny of the reliability of such evidence is necessary at sentencing. *See id.* at 92–93.

[191] A full arrest or search requires probable cause. Henry v. United States, 361 U.S. 98, 102 (1959). A full search also requires probable cause. Chambers v. Maroney, 399 U.S. 42, 51 (1970) ("In enforcing the Fourth Amendment's prohibition against unreasonable searches and seizures, the Court has insisted upon probable cause as a minimum requirement for a reasonable search permitted by the Constitution."). A temporary investigative seizure requires reasonable suspicion. Terry v. Ohio, 392 U.S. 1, 27 (1968). A limited frisk, or pat-down of a suspect's outer clothing requires reasonable suspicion that a suspect is armed and dangerous. *Id.* at 30–31.

[192] *See* United States v. Hensley, 469 U.S. 221, 227 (1985) (extending rule permitting stops based on reasonable suspicion for crimes occurring or about to occur to completed crimes); *see also* Brown v. Texas, 443 U.S. 47, 51 (1979); *Terry*, 392 U.S. at 21; Florida v. J.L., 529 U.S. 266, 271 (2000); United States v. Cortez, 449

An accused person may challenge the sufficiency of suspicion at a pretrial suppression hearing.[193] Whether the requisite degree of suspicion is present is assessed under a flexible, totality of circumstances test.[194] If a search or seizure is found to be unlawful, the general remedy is exclusion—or suppression—of evidence recovered via the illegal intrusion.[195]

Fourth Amendment law requires judges to defer to officers' judgment in assessing the reasonableness of searches and seizures,[196] under the rationale that police are trained and have expertise necessary to assess suspicious activity and actors.[197] Deference to police judgment is a much-criticized aspect of criminal procedure doctrine.[198] Courts defer to officer judgments indiscriminately, without examination of whether an individual officer's specific training and experience justifies it.[199] Moreover, uncritical deference to officers has resulted in erosion of the Fourth Amendment's promised protections.[200] Officers abuse the discretion that they are allowed to conduct flagrant and rampant Fourth Amendment violations, often involving harassment and violence.[201] These

U.S. 411, 417–18 (1981) (finding that, based upon the totality of the circumstances, "detaining officers must have a particularized and objective basis for suspecting the particular person stopped of criminal activity").

[193] *See* LAFAVE, *supra* note 31, § 3.1(d).

[194] *Cortez*, 449 U.S. at 417–18.

[195] *See* Mapp v. Ohio, 367 U.S. 643, 655–56 (1961) (extending rule making evidence obtained by an unlawful search inadmissible in federal courts to state courts); Wong Sun v. United States, 371 U.S. 471, 485–88 (1963).

[196] *Terry*, 392 U.S. at 27.

[197] *Id.* ("[I]n determining whether the officer acted reasonably . . . due weight must be given . . . to the specific reasonable inferences which he is entitled to draw from the facts in light of his experience."); *see also Cortez*, 449 U.S. at 418 ("[A] trained officer draws inferences and makes deductions—inferences and deductions that might well elude an untrained person.").

[198] *See, e.g.*, L. Song Richardson, *Police Efficiency and the Fourth Amendment*, 87 IND. L.J. 1143, 1179 (2012) [hereinafter Richardson, *Police Efficiency*] (noting that the "broad discretion" police officers currently enjoy does not align with the Framers' intentions); Harris, *supra* note 56, at 4–6 (predicting that deference to police judgment would result in erosion of Fourth Amendment protections); Miller, *supra* note 56, at 224 ("[T]he Court . . . removes both judicial and public scrutiny through deference to some inarticulable police 'sixth sense' about crime."); I. Bennett Capers, *Crime, Legitimacy, and Testilying*, 83 IND. L.J. 835, 866 (2008) ("[O]fficers know they can misrepresent their motives for conducting stops without consequences.").

[199] Richardson, *Police Efficiency*, *supra* note 198, at 1155 ("[C]ourts repeatedly defer to the judgments of *all* officers, with no inquiry into the particular officer's training, experience, and skill.").

[200] *See, e.g.*, Tracey Maclin, Terry v. Ohio*'s Fourth Amendment Legacy: Black Men and Police Discretion*, 72 ST. JOHN'S L. REV. 1271, 1309 (1998) (discussing how *Terry* "gave officers enormous discretion and diminished the constitutional freedom of the individual").

[201] *See* Harris, *supra* note 56, at 5 (explaining that police have been increasingly willing and able to "automatically frisk" people they stop, even without the presence of any individualized circumstances which point to danger); David H. Gans, *"We Do Not Want to Be Hunted": The Right to Be Secure and Our Constitutional Story of Race and Policing*, 11 COLUM. J. RACE & L. 239, 309 (2021) (explaining that stop-and-frisk policing "creates the potential for a tragic violent encounter between the police and the populace").

harms are disproportionately suffered by members of marginalized communities.[202] On top of this, there is little evidence that the original rationale for granting police officers such deference is justified: research shows that police officers are not particularly effective at uncovering crime and routinely judge criminality incorrectly.[203]

Against this backdrop, turning to policing technology may seem like an ideal solution. But automating suspicion contributes to unjustified Fourth Amendment intrusions in unexpected and different ways. Technologies that point to a place, like crime mapping software or ShotSpotter, encourage police to assume a person is acting criminally based on their physical location, even when other factors such as descriptive information, suspicious conduct, or demeanor, suggest the opposite. Technologies that point to a person, like the predictive program used in Bobby Jones's case, encourage police to assume that crime is occurring, even if it is not.

As described in Part I, law enforcement often assumes the accuracy of automated suspicion technology, even when reliability has not been established.[204] In other words, policing technology may be even *worse* at identifying crime and criminal actors than police officers, even if those who rely on it are unaware of this.

Moreover, policing technology confounds the remaining checks against unjustified searches and seizures. As the Supreme Court has consistently pronounced, reasonable suspicion and probable cause must be supported by

[202] *See* L. Song Richardson, *Arrest Efficiency and the Fourth Amendment*, 95 MINN. L. REV. 2035, 2052 (2011) [hereinafter Richardson, *Arrest Efficiency*] (explaining that implicit bias leads to police targeting of Black people).

[203] KRISTIN HENNING, THE RAGE OF INNOCENCE: HOW AMERICA CRIMINALIZES BLACK YOUTH 212 (Pantheon 2021) (noting that Black and Latino male youth in New York City are often stopped by police despite being innocent of any crime and that the vast majority of those stopped for weapons offenses did not have a weapon); Richardson, *Police Efficiency*, *supra* note 198, at 1145 (summarizing rates at which police find evidence of criminal activity when conducting a stop-and-frisk as ranging from as low as 1.89% to just 23.53% for some cities); *see also* Richardson, *Arrest Efficiency*, *supra* note 202, at 2063 (arguing that: "[T]he nature of their jobs may lead officers to perform no better than civilians when it comes to differentiating criminal from noncriminal activity. They perhaps may perform even worse in situations where nonwhites are involved."); Shima Baradaran Baughman, *Crime and the Mythology of Police*, 99 WASH. U. L. REV. 65, 110 (2021) ("[O]n a good year, police solve less than a quarter of reported cases.").

[204] *See supra* Part I.

reliable information.[205] Police cannot rely on information that is wholly untrustworthy to justify a stop.[206]

At a suppression hearing, a judge determines whether reasonable suspicion or probable cause supported the intrusion at issue and can evaluate the reliability of information police relied on in support of a search or seizure as a part of that process.[207] Assessing the reliability of police observations is straightforward, if not always easy. A judge can examine what the officer observed, evaluate the officer's training and experience in similar cases, and evaluate the officer's credibility.[208]

The reliability assessment may be more difficult when police rely on information that is not obtained firsthand. As a result, the Court has outlined some guiding frameworks for assessing reliability in a Fourth Amendment inquiry,[209] most often in analyzing whether reasonable suspicion and probable cause can be supported by information provided by third parties.[210] The Court has also addressed how to assess the reliability of drug-sniffing dogs that purport to alert police to the presence of drugs.[211]

Although the Supreme Court's analysis of reliability of third-party information appears substantially different from its drug dog reliability analysis, closer examination reveals a through line between the two frameworks.[212] As described next, the Court looks to external indicators of reliability separate from the information provided or its source. Examples include the existence of corroborating evidence, evidence of a source's prior reliability, and reliability

[205] Alabama v. White, 496 U.S. 325, 328–29 (1990); *see also* Grossman, *supra* note 156, at 349 (describing the Supreme Court's pronouncement that probable cause and reasonable suspicion must be supported by reliable information as "constant").

[206] *See* Grossman, *supra* note 156, at 349 (explaining that if "the reliability of the person providing the information or of the information itself" cannot be demonstrated, "no search or seizure based on probable cause or reasonable suspicion is permitted").

[207] LAFAVE, *supra* note 31, § 11.2(d).

[208] *Id.*

[209] *See* Grossman, *supra* note 156, at 354–55 (explaining that the Supreme Court "has enumerated various methods through which the government can demonstrate that the information or the person supplying evidence of a crime or criminal activity is sufficiently reliable to permit the government to search or seize").

[210] *E.g.*, Aguilar v. Texas, 378 U.S. 108, 114 (1964); Spinelli v. United States, 393 U.S. 410, 415–16 (1969); Illinois v. Gates, 462 U.S. 213, 230 (1983); *White*, 496 U.S. at 328–29; Florida v. J.L., 529 U.S. 266, 270–71 (2000); Florida v. Harris, 568 U.S. 237, 244–45 (2013); Navarette v. California, 572 U.S. 393, 398 (2014); *see also* Jones v. United States, 362 U.S. 257, 271 (1960); Adams v. Williams, 407 U.S. 143, 147 (1972).

[211] *See generally Harris*, 568 U.S. 237 (addressing how to assess the reliability of drug-sniffing dogs that purport to alert police to the presence of drugs).

[212] *See infra* Part II.A.1.

assessments by outside organizations to scrutinize the reliability of information police officers do not obtain through their own observations or inferences.[213]

1. *Assessing the Reliability of Third-Party Information in Reasonableness Determinations*

The Supreme Court assesses reliability of third-party information used to justify a search or seizure differently depending on whether the informant is a citizen informant with a known identity, a criminal informant with a known identity, or an anonymous informant. For each type of informant, however, the Court evaluates reliability by looking beyond the tip itself for indicators of the tip's reliability.

In order to determine the reliability of information provided to police by known informants, the Court has looked to indicia of reliability external to the tip itself. Courts treat a known identity, absent criminal history or affiliation, as indicating a tip's reliability even without assessing the tip's content.[214] Known informants typically enjoy a presumption of reliability even if the information they provide cannot be corroborated, under the rationale that the possibility of prosecution for falsifying a report is a deterrent to lying and that informants who voluntarily provide identifying information have no apparent motive to lie.[215]

Criminal informants are considered less reliable than non-criminal informants because they may be seeking a benefit and because their criminal conduct is perceived as indicating inherent unreliability.[216] Courts evaluate the reliability of criminal informant tips by assessing the reliability of prior tips provided by the same informant, the informant's basis of knowledge, and, critically, whether the information provided can be corroborated.[217]

The Supreme Court has viewed anonymous tipsters with some skepticism because they are not likely to face repercussions for falsifying a report and because their motivations for providing police with information are not easily

[213] *See infra* Part II.A.1.

[214] *See* Ariel C. Werner, Note, *What's in a Name? Challenging the Citizen-Informant Doctrine*, 89 N.Y.U. L. REV. 2336, 2348–50 (2014) (explaining that courts generally treat known citizen informants as reliable).

[215] *See id.* at 2360 (explaining that the typical justifications for presuming reliability of a known citizen-informant include a lack of motive to lie and the deterrent effect of knowing that serious repercussions for falsifying a report exist).

[216] *See* Rich, *supra* note 33, at 908. *But see* Werner, *supra* note 214, at 2365–66 (analyzing competing incentives of criminal informants to be truthful or falsify information).

[217] *See* Werner, *supra* note 214, at 2338–39 (summarizing factors considered to assess criminal informant reliability).

discernable.[218] Consequently, the Court looks to factors external to the source of the tip that can suggest whether a tip is reliable. Specifically, the Court has examined whether a tip contains predictions of future conduct that can be independently corroborated.[219] If a tip predicts future actions by a target that are corroborated by police observation, that information is likely to be found sufficiently reliable to justify an intrusion.[220] The Court has also suggested that an anonymous tipster's basis of knowledge and the level of detail provided can support a tip's reliability.[221]

2. *Assessing the Reliability of Drug-Sniffing Dog Alerts in Reasonableness Determinations*

The Supreme Court has also addressed how to evaluate the reliability of drug-sniffing dogs in Fourth Amendment reasonableness determinations.[222] In *Florida v. Harris*, the Court considered whether a positive alert by a drug-sniffing dog was sufficient to establish probable cause to search a vehicle.[223] In *Harris*, a trained police drug-sniffing dog signaled an alert for the presence of drugs at the driver's side of Clayton Harris's truck during a traffic stop.[224] In response, the officer who initiated the stop searched the vehicle, discovered chemical substances that could be used to produce methamphetamine, and arrested Harris.[225] Harris moved to suppress the items recovered from his truck, arguing that the dog's alert did not establish probable cause for the search.[226] He argued that the dog's certification, which had expired prior to the stop, and its performance in the field did not establish that the dog could reliably perform drug sniffs.[227] The trial court denied the motion to suppress and the Florida Supreme Court reversed, finding that a canine alert will not establish probable

[218] Illinois v. Gates, 462 U.S. 213, 227 (1983); Florida v. J.L., 529 U.S. 266, 270 (2000).

[219] *See Gates*, 462 U.S. at 227; Alabama v. White, 496 U.S. 325, 331 (1990).

[220] *See White*, 496 U.S. at 331 (finding tip sufficiently reliable to justify *Terry* stop where some information provided was corroborated by police observation). In practice, however, the Supreme Court has found intrusions to be justified even where information provided by an informant was not corroborated or was contradicted. *Id.* (upholding *Terry* stop where some information provided by the tipster was inaccurate and some was not corroborated); *see also* Navarette v. California, 572 U.S. 393, 404 (2014) (upholding *Terry* stop where anonymous tipster's claim that a vehicle was being driven recklessly was not corroborated).

[221] *See, e.g.*, *White*, 496 U.S. at 331; *J.L.*, 529 U.S. at 270; *Navarette*, 572 U.S. at 398.

[222] Florida v. Harris, 568 U.S. 237, 240 (2013).

[223] *Id.*

[224] *Id.*

[225] *Id.* at 241.

[226] *Id.*

[227] *Id.* at 242.

cause unless the State provides documentation of the dog's field performance supporting its reliability.[228]

The Supreme Court reversed.[229] Writing for a unanimous Court, Justice Elena Kagan made two primary critiques of the Florida Supreme Court's reasoning. First, she found that the Florida high court's analysis was inconsistent with the established totality of circumstances approach for determining probable cause.[230] The Court described probable cause as a "flexible," "practical and common-sensical standard" that does not require the reliability of a drug-sniffing dog to be established through a "strict evidentiary checklist."[231]

Second, Kagan criticized the Florida Supreme Court's emphasis on the dog's field performance, reasoning that field records may be inaccurate or incomplete, and alerts that do not result in the recovery of drugs may not actually reflect a false positive.[232] Rather, they may signify that a dog has alerted to the residual scent of drugs no longer present.[233] Kagan suggested the reliability is measured better by performance in a controlled environment where ground truth—whether or not drugs are present—is known to the trainer.[234]

The Supreme Court once again looked for external indicators of the dog alert's reliability. It held that a dog's recent completion of a training program or certification by a "bona fide" organization creates a presumption that a dog's alert is sufficiently reliable to establish probable cause to search.[235]

As scholars have complained, Justice Kagan's logic contains a defect.[236] Despite acknowledging that probable cause is determined under an "all-things-considered" totality of circumstances approach, her test—under which a certified dog enjoys a presumption of reliability notwithstanding other factors, like the conduct of the target of the search or problems with the dog's field performance—places too great a weight on a drug-dog's certification.[237]

[228] *Id.* at 244–45.

[229] *Id.* at 250.

[230] *Id.* at 244.

[231] *Id.*

[232] *Id.* at 245–46.

[233] *Id.*

[234] *See id.*

[235] *Id.* at 246–47.

[236] *E.g.*, Rich, *supra* note 33, at 915–18.

[237] *Harris*, 568 U.S. at 246–47; *see also* Rich, *supra* note 33, at 915 (explaining that *Harris* overvalues one data point and will lead to undervaluing of others in the totality of circumstances analysis).

The reasoning in *Harris* also fails in other ways. The Court did not explain which organizations should be considered "bona fide."[238] The Court also seemed not to recognize that different organizations test drug dog reliability in widely varying ways.[239] Consequently, certification may suggest very different levels of reliability depending on the certifying body.[240] Moreover, a wide variety of factors outside of performance in a certification program can affect a dog's performance in the field. Dog handler cues are well known to influence whether a dog alerts and dogs can alert to lawful substances that contain the same compounds as illegal narcotics.[241] In other words, certification alone is not a good measure of a dog's reliability.

Finally, *Harris* nods to the idea that the accused should be able to challenge a dog's reliability, but it does so without meaningful engagement with how to accomplish this practically.[242] Accused persons are likely to be impeded in launching such challenges by both adversarial and administrative hurdles. Information about a dog's reliability is controlled by police and prosecutors who may be unwilling to disclose details regarding a drug dog's training and performance.[243] Moreover, relevant information may be unavailable simply because of poor or incomplete record keeping.[244]

3. *Application of Existing Fourth Amendment Reliability Frameworks to Policing Technology*

Although at a surface level, automated suspicion technology may appear to act like a third-party tipster, it shares only superficial characteristics with human informants.[245] As others have observed, the Supreme Court's analysis of tipster reliability does not map on well to policing technologies.[246] Much of the Court's

[238] *See Harris*, 568 U.S. at 246–47.

[239] Jeremiah K. Geffe, *License to Sniff: The Need to Regulate Privately Owned Drug-Sniffing Dogs*, 19 J. GENDER, RACE & JUST. 167, 191–92 (2016).

[240] Lee Epstein, Barry Friedman & Geoffrey R. Stone, *Foreword: Testing the Constitution*, 90 N.Y.U. L. REV. 1001, 1028–31 (2015) (describing significant variation in dog certification and training programs).

[241] *Id.* at 1035–36.

[242] *See Harris*, 568 U.S. at 247.

[243] *See* Kit Kinports, *The Dog Days of Fourth Amendment Jurisdiction*, 108 NW. U. L. REV. COLLOQUY 64, 65–66 (2013) (describing challenges the defense may face in obtaining access to records regarding a drug dog's reliability, including the government taking the position that such information is not discoverable).

[244] Robert C. Bird, *An Examination of the Training and Reliability of the Narcotics Detection Dog*, 85 KY. L.J. 405, 415 (1997) (describing limited record-keeping at one agency utilizing drug dogs).

[245] *See* Rich, *supra* note 33, at 909 (noting that ASAs may appear similar to informants because they are a non-law enforcement source of information but that they are otherwise not categorizable as informants).

[246] *See id.* at 908.

informant reliability analysis is directed toward assessing a tipster's motivations to be truthful or to falsify, evaluating factors such as whether a tipster's identity is known, whether a tipster can be held responsible for a false report, whether a tip is based on first-hand observation, and whether a tip accurately predicts its subject's future conduct.[247]

These conventional methods of discerning the truthfulness of a tip simply do not apply to most policing technology.[248] Technology may replicate the biases of its developers, but it has no motive to lie in any traditional sense.[249] Unlike human tipsters, the fact that an automated suspicion technology's "identity" (or, at least, its developers' identities) are known or discernible, says nothing about the reliability of its output because policing technologies cannot typically be held responsible for providing erroneous information.[250] Not all policing technology bases outputs on first-hand observation or "insider information." Rather most policing technology outputs are based on analysis of factors and patterns in large data sources.[251] And, policing technology does not always predict future conduct.[252]

Greater insight about how to assess reliability of policing technology output can be drawn from the Supreme Court's analysis of drug-sniffing dog reliability in *Harris*. While drug-sniffing dogs are unlike automated policing technology in most ways, as Professor Michael Rich has observed, they do operate as black boxes.[253] The dogs provide an output—an alert to the presence of drugs—but

[247] *E.g.*, Alabama v. White, 496 U.S. 325, 332 (1990).

[248] *See* Ferguson, *Predictive Policing and Reasonable Suspicion*, *supra* note 100, at 305 (explaining that the "core logic of the tip cases falls away" when applied to predictive analytics because such tools do not identify a particular suspect); *see also* Sinha, *Automated Gunshot Detection*, *supra* note 34, at 99–102 (analyzing how traditional informant reliability analysis fails when applied to ShotSpotter).

[249] *See* Rich, *supra* note 33, at 909 (explaining that developer biases are not relevant to the reasonableness analysis in the same ways they are when a human tipster is concerned). Generative artificial intelligence like ChatGPT, which has been found to produce false information, however, challenges this assumption. *See* Pranshu Verma & Will Oremus, *ChatGPT Invented a Sexual Harassment Scandal and Named a Real Law Prof as the Accused*, WASH. POST (Apr. 5, 2023, 2:07 PM), https://www.washingtonpost.com/technology/2023/04/05/chatgpt-lies/ (describing case in which ChatGPT created a false sexual harassment story involving a real professor).

[250] *See* Rich *supra* note 33, at 909–10.

[251] *Id.*

[252] *See* Sinha, *Automated Gunshot Detection*, *supra* note 34, at 100 (noting that ShotSpotter does not supply predictive information).

[253] *See* Rich, *supra* note 33, at 912 (noting "[d]rug-sniffing dogs are the prototypical black boxes" because "explaining how the input of the residue of an illegal drug is translated in a dog's brain into the output of an 'alert' is beyond the scope of available expert testimony").

the specific process they engage in to reach their determination is opaque to humans.[254]

By reasoning that a dog's performance in controlled environments is a better measure of reliability than its field performance, the Supreme Court in *Harris* recognized that testing under circumstances where ground truth is known is important to establishing the reliability of black box methods.[255] Reliability of black box policing technologies cannot be established by review of field performance, because ground truth is typically unknown in the field.[256] In other words, as *Harris* explains, post-deployment performance, while relevant, cannot establish reliability because myriad factors can obscure an understanding of whether field performance is accurate.[257] These include the scent of drugs no longer present, other scents that contain similar compounds to illegal drugs, handler influences, and variation in ability to detect different drugs, among others.[258]

Although the Court's focus on testing under controlled conditions makes general sense, its analysis lacks depth. No method is always reliable;[259] any given method is only reliable for particular purposes under certain conditions.[260] In order for a policing technology to be considered reliable under any particular set of conditions, the technology must be tested in a controlled environment under those conditions, or, reasonable approximations of those conditions.[261]

Blanket treatment of certification as a proxy for reliability fails to account for these limitations. Even generally well-trained drug dogs will not perform reliably under all field conditions. Knowing whether a dog will perform reliably under a certain set of conditions requires assessing their performance under equivalent conditions. As previously described, however, dog training varies widely by program and organization.[262]

[254] *Id.* (explaining that humans know the inputs and the receive the outputs but that "[humans] cannot fully understand how the internal mechanism works").

[255] Florida v. Harris, 586 U.S. 237, 246 (2013).

[256] Sinha, *Automated Gunshot Detection*, *supra* note 34, at 106.

[257] *Harris*, 586 U.S. at 245–46.

[258] Epstein et al., *supra* note 240, at 1031–36.

[259] *See supra* note 183.

[260] *See supra* note 183.

[261] *See* GARVIE, *supra* note 76, at 15 ("For a forensic technique to be considered scientifically valid, it must be subjected to empirical testing, under conditions representative of its operational use.").

[262] *See supra* note 240 and accompanying text.

B. Trial Stage Reliability

At the trial stage, if a party seeks to admit technological evidence, their opponent may seek exclusion of that evidence as irrelevant or unreliable.[263] Generally, one of two standards governs admissibility of such evidence at a trial. A minority of states apply the standard set out in *Frye v. United States*.[264] Federal courts and the majority of states follow the standard set out in *Daubert v. Merrell Dow Pharmaceuticals, Inc*.[265]

Both standards require a method or technique to be reliable in order for evidence derived from it to be admissible, but assess reliability differently.[266] The *Frye* test outsources the reliability determination, requiring that a method be "generally accepted" as reliable by a relevant community of scientists in order to be admissible.[267] The *Frye* test has been widely criticized for not requiring judges to directly assess reliability[268] and because admissibility often turns on who a judge deems to constitute the relevant scientific community.[269]

Under the *Daubert* test, admissibility requires scientific evidence to be relevant and reliable.[270] In *Kumho Tire Company, Limited v. Carmichael*, *Daubert*'s holding was extended beyond scientific evidence to technological evidence.[271] *Daubert* and *Kumho Tire Co.* offer five non-exhaustive factors for evaluating a scientific method or technology's reliability: (1) whether the technology can and has been tested, (2) whether the technology has been subjected to peer review and publication, (3) the known or potential error rate of the technology, (4) whether standards exist that control the field, and (5) whether the technology is generally accepted by the relevant scientific community.[272]

[263] FED. R. EVID. 702.

[264] 293 F. 1013 (D.C. Cir. 1923).

[265] 509 U.S. 579 (1993); *see also* Rochkind v. Stevenson, 236 A.3d 630, 633 (Md. 2020) (noting that "[a] supermajority of states . . . replaced their respective . . . standards with *Daubert*").

[266] *Compare Frye*, 293 F. at 1014 (suggesting that reliability be measured by "general acceptance" in the relevant field), *with Daubert*, 509 U.S. at 593–94 (setting out factors for assessing the reliability of scientific evidence).

[267] *Frye*, 293 F. at 1014.

[268] *See, e.g.*, Sinha, *Junk Science*, *supra* note 62, at 83 (describing the *Frye* approach as "a 'hands off' one that allows judges to avoid meaningfully engaging with whether a given scientific discipline is valid or not").

[269] *See* Roselle L. Wissler, Keelah E.G. Williams & Michael J. Saks, *Dual-Processing Models of Admissibility: How Legal Tests for the Admissibility of Scientific Evidence* Resemble Cognitive Science's *System 1 and System 2*, 17 VA. J.L. & TECH. 354, 357 (2013) (explaining that *Frye* requires "judges to defer to evidence evaluators outside of the court, and the court was to piggyback on their judgments").

[270] *Daubert*, 509 U.S. at 590–91.

[271] Kumho Tire Co. v. Carmichael, 526 U.S. 137, 149 (1999).

[272] *Daubert*, 509 U.S. at 593–94.

The *Daubert* test has since been codified in Federal Rule of Evidence 702 and analogous state laws.[273] Importantly, such rules govern admissibility only at the trial stage, not at pretrial hearings such as suppression hearings.[274]

The *Daubert* test, too, has been exhaustively critiqued.[275] While a full recitation of the criticisms is voluminous and beyond the scope of this Article, a brief overview is helpful. Judges have been criticized for failing to apply the test or applying it superficially[276] and for being more lenient in admitting evidence offered by prosecutors than by the accused.[277] The standard has also been criticized for its malleability.[278] Even when applied correctly, the *Daubert* factors do not all directly measure reliability and while they appear clear on the surface, many are fuzzy.[279] This allows prosecutors and those with career-based, financial, or other interests in admitting evidence to create an appearance that the *Daubert* factors are satisfied even when little substantive evidence of reliability exists.[280]

Despite the criticisms, the *Daubert* test retains some utility. Most obviously, it was devised specifically for assessment of expert evidence. And it can be effective in determining the reliability of technological evidence when applied properly.[281]

Even though these tests were developed for the explicit purpose of assessing the reliability of scientific or technological evidence—and judges have used

[273] FED. R. EVID. 702 advisory committee's note to 2000 amendment; *see also* Sinha, *Junk Science*, *supra* note 62, at 82 n.184 (collecting state rules analogous to Federal Rule of Evidence 702).

[274] FED. R. EVID. 104(a), 1101(d); *see also* United States v. Ozuna, 561 F.3d 728, 736 (7th Cir. 2009) ("We see no persuasive reason to disregard the Rules of Evidence and impose a new requirement on district court judges to conduct a *Daubert* analysis during suppression hearings.").

[275] *See, e.g.*, Jules Epstein, *Preferring the "Wise Man" to Science: The Failure of Courts and Non-Litigation Mechanisms to Demand Validity in Forensic Matching Testimony*, 20 WIDENER L. REV. 81, 84–86 (2014); Jonathan J. Koehler, *Forensics or Fauxrensics? Ascertaining Accuracy in the Forensic Sciences*, 49 ARIZ. ST. L.J. 1369, 1389, 1395 (2017). *See generally* Brandon L. Garrett & M. Chris Fabricant, *The Myth of the Reliability Test*, 86 FORDHAM L. REV. 1559 (2018) (explaining that the *Daubert* test has done little to keep unreliable evidence out of trials); Paul C. Giannelli, *Forensic Science:* Daubert*'s Failure*, 68 CASE W. RSRV. L. REV. 869 (2018) (same).

[276] David L. Faigman, Edward K. Cheng, Jennifer L. Mnookin, Erin E. Murphy, Joseph Sanders & Christopher Slobogin, 1 MOD. SCI. EVIDENCE § 1:30 (2021–2022 ed.) § 1:30 ("[C]ourts have been, at best, lackadaisical and, at worst, disingenuous, in carrying out their gatekeeping duties . . . .").

[277] Stephanie L. Damon-Moore, Note, *Trial Judges and the Forensic Science Problem*, 92 N.Y.U. L. REV. 1532, 1557 (2017).

[278] Sinha, *Radically Reimagining*, *supra* note 62, at 927–37.

[279] *Id.*

[280] *See id.*

[281] Sinha, *Junk Science*, *supra* note 62, at 101.

them to do so at the trial stage for decades—courts typically do not utilize either the *Frye* or *Daubert* tests to analyze policing technology reliability in Fourth Amendment reasonableness determinations.[282] The reasons for this have been under explored until now. As Part III reveals, however, this approach is possibly explained by the lack of clarity in the Supreme Court's guidance for how courts should address reliability of such technology in Fourth Amendment reasonableness determinations and the fact that the *Daubert* test and Rules of Evidence do not apply at suppression hearings. Part III now examines how courts address reliability of policing technology in evaluating the legality of searches and seizures.

III. EVALUATING POLICING TECHNOLOGY RELIABILITY IN FOURTH AMENDMENT REASONABLENESS DETERMINATIONS

To ground the examination of automated Fourth Amendment decision-making, this Part presents the results of a comprehensive original analysis of state and federal opinions. Here we see what courts have done when they addressed—or when litigants asked them to consider—reliability of a policing technology in the context of a reasonable suspicion or probable cause determination. The analysis reveals that Fourth Amendment doctrine has not kept pace with technology-driven policing. Although a hallmark of Fourth Amendment law is that the information provided to or observed by law enforcement to justify a search or seizure must be reliable,[283] the Supreme Court has not announced how—or whether—courts should address the reliability of policing technologies that police increasingly rely upon to support intrusions. Against the backdrop of this doctrinal gap, the analysis reveals that lower courts apply a range of approaches when confronted with questions about policing technology reliability in reasonable suspicion and probable cause determinations. With troubling frequency, however, courts fail to insist on meaningful scrutiny—or exhibit any skepticism about—policing technology reliability.[284]

[282] *See infra* Part III.B.5; United States v. Ozuna, 561 F.3d 728, 736 (7th Cir. 2009) (declining to require district courts to apply the *Daubert* test at suppression hearings).

[283] Illinois v. Gates, 462 U.S. 213, 230 (1983); Alabama v. White, 496 U.S. 325, 328–29 (1990); Florida v. J.L., 529 U.S. 266, 270–72 (2000); Florida v. Harris, 568 U.S. 237, 246–47 (2013); Navarette v. California, 572 U.S. 393, 397 (2014).

[284] *See infra* Part III.B.

A. Research Methodology

To evaluate how courts have addressed the reliability of policing technologies in Fourth Amendment determinations, this study compiles federal and state cases in which courts were asked to address or considered, on their own accord, the reliability of a policing technology in the context of a reasonable suspicion or probable cause determination.[285] A team of research assistants and I cast a wide net and attempted to collect all possible cases decided before July 2023 that might meet the study criteria.[286]

Each opinion was then individually analyzed to determine if it met the study criteria, that is, if the reliability of policing technology was raised or considered as part of a reasonable suspicion or probable cause determination or review of such determination by another court. Cases in which a technology's reliability was neither raised nor considered specifically as a part of the Fourth Amendment determination were eliminated.[287] In the end, the final data set comprised 170 cases.

As with any search, results were limited by the terms used and thus, the initial searches may not have captured every case that met the study criteria.[288] Accordingly, the study is not designed to be a perfect reflection of how courts approach determining policing technology reliability in Fourth Amendment reasonableness determinations. Nevertheless, the final data set presents a detailed and comprehensive picture of how courts have been addressing this issue to date.

Courts in the study frequently declined to address policing technology reliability, even when the accused specifically requested them to do so.[289]

[285] The data set includes trial level and appellate criminal cases, as well as some civil cases and post-conviction collateral appeals substantively considering Fourth Amendment issues (such as Section 1983 or habeas claims). Searches were not limited by publication status. As a result, the data set contains cases reported in official reporters and cases reported only in Westlaw.

[286] Several research assistants conducted the initial searches. Searches were conducted using the Westlaw case search database.

[287] For example, the search captured some cases in which the accused simultaneously moved to suppress evidence alleging a Fourth Amendment violation and also moved to exclude evidence relating to a policing technology as inadmissible at trial under *Daubert* or Federal Rule of Evidence 702. *See, e.g.*, United States v. Brooks, No. 4:11-cr-96, 2012 WL 12895351, at *1–5 (S.D. Iowa Jan. 24, 2012) (describing the defendant's separate motions to suppress and requesting a *Daubert* hearing and analyzing each challenge separately), *aff'd*, 715 F.3d 1069 (8th Cir. 2013). Such cases were removed from the data set.

[288] For example, the initial searches may not have captured cases involving lesser-known technologies.

[289] *E.g.*, United States v. Rickmon, 952 F.3d 876, 879, 879 n.2 (7th Cir. 2020) (declining to reach the question of ShotSpotter's reliability, despite Rickmon's challenge to technology's reliability). *But see* United

Because a central aim of the study is to gain an understanding of how courts address this issue when it is raised, cases in which courts declined to address reliability even when it was squarely presented to them were kept in the data set. If a court did address policing technology reliability, the cases were broadly categorized based on how the court conducted its analysis.

B. Research Findings

The analysis revealed that courts that did address reliability typically took one of four common approaches. First, some courts simply presumed the technology in question was reliable without conducting any reliability analysis whatsoever.[290] Second, several courts determined that a technology was reliable based solely on a law enforcement officer's inexpert assertions.[291] Third, a few courts attempted to apply established frameworks for assessing reliability, such as the framework for assessing reliability of anonymous tips or citizen informants.[292] In many cases where courts attempted to apply such frameworks, they did so without grappling with how such a method might not be suitable when applied to technology.[293] Finally, many courts conducted ad-hoc reliability analyses, relying on factors identified on a case-by-case basis.[294] These findings and initial conclusions that can be drawn from them are further described next.

1. Declining to Address Reliability

In nearly twelve percent of cases in the final data set, courts simply declined to address policing technology reliability even where it was directly challenged by the accused.[295] In declining to address reliability, some courts expressly identified a rationale inconsistent with established Fourth Amendment law. Several courts treated the question of whether reasonable suspicion or probable cause justified a search or seizure as distinct from whether the technology that justified the intrusion was reliable, rather than considering reliability as one

States v. Posado, 57 F.3d 428, 436 (5th Cir. 1995) (applying the *Daubert* standard to evaluate the admissibility of polygraph results at a suppression hearing).

[290] *See infra* Part III.B.2.

[291] *E.g.*, United States v. Hawkins, 37 F.4th 854, 858 n.2 (2d Cir. 2022) (noting the defendant's argument that ShotSpotter is unreliable but finding that it works with a "reasonably high degree of accuracy" based on officer testimony); *see infra* Part III.B.3.

[292] *See infra* Part III.B.4.

[293] *See infra* Part III.B.4.

[294] *See infra* Part III.B.5.

[295] Courts declined to address reliability in 20 out of 170 cases. *E.g.*, Ohio v. Carter, 183 N.E.3d 611, 616, 628–29 (Ohio Ct. App. 2022) (declining to reach the question of ShotSpotter's reliability despite Carter's challenge).

component of a totality of circumstances analysis relevant to evaluating an intrusion's legality.[296] In *United States v. Martin*, for example, a case in which the accused argued that a GPS tracker used to locate him was unreliable and thus could not establish probable cause or reasonable suspicion to stop him, the Eighth Circuit found officers' reliance on the tracker to be reasonable but without conducting any evaluation of the device's reliability.[297]

Courts also conflated the Fourth Amendment reliability requirement with the trial stage reliability requirement. Some appeared to believe that no reliability analysis was required as a part of Fourth Amendment reasonableness determinations because the rules of evidence governing admissibility of expert evidence at trial do not apply to suppression hearings.[298] Indeed, in *Martin*, the Eight Circuit rationalized its decision not to tackle the GPS device's reliability by suggesting that admissibility and reliability are trial-stage questions.[299]

Other courts that declined to address policing technology reliability evaluated the reasonableness of a search or seizure as if the technology played no role in police officers' decision-making. Some decided not to assess policing technology reliability because additional factors supported reasonable suspicion or probable cause after setting aside the output of the policing technology.[300] In other words, these courts failed to acknowledge that exculpatory evidence—such as the unreliability of technology used to justify a search or seizure decision—could weigh against a finding of reasonableness.[301]

[296] *See, e.g.*, *Carter*, 183 N.E.3d at 628 ("[T]he issue herein is whether [the officers] had reasonable suspicion to stop Carter and conduct the pat down, and we need not reach the scientific reliability of the ShotSpotter system."); United States v. Martin, 15 F.4th 878, 882 (8th Cir. 2021) (finding officers' reliance on GPS tracker to conduct a stop "reasonable" without addressing the device's reliability); *see also* United States v. Robertson, 39 F.3d 891, 894 (8th Cir. 1994) (describing argument that Forward Looking Infrared Device (FLIR) was not accurate as "essentially irrelevant to the probable cause inquiry" and finding officers' reliance on FLIR reading to conclude that marijuana was present inside the accused's home was a reasonable inference without analyzing the FLIR's reliability).

[297] *Martin*, 15 F.4th at 881–82.

[298] *See Carter*, 183 N.E.3d at 629 (explaining that both parties agreed that rules of evidence do not apply to suppression hearings in declining to reach the question of ShotSpotter's reliability).

[299] *Martin*, 15 F.4th at 882.

[300] *See, e.g.*, United States v. Charles, No. CR.A. 03-15-SLR, 2003 WL 21730639 (D. Del. July 23, 2003) (finding it unnecessary to address the reliability of radar equipment used to check accused's speed because officer observations gave reasonable suspicion for a stop).

[301] *Compare id.*, *with* Kuehl v. Burtis, 173 F.3d. 646, 650 (8th Cir. 1999) (stating that the "Fourth Amendment requires that we analyze the weight of all the evidence," including exculpatory evidence in a totality of circumstances analysis).

2. *Presuming Reliability*

In just over a quarter of cases, courts simply presumed the reliability of the policing technology at issue without analysis.[302] At first blush, it may seem that there is little distinction between cases where courts decline to assess policing technology reliability entirely and those where reliability is assumed. But courts that chose not to assess reliability took an approach arguably in greater tension with the Supreme Court's repeated pronouncement that information supporting reasonable suspicion and probable cause must be reliable than courts that presumed a technology's reliability. Unlike in cases where courts declined to address reliability at all, courts in this set of cases at least recognized that policing technology reliability was relevant to the legality of a search or seizure that police justified using such technology. Nevertheless, rather than scrutinizing the technology at issue, in these cases, courts simply asserted that a policing technology was reliable without offering an explanation for that finding.[303]

3. *Adopting Officers' Inexpert Reliability Determinations*

In over ten percent of cases comprising the final data set, courts determined that a policing technology was reliable based solely, or near-solely, on a law enforcement agent's conclusion about the technology's reliability.[304] Officer conclusions about reliability, however, were consistently superficial, typically based on their own or other officers' field observations of accuracy.[305]

Courts in this set of cases did not appear to recognize that officer claims are not accurate reliability measures. As Justice Kagan explained in *Florida v. Harris*, reliability cannot generally be measured by field performance because, in the typical case, officers do not know ground truth.[306] Assessing reliability of

[302] We observed that courts presumed reliability in forty-three cases in the data set. *See, e.g.*, United States v. King, 439 F. Supp. 3d 1051, 1055, 1055 n.2 (N.D. Ill. 2020) (assuming ShotSpotter to be reliable in finding that a human tip that contradicted ShotSpotter was insufficiently reliable to support a stop).

[303] *See, e.g.*, *id.*

[304] We observed that courts based their reliability determination on police officer assertions in 18 opinions in the 170-case data set. *E.g.*, State v. Bellamy, No. A-2978-16T2, 2018 WL 2925724, at *8–9 (N.J. Super. June 12, 2018) (finding ShotSpotter reliable based on officer's claim (1) of familiarity with the system; (2) that ShotSpotter "identifies and pinpoints" gunfire; and (3) that he has never responded to an alert "that was proven inaccurate").

[305] *See supra* note 304.

[306] *See supra* notes 255–57 and accompanying text; JILLIAN B. CARR & JENNIFER L. DOLEAC, THE GEOGRAPHY, INCIDENCE, AND UNDERREPORTING OF GUN VIOLENCE: NEW EVIDENCE USING SHOTSPOTTER DATA 4–5 (2016), https://www.brookings.edu/wp-content/uploads/2016/07/Carr_Doleac_gunfire_underreporting.pdf ("[I]t is typically impossible to distinguish false positives from gunshots that cannot be corroborated by other evidence . . . .").

field data overvalues observed true positives without sufficiently incorporating the effect of false positives and false negatives on reliability. False positives and false negatives in field data may never be uncovered, which means that relying on field performance to assess reliability can create significant misimpressions about a technology's accuracy, consistency, and error rate.[307]

Officers may well believe the technology they rely on is reliable based on field experience or information provided by other officers. Even so, the average officer, whose primary responsibilities relate to on-the-ground policing, is not qualified to assess the validity of complex technologies.[308] Police officers are not typically trained in how policing technologies work, are not involved in the development of such technologies, and do not participate in assessments of such technologies. Even if they were capable of doing so, officers acting in the moment typically do not engage in sophisticated analyses of the reliability of policing technologies before initiating a seizure or search; they often assume it.[309]

Courts' reliance on such unscientific assessments, thus, stretches the requirement that Fourth Amendment intrusions be supported by reliable information where suspicion is substantially based on policing technology. Courts in the study nevertheless frequently failed to scrutinize the bases of officer observations about policing technology or whether such observations could reasonably establish reliability.[310] Courts did not ask for validation testing data, error rate calculations, scientific literature, or other information that could be used to vet officers' claims of accuracy.[311]

4. *Applying Existing Reliability Frameworks*

A small fraction of courts, about five percent, applied existing frameworks for assessing reliability. Most courts in this group used frameworks developed to assess the reliability of non-technological information to analyze policing

[307] *See* Sinha, *Automated Gunshot Detection*, *supra* note 34, at 79–80 (explaining why field performance cannot establish reliability).

[308] *See* GARVIE, *supra* note 76, at 26 (quoting an officer explaining that "[t]here is no specific training or certification to use the facial recognition database," a type of technology used by officers).

[309] *See* Elizabeth E. Joh, *The Unexpected Consequences of Automation in Policing*, 75 SMU L. REV. 507, 513 (2022) [hereinafter Joh, *Unexpected Consequences of Automation*] (describing how officers take "cognitive shortcuts" to rely on the outputs of policing technologies to conduct stops even under circumstances in which such reliance is unjustified).

[310] State v. Bellamy, No. A-2978-16T2, 2018 WL 2925724, at *8–9 (N.J. Super. June 12, 2018) (taking officer's claims about ShotSpotter's accuracy at face value).

[311] *See, e.g.*, *id.*

technology reliability. For example, many courts analogized policing technologies to third-party tipsters.[312] This was not a surprising result, given that some technologies provide similar types of information as human tipsters and the Supreme Court has laid out frameworks for assessing reliability of third party tipsters.[313] In accordance with such frameworks, courts in this group often looked for evidence corroborating the policing technology to support reasonable suspicion or probable cause.[314] In several cases in which courts attempted to adapt analyses designed for tipsters to policing technology, however, they failed to address how such analysis may fail when applied to technology.[315]

Only one court used either the *Daubert* or *Frye* tests or analogous trial-stage admissibility frameworks to assess policing technology reliability.[316] This was a somewhat surprising result given that such frameworks are designed to assist courts in assessing the reliability of expert evidence, albeit at the trial stage. While courts are not required to apply rules of evidence at suppression hearings, nothing prevents judges from applying such tests to assess reliability at that stage.[317] Moreover, because courts are familiar with it, it would seem reasonable for them to consider using this approach when the substantive question before them—the reliability of technological evidence—is the same.

[312] Nine of 170 courts attempted to apply existing Fourth Amendment reliability frameworks to policing technologies. *E.g.*, United States v. Vallo, 608 F. Supp. 3d 1071, 1078–79 (D.N.M. 2022) (analogizing ShotSpotter to an anonymous tip), *appeal dismissed*, No. 22-2097, 2022 WL 18781016 (10th Cir. Aug. 29, 2022); *In re* Matter of Search of Multiple Email Accts., 585 F. Supp. 3d 1, 20 (D.D.C. 2022) (recognizing the importance of assessing clustering software's reliability to probable cause determination and analogizing to confidential informant); State v. Gaddy, 93 P.3d 872, 876–77 (Wash. 2004) (analogizing state records database to a citizen informant); State v. Henz, 514 P.3d 1, 8 (N.M. Ct. App. 2022) (finding that Tumblr and Google "functioned similarly to an identified citizen informant" by providing the National Center for Missing and Exploited Children with reports that users had posted child pornography to their platforms).

[313] *See supra* Part II.A.1.

[314] *See, e.g.*, *Vallo*, 608 F. Supp. 3d at 1078–79 (finding that even if a ShotSpotter alert for gunfire were reliable, some corroboration of the alert—which was not present—would be necessary to establish reasonable suspicion).

[315] *See, e.g.*, State v. Sowin, 2020 WI App 70, ¶¶ 10–12 (Wis. Ct. App. 2020) (reasoning that a cyber tipster is akin to a citizen informant and thus deserving of a presumption of reliability without examining the differences between human and electronic tipsters).

[316] State v. Smith, 130 Wash. 2d 215, 222 (1996). Another court applied elements of the *Daubert* test, but not the entire framework, to evaluate whether a policing technology was sufficiently reliable to establish probable cause. United States v. Sigouin, 494 F. Supp. 3d 1252, 1268 (S.D. Fla. 2019) (factoring software's testing, false positive error rate, and one peer-reviewed article into its determination of whether a software was sufficiently reliable to support an affidavit for a search warrant).

[317] *See* United States v. Ozuna, 561 F.3d 728, 737 (7th Cir. 2009) (explaining that judges may, but are not required to, conduct a *Daubert* analysis at a suppression hearing).

5. *Ad-Hoc Reliability Analysis*

Nearly half of courts in the study analyzed policing technology reliability using ad-hoc, case-specific factors.[318] In many cases, courts' approaches were soundly reasoned. For example, relying on Supreme Court precedent emphasizing corroboration as a key determinant of reliability, some courts weighed whether police officers were able to independently verify the output of a technology heavily in reliability determinations.[319]

As implied by the category designation, courts' approaches varied considerably. Judges evaluated a variety of factors, including police accuracy assessments, scientific literature, assessments of testing, and error rates.[320] A result of the variability in approaches was that some courts applied reliability criteria that may be in tension with each other. For example, in two cases in which the reliability of a database was at issue, the Ninth Circuit found that reliability may be presumed unless the database is subject to "systemic errors," while the Tenth Circuit found reliance on a database is reasonable unless there is "reason to worry" about the database's reliability.[321] The level of suspicion at issue was different in each case, with the Ninth Circuit considering whether probable cause existed and the Tenth Circuit conducting a reasonable suspicion test.[322] Though greater reliability is required to establish probable cause than

[318] We observed this approach in 80, or 47%, of 170 cases. *E.g.*, State v. Police, 273 A.3d 211, 227–30 (Conn. 2022) (relying on applicable precedent as well as scientific and legal literature to find that DNA analysis of complex DNA mixture that purported to identify a suspect by DNA profile was insufficiently reliable to satisfy the Fourth Amendment's particularity requirement and "John Doe" warrant based on such DNA analysis was void); Commonwealth v. Ford, 182 N.E.3d 1013, 1018 (Mass. App. Ct. 2022) (finding ShotSpotter alert sufficiently reliable to support reasonable suspicion because the timing and location of subsequent alerts "indicated a specific linear trajectory" and because alert was corroborated by officer observation).

[319] *E.g.*, United States v. Pipes, 909 F. Supp. 689, 694 n.7 (D. Neb. 1995) (determining that there were "no serious questions" raised about the reliability of a computer program used to determine vehicle speed, partly because the speed calculation was confirmed by the officer's speedometer reading and observation), *aff'd*, 125 F.3d 638 (8th Cir. 1997); United States v. Thomas, 788 F.3d 345, 353 (2d Cir. 2015) (finding that software for detecting child pornography exhibited "sufficient indicia of reliability" based in part on law enforcement's corroboration of information received); United States v. Collins, 753 F. Supp. 2d 804, 811, 813 (S.D. Iowa 2009) (finding Peer Spectre, an automated software program used to trace Internet Protocol (IP) addresses that may be sharing contraband files, sufficient to establish probable cause where it was corroborated by other evidence).

[320] *E.g.*, *Sigouin*, 494 F. Supp. 3d at 1267–68 (finding software used to detect child pornography reliable based on widespread law enforcement use, review of peer-reviewed article containing summary of software's testing and error rate, and corroboration).

[321] *Compare* Gonzalez v. United States Immigr. & Customs Enf't, 975 F.3d 788, 822 (9th Cir. 2020) (officers may rely on databases to support probable cause unless the database is subject to systemic errors), *with* United States v. Esquivel-Rios, 725 F.3d 1231, 1235–36 (10th Cir. 2013) (declaring that officers may rely on databases to support reasonable suspicion unless there is "reason to worry" about the database's reliability).

[322] *Gonzalez*, 975 F.3d at 822; *Esquivel-Rios*, 725 F.3d at 1238.

reasonable suspicion,[323] it is not clear that the Tenth Circuit's reliability test is less onerous than the Ninth Circuit's.

C. Research Limitations

This study appears to be the first deliberate attempt to convey a snapshot of how courts address the reliability of policing technologies in Fourth Amendment determinations. A natural consequence is that the analysis comes with unavoidable limitations. First, the data set only includes cases that are reported or otherwise available on Westlaw, which has varied coverage across courts.[324] Cases that are not available on Westlaw were not included. As a result, the data set does not reflect how all courts have evaluated policing technology reliability in the context of Fourth Amendment reasonable suspicion and probable cause determinations.

A second unavoidable constraint is that there is no obvious or straightforward way to identify cases in which policing technology was *used* but never *raised* in litigation. As such, the data set is limited to cases in which courts were affirmatively asked to address or considered, on their own accord, the reliability of a policing technology in the context of a reasonable suspicion or probable cause determination. Consequently, it does not capture the substantial number of cases in which policing technology was used to justify a search or seizure decision, but that technology's reliability was neither challenged nor considered by a judge evaluating the legality of an intrusion.[325] Two takeaways of this limitation are that, notwithstanding the Supreme Court's pronouncement that reasonable suspicion and probable cause be based on reliable information, (1) litigants frequently fail to challenge policing technology reliability, perhaps

[323] Alabama v. White, 496 U.S. 325, 330 (1990) ("[R]easonable suspicion can arise from information that is less reliable than that required to show probable cause.").

[324] A substantial number of cases, particularly state trial court rulings, do not make it into conventional case search engines like Westlaw. *See, e.g.*, *Case Law Research*, UNLV WIENER ROGERS L. LIB., https://law-unlv.libguides.com/caselaw/published-versus-unpublished ("Most states, including Nevada, do not publish state trial court cases"). This is for a variety of reasons, including that judges often deliver oral rulings without accompanying written opinions, and because database coverage varies across courts. https://1.next.westlaw.com/Browse/Home/Cases/FederalCases?transitionType=Default&contextData=(sc.Default) (choose the "(I)" (Information) icon).

[325] For example, in *United States v. Lopez-Navalles*, border patrol agents relied in part on sensor technology that purports to detect when something weighing over forty pounds crosses the border as justification for conducting a stop of Adan Lopez-Navalles's vehicle. No. 90-10094, 1991 WL 67170, at *3 (9th Cir. May 1, 1991). Although the agents relied on the sensor to establish reasonable suspicion for the stop, Lopez-Navalles did not challenge the sensor technology's reliability in his Fourth Amendment challenge—although he did challenge its reliability for purposes of admissibility at trial—and the court did not address it independently. *See id.* at *3, *3 n.1.

because they are not always aware that technology has been used,[326] and (2) judges frequently fail to address the reliability of policing technology or assume its reliability to a much greater extent than represented in this analysis. This limitation itself highlights the importance of the snapshot the dataset does provide, as well as the need for more research on police reliance on technology for developing suspicion and courts' review of such reliance.

Third, ambiguity in courts' analyses complicated the effort to systematically label and categorize their approaches. Courts were frequently not explicit in describing how they evaluated reliability. There is also some overlap between categories, and in some cases, courts took an approach that had features of multiple categories.[327] For example, some courts presumed reliability of a policing technology, but appeared to do so based on a law enforcement claim that the technology was accurate.[328] Other courts did not explicitly address reliability, but asserted that police reliance on the technology at issue was reasonable, implying that the technology was reliable to at least some degree.[329] In many such cases, courts did not analyze how a lack of reliability might influence the totality of circumstances analysis by undercutting other evidence supporting reasonable suspicion or probable cause.[330] For purposes of labeling and categorizing, instances with ambiguity necessitated making judgment calls about which reasonable minds might differ. In such instances, the category that most substantially aligns with the court's reasoning was selected. Part IV considers the implications of the results and a path forward.

[326] Ram, *supra* note 111, at 666–82 (describing efforts to prevent disclosure of police use of various technologies to the defense); *see also, e.g.*, Nicky Wolf, *2,000 Cases May Be Overturned Because Police Used Secret Stingray Surveillance*, GUARDIAN (Sept. 4, 2015), https://www.theguardian.com/us-news/2015/sep/04/baltimore-cases-overturned-police-secret-stingray-surveillance (describing alleged collusion between prosecutors and police to withhold disclosures regarding police use of Stingray tracking devices in thousands of cases).

[327] *See, e.g.*, United States v. Rickmon, 952 F.3d 876, 879 n.2, 882 (7th Cir. 2020) (explaining that the court "need not reach the reliability of ShotSpotter," but analogizing ShotSpotter to an anonymous tipster in analyzing reasonable suspicion).

[328] *See, e.g.*, State v. Bellamy, No. A-2978-16T2, 2018 WL 2925724, at *4 (N.J. Super. June 12, 2018) (asserting that ShotSpotter is "objectively more reliable than an anonymous report" based in part on detective's "familiar[ity] with the ShotSpotter system" and claim that he "has never responded to a ShotSpotter report of gunfire that was proven inaccurate").

[329] *See, e.g.*, Ohio v. Carter, 183 N.E.3d 611, 628–29 (Ohio Ct. App. 2022) (concluding, without reaching the question of ShotSpotter's reliability, that officers had reasonable suspicion to conduct a stop in part based on a ShotSpotter alert).

[330] *See, e.g.*, *id.*

IV. IMPLICATIONS AND A PATH FORWARD

The above analysis reveals that lower courts are often unable or unwilling to incorporate the Supreme Court's pronouncement that privacy intrusions must be supported by reliable information when policing technology provides such information. This Part first considers the implications of this finding. It then proposes a new normative framework for addressing reliability of technology in Fourth Amendment reasonableness determinations. Drawing on the Supreme Court's existing Fourth Amendment reliability case law, especially its decision in *Florida v. Harris*, it recommends that courts evaluate external, independent, and disinterested evidence of policing technology reliability in probable cause and reasonable suspicion determinations. Where no such evidence exists, it recommends courts apply the test first established in *Daubert* and extended to technology by *Kumho Tire Co.* to evaluate policing technology reliability. Because many cases in which policing technology justifies a search or seizure do not result in formal charges or criminal litigation, this Part also considers targeted policy interventions that may alleviate harms of increasing police reliance on technology.

A. Implications

1. Doctrinal Implications

Taking the sum-total of the study results together reveals a clear big-picture conclusion: at best, many courts paid lip service to the idea that reasonable suspicion and probable cause need be supported by reliable information; at worst, their approaches are out of step with the Fourth Amendment.

The five broad approaches that courts have taken when confronted with questions about policing technology reliability in Fourth Amendment reasonableness determinations can be broadly grouped into two sets. A sizeable percentage of courts approached policing technology reliability in a way that conflicts with, or is inattentive to, the Supreme Court's consistent pronouncement that reasonable suspicion and probable cause must be based on reliable information. Nearly forty-eight percent of courts in the study declined to address a technology's reliability, presumed it without analysis, or found it to be reliable based merely on a law enforcement officer's inexpert assertion, without conducting additional, independent reliability analysis. A similar percentage of courts—roughly fifty-two percent—did attempt to address reliability meaningfully. But these courts did so in inconsistent ways, and often, in ways poorly suited to addressing technological reliability.

What is less clear is *why* courts are reticent to substantively address policing technology reliability in the face of consistent precedent that seems to dictate a contrary approach. The study findings, however, suggest some possibilities. First, courts may simply not understand that because reliability is relevant to reasonable suspicion and probable cause determinations, policing technology reliability should be addressed. In other words, courts may not understand when a reliability analysis is required. Indeed, the study findings support this conclusion.[331] Some courts seemed to take the concerning view that whether an intrusion was reasonable turned entirely on what officers knew at the time regardless of a technology's reliability.[332] One appellate court succinctly clarified the flaw in such reasoning, explaining, in the context of a stop based upon reasonable suspicion, that it "conflates two different principles of reasonable suspicion. It is true that reasonable suspicion is determined by the totality of the circumstances known to the officers at the time they seized the defendant . . . However, . . . [t]he reliability analysis examines the *source* of the information that was relayed to police, not what the officers personally observed or knew."[333] Courts in the study routinely failed to make this precise distinction.[334]

Courts may also be reluctant to conduct in-depth reliability analyses. Trial stage admissibility hearings can be time-consuming and resource intensive. They frequently last several days, weeks, or longer and involve expensive and prolonged expert testimony from witnesses called by the parties.[335] In contrast,

[331] *See supra* note 298 (discussing *Carter*, 183 N.E.3d at 629).

[332] *See, e.g.*, Molina ex rel. Molina v. Cooper, 325 F.3d 963, 971 (7th Cir. 2003) (finding it reasonable for officers to rely on positive field test for controlled substances that turned out to be false because there was no reason for officers to know that field tests were unreliable). In other instances, courts took the related view that even where technology was unreliable, intrusions would be justified under the good faith exception. *See, e.g.*, United States v. Carter, No. 0:20-000352020, 2020 WL 6136480, at *8 (D. Minn. Sept. 18, 2020) (determining that police reliance on a warrant would fall within good faith exception even if ion scanner results, which were cited as supporting probable cause in the warrant affidavit, were not reliable). Of course, reliance on the good faith exception in any given case might be undermined by a *prior* judicial determination that the policing technology in question lacks reliability. *See* United States v. Chatrie, 590 F. Supp. 3d 901, 941 (E.D. Va. 2022) (finding that while good faith exception applied to overbroad geofence warrant, it "may not carry the day in the future").

[333] People v. Jones, 220 N.E.3d 475, 489–90 (Ill. App. 2023) (emphasis in original) (citations omitted).

[334] *See, e.g.*, *Cooper*, 325 F.3d at 971.

[335] *See, e.g*., United States v. Gissantaner, 417 F. Supp. 3d 857, 860 (W.D. Mich. 2019) (describing *Daubert* hearing concerning admissibility of DNA evidence that involved several witnesses, voluminous briefing, and several days of testimony over the course of more than a year), *rev'd*, 990 F.3d 457 (6th Cir. 2021); United States v. Tibbs, No. 2016-CF1-19431, 2019 WL 4359486, at *1 (D.C. Super. Ct. Sept. 5, 2019) (describing "an extensive evidentiary hearing" on the admissibility of firearms and tool mark identification evidence "that involved detailed testimony from a number of distinguished expert witnesses, review of all of the leading studies in the discipline, pre- and post-hearing briefing, and lengthy arguments by . . . counsel").

suppression hearings typically take just a few hours or less and are often held immediately prior to the start of a trial.[336] Judges may recognize that a comprehensive reliability inquiry has the potential to slow down a single criminal prosecution substantially and have a domino effect on others.

Finally, courts may simply not understand how to conduct a reliability analysis when policing technology is at issue. Judges may be uncomfortable with or incapable of serious analysis of the technical aspects of complicated policing technologies. Many judges do not have the technical know-how to engage with scientific reliability or, possibly, may not be sufficiently well-versed in scientific concepts to even recognize that establishing reliability—and thus, evaluating it—requires a different approach than does assessing reliability of more traditional information.[337]

Courts' laxity in addressing policing technology reliability in assessing the legality of searches and seizures has significant doctrinal implications. By allowing police to predicate seizures and searches on the outputs of technology they decline to scrutinize, judges allow police to take actions that water down privacy protections and circumvent the Fourth Amendment's requirement that suspicion be particularized as to crimes and persons. As was true in Bobby Jones's case, policing technology that directs police to a purportedly suspicious person encourages police to assume that crime is occurring, even if it is not.[338] Of course, this is also true in non-technology driven contexts, as when a person provides police with information about a potential suspect. But in such cases, it is expected that courts vet reliability of such information sources according to established precedent. In cases where technology is the information source and precedent does not provide clear guidance, reliability is not being analyzed to the same extent as human informants.

Reciprocally, as in the case of the man described in the opening vignettes as walking beside a building in Chicago, technology that pinpoints locations of purported suspicious activity dilutes the individualized suspicion requirement.[339] In cases where policing technology has directed police to a specific location, courts have treated a person's presence in the vicinity of that location as sufficient to establish reasonable suspicion for a stop, even where

[336] *See* LAFAVE, *supra* note 31, § 11.2(d).

[337] *See* Damon-Moore, *supra* note 277, at 1536 n.22 (explaining that judges "commonly lack scientific training and may struggle to rigorously assess reliability of expert evidence as a result").

[338] *See supra* notes 16–25, 41, 48 and accompanying text.

[339] *See supra* notes 9–15, 40, 45–47 and accompanying text.

little else, such as individual features, characteristics, or conduct, connected the person to the alert.[340] In *United States v. Martin*, the Eighth Circuit upheld a stop that took place after a GPS tracker directed police to the accused's location, but declined to assess the reliability of the tracking device as requested by the defense.[341] Aside from Martin's location in the vicinity of the alleged crime, which would not have been known to police without the GPS device, there was little to suggest that police would have had reasonable suspicion to conduct the stop.[342] Although the court suggested that Martin and the vehicle he was found in matched descriptions provided by witnesses, closer analysis of the facts reveals this to be inaccurate.[343] Given the misalignment between the suspect description and both Martin and his vehicle's appearances, without the GPS tracking, the reliability of which was never assessed, police may not have had reasonable suspicion to stop Martin.

In *United States v. Rickmon*, the Seventh Circuit found that reasonable suspicion to stop Terrill Rickmon existed based on his presence near the alleged location of a ShotSpotter alert.[344] The officer who stopped Rickmon's vehicle admitted that his only reason for conducting the stop was its proximity to the alleged location of gunfire, not because he had reason to believe the vehicle's occupants were responsible for gunfire or were otherwise acting suspiciously.[345]

2. *Non-Doctrinal Implications*

It might appear that courts' failure to evaluate policing technology reliability when determining the lawfulness of Fourth Amendment intrusions has purely doctrinal implications. But what begins as a doctrinal problem has far-reaching real-world consequences. Courts' lackluster approach to vetting policing

[340] *See, e.g.*, United States v. Rickmon, 952 F.3d 876, 886 (7th Cir. 2020) (Wood, C.J., dissenting) (disagreeing with majority's finding of reasonable suspicion based in part on a ShotSpotter alert because: "The only thing that distinguished the car [the officer] chose to stop was that it existed, and it was the only car on the street at that early hour of the morning. None of the information he had received even hinted at the shooter's car's make, color, age, style, or anything else.").

[341] 15 F.4th 878 (8th Cir. 2021), *cert. denied*, 142 S. Ct. 1432 (2022).

[342] *See id.* at 880–82.

[343] The suspect's vehicle was described as a dark green coupe (a two-door vehicle), but Martin was observed in a dark blue sedan (a four-door vehicle); Martin and the vehicle's driver were described as acting "unusual," but what police believed to be "unusual" was only their failure to react to a heavy police presence; and the court gave no indication that Martin matched the detailed suspect description aside from his being Black. *Id.* at 880, 882.

[344] *Rickmon*, 952 F.3d at 879–80, 88.

[345] *Id.* at 880; *see also id.* at 886 (Woods, C.J., dissenting) ("[The officer] frankly admitted that he would have stopped literally any car he saw . . . .").

technology reliability has the potential to influence—and, in many cases, is already influencing—policing practices.[346] Judges' failure to vet policing technology when analyzing Fourth Amendment questions may incentivize ever greater reliance on policing technology, including technologies that lack scientific reliability.[347]

Judges vetting the basis of Fourth Amendment intrusions serves as a check on police conduct.[348] But where there is no such vetting, there can be no corollary disincentive to unlawful intrusions.[349] Without scrutiny, police are left to use policing technologies as they please. As partisan actors, their tendency will naturally be toward extending and expanding uses of policing technology.[350]

This is a consequence that raises corollary concerns for communities that already suffer over-policing. Policing technologies are disproportionately deployed against communities of color.[351] ShotSpotter is disproportionately installed in predominantly Black, Latine, and other minority communities.[352] GPS monitors are frequently worn by individuals either on pretrial release or post-conviction supervision.[353] As with other aspects of the criminal legal system, Black and Latine people and members of other marginalized communities are over represented in the population of those under GPS

[346] *See* Utah v. Strieff, 579 U.S. 232, 245–46 (2016) (Sotomayor, J., dissenting) ("When courts admit only lawfully obtained evidence, they encourage 'those who formulate law enforcement polices, and the officers who implement them, to incorporate Fourth Amendment ideals into their value system.' But when courts admit illegally obtained evidence as well, they reward 'manifest neglect if not an open defiance of the prohibitions of the Constitution.'" (internal citations omitted)).

[347] *See id.*

[348] *See id.* at 249 ("[T]he exclusionary rule gives [police officers] an 'incentive to err on the side of constitutional behavior.'" (internal citation omitted)).

[349] *See* Lego v. Twomey, 404 U.S. 477, 489 (1972) ("[T]he exclusionary rules are very much aimed at deterring lawless conduct by police and prosecution."); *see also Strieff*, 579 U.S. at 245 (Sotomayor, J., dissenting) ("The exclusionary rule 'removes an incentive for officers to search us without proper justification.'" (internal citation omitted)).

[350] *See* Joh, *Unexpected Consequences of Automation*, *supra* note 309, at 507, 526 (describing how automation "[led] to changes in police behavior" and expanded uses of ShotSpotter beyond those recommended by the company that produces it); *see also* Johnson v. United States, 333 U.S. 10, 13–14 (1948) (describing police officers as "engaged in the often competitive enterprise of ferreting out crime"); United States v. Jones, 565 U.S. 400, 416 (2012) (Sotomayor, J., concurring) (commenting that "unrestrained [government] power to assemble data . . . is susceptible to abuse").

[351] Vincent M. Southerland, *The Master's Tools and a Mission: Using Community Control and Oversight Laws to Resist and Abolish Police Surveillance Technologies*, 70 UCLA L. Rev. 2, 18–23 (2023) (explaining that policing technologies have historically been used for surveillance and control of communities of color); *see also* Leaders of a Beautiful Struggle v. Baltimore Police Dep't, 2 F.4th 330, 347 (4th Cir. 2021) (acknowledging over-surveillance and over-policing of Black and other "disadvantaged" communities).

[352] Sinha, *Automated Gunshot Detection*, *supra* note 34, at 87.

[353] Arnett, *supra* note 141, at 644, 672.

surveillance.[354] Predictive policing tools rely on data that is a product of over-policing of marginalized communities, and thus, will reproduce already biased policing practices.[355]

While individuals and communities may feel the effects of policing technology, they will not always be well-positioned to impose restrictions on or curtail its use. Practical roadblocks inhibit civilians from voicing concerns. They may not have access to the information necessary to expose such technologies' flaws.[356] Data is very frequently not publicized; policing technologies are often secretly purchased and deployed with little or no oversight by local legislatures or input from non-law enforcement parties, including civilians whom such technology targets.[357] And, there are not always clear regulatory structures that allow for notice and comment or other feedback mechanisms.[358]

Police use of unvetted technology can also add to distrust between law enforcement and community members.[359] While members of disadvantaged communities may intuitively be aware that policing technologies are flawed and disproportionately applied against them, technology also has a "tech-washing" effect, creating an appearance of fairness and neutrality.[360] As Professor Monica Bell explains, "The apparent neutrality of most modern laws and policies means that even those who are disadvantaged under them might not fully perceive them as discriminatory."[361] Individuals may not feel capable of voicing their concerns,

[354] Weisburd, *Punitive Surveillance*, *supra* note 140, at 155; Weisburd, *Sentenced to Surveillance*, *supra* note 141, at 759.

[355] *See* LEE ET AL., *supra* note 57, at 7–8 ("[U]se of predictive algorithms in place-based crime forecasting produced harmful, self-perpetuating feedback loops of crime predictions, in which officers would repeatedly patrol neighborhoods that had been disproportionately targeted by law enforcement in the past . . . .").

[356] *See, e.g.*, Manes, *supra* note 187; Justin Fenton, *Baltimore Police Used Secret Technology to Track Cellphones in Thousands of Cases*, BALT. SUN (Apr. 9, 2015, 6:52 AM), https://www.baltimoresun.com/maryland/baltimore-city/bs-md-ci-stingray-case-20150408-story.html (explaining Baltimore Police Department's shielding of surveillance data).

[357] LEE ET AL., *supra* note 57, at 9.

[358] *See, e.g.*, Tom Schuba, *Activists Slam City for Extending ShotSpotter Contract Amid Mounting Criticism of the Gunshot Detection System*, CHI. SUN TIMES (Aug. 19, 2021), https://chicago.suntimes.com/crime/2021/8/19/22633412/activists-slam-city-shotspotter-contract-gunshot-detection-system-policing (describing community complaints about inability to provide comments regarding ShotSpotter).

[359] *See* Meares, *supra* note 60, at 164–65.

[360] LEE ET AL., *supra* note 57; *see also* Arnett, *supra* note 119, at 1411–12 (describing deeply-held societal belief in technology's value that allows it to be touted as fair and neutral).

[361] Monica C. Bell, *Police Reform and the Dismantling of Legal Estrangement*, 126 YALE L.J. 2054, 2115 (2017). The disparate application of policing technology to certain communities but not others can be conceived of as one aspect of "structural exclusion," the way in which neutral-seeming policies are distributed to the exclusion of disadvantaged population. *See id.* Structural exclusion, in turn, is one component of what Bell

either because of the difficulty of piercing the tech veil of neutrality or because there is no forum for providing their opinions. The gap between individuals' intuitive knowledge and ability to remedy concerns about police deployment of technology against them may exacerbate feelings of societal alienation and mistrust of law enforcement.[362]

Judicial failure to scrutinize policing technology reliability when evaluating the legality of searches and seizures thus sends law enforcement the message that over-policing enabled by technology can occur without consequence.

B. Toward Doctrinal Clarity

To prevent further dilution of the Fourth Amendment's protections and to mitigate the harms described above, courts must begin to consistently scrutinize the reliability of policing technology used to justify searches and seizures. This section proposes a new doctrinal framework to allow courts to conduct that analysis. The framework draws from and extends the Supreme Court's existing reliability precedent to map out a straightforward means for assessing policing technology reliability in evaluations of search and seizure legality. Under this framework, courts would look for external, independent, and disinterested evidence of a policing technology's reliability in probable cause and reasonable suspicion determinations. Where none exists, courts should turn to the trial-stage reliability test applicable to technology: the test laid out in *Daubert*.[363]

Examining its decisions that address the need for reasonable suspicion and probable cause to be supported by reliable information makes clear that the Court consistently emphasizes that external, independent evidence can establish reliability—if not so directly.[364] While most of those cases analyze reliability in the context of third-party tipsters and correspond poorly to policing technology,[365] the Supreme Court has considered reliability in one case in which suspicion was supplied by something approximating a black box, a drug-sniffing dog.[366] As previously described, in that case, *Florida v. Harris*, the Court looked

describes as "legal estrangement," the idea that law and its enforcement operate to exclude communities of color and other disadvantaged communities from society. *Id.* at 2067–68.

[362] *See id.* at 2100 (theorizing disadvantaged communities' distrust of law enforcement as "legal estrangement" resulting from exclusion); *see also id.* at 2054 ("Legal estrangement is a theory of detachment and eventual alienation from the law's enforcers, that reflects the intuition among many people in poor communities of color that the law operates to exclude them from society.").

[363] 509 U.S. 579, 592–94 (1993).

[364] *See supra* Part II.A.1.

[365] *See supra* notes 245–52 and accompanying text.

[366] *See generally* Florida v. Harris, 568 U.S. 237 (2013).

to a certifying body for evidence of the dog's reliability.[367] While the Court's reasoning was flawed, the core of its logic can be distilled and adapted to policing technology.

Without saying so explicitly, *Harris* recognizes that finding a black box reliable requires external and independent evidence supporting reliability.[368] A drug-sniffing dog cannot explain its decision-making, nor can a judge assess it, necessitating a reliability proxy that is trustworthy on its own. Accordingly, the Court reasoned that an appropriate reliability measure should focus on a drug-sniffing dog's performance in a controlled environment—"standard training and certification settings"—where ground truth is known, and a dog's performance can be assessed against it.[369]

External evidence of reliability. The Court in *Harris* recognized that, in a black box context, evidence of reliability cannot be supplied by the black box itself. Absent additional information, like the output of a policing technology, a drug dog's output—a positive alert or failure to alert—provides little information about its accuracy. For reasons already described, the Court's choice of reliability metrics—"evidence of a dog's satisfactory performance in a certification or training program" conducted by a "bona fide organization"[370]—is an imperfect one.[371] Nevertheless, by looking to a drug dog's certification, the Court appropriately sought *external* evidence of reliability. "External," accordingly, signifies evidence separate from a technology itself, and from its outputs.

Courts can apply this requirement straightforwardly to policing technology. Consider a case in which law enforcement relies on facial recognition to identify and arrest a suspect. A court considering the legality of the arrest under the Fourth Amendment would begin its determination by analyzing whether the facial recognition technology was reliable enough to support probable cause. It would need to look outside of the software's output—here, a purported "match" to the accused—for such evidence.

Because scientific validity is the measure of a technology's reliability,[372] external evidence of policing technology reliability should demonstrate whether

[367] *Id.* at 246–47.
[368] *See id.*
[369] *Id.* at 246.
[370] *Id.* at 246–47.
[371] *See supra* notes 238–41 and accompanying text.
[372] *Supra* notes 181–82 and accompanying text.

the technology in question is grounded in good science.[373] That means courts should evaluate the degree of empirical testing a technology has undergone and whether that testing was conducted in conditions reflective of the technology's intended use.[374]

Independent evidence of reliability. The Supreme Court in *Harris* also alluded to a second requirement for properly assessing black box reliability. The Court did not suggest that lower courts should take a drug dog handler's word about a dog's reliability for granted.[375] Albeit laxly, the Court in *Harris* paid some attention to the idea that a suitable reliability metric must have some degree of independence from the parties in litigation. The Court looked beyond the parties to an official certifying organization for evidence of reliability.[376]

Here too, however, the underlying logic of *Harris* can be surfaced to construct a more effective test for assessing policing technology reliability in probable cause and reasonable suspicion determinations. The Court was right to look beyond the parties. It just failed to look far enough. It implicitly assumed that a law enforcement certification organization was sufficiently removed from the prosecution of a specific case to provide trustworthy information about a drug dog's reliability.[377] In order for a court's reliability analysis to be meaningful under this new framework, however, evidence of a policing technology's reliability must also be truly independent. As used here, "independent" evidence is not developed by actors who are parties to the case *or* their affiliates, such as law enforcement, prosecution, or defense entities.

Turning back to the facial recognition example above, the court would need to consider evidence of the software's reliability developed by entities unaffiliated with law enforcement or the defense. These might include national vetting organizations or other independent organizations.

Disinterested evidence of reliability. While related to "independence," "disinterested" means more than evidence not developed by parties to the case or their affiliates; it goes further. "Disinterested" connotes neutrality and lack of

[373] *Cf.* Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 590 (1993) (explaining that scientific evidence "must be supported by appropriate validation—*i.e.*, 'good grounds,' based on what is known").

[374] *See supra* notes 183–86 and accompanying text (explaining the importance of establishing validity as applied).

[375] Florida v. Harris, 568 U.S. 237, 246–47 (2013).

[376] *Id.*

[377] *See id.* at 247 ("After all, law enforcement units have their own strong incentive to use effective training and certification programs, because only accurate drug-detection dogs enable officers to locate contraband without incurring unnecessary risks or wasting limited time and resources.").

bias in favor of or against a party or its affiliates. In *Harris*, the Court paid little attention to the need for evidence of reliability to be disinterested. As noted, it undermined its own logic and conclusion by assuming that law enforcement itself could be the sole provider of a drug dog's reliability.[378] Yet, interested parties attempt to influence courts' reliability assessments.[379] They do so in a variety of ways, including by supplying courts with evidence that ostensibly supports reliability, but upon close inspection, does the opposite.[380] Interested parties also sponsor research meant to prop up a technology's reliability.[381] Disinterested evidence is, thus, critical to establishing an effective framework for assessing reliability.

In the facial recognition example above, disinterested evidence might comprise research that is not sponsored or otherwise funded by law enforcement (or defense) bodies. It might also include studies produced by entities unaffiliated with organizations that sell or market products primarily to one side of the adversarial equation.

In *Harris*, the Supreme Court looked for a measure of reliability that avoided creating new evidentiary requirements to determining probable cause.[382] In her opinion for the Court, Justice Kagan sought to preserve the flexibility of the totality-of-circumstances approach to determining probable cause.[383] The adapted framework set forth above does the same.

Practically speaking, courts could implement reliability evaluations of policing technology for purposes of a probable cause or reasonable suspicion determination in several ways. Courts would maintain discretion over how to receive evidence of policing technology reliability. They could ask the parties to supply research, such as empirical studies assessing a technology's validity

[378] *See id.*

[379] *See* Sinha, *Radically Reimagining*, *supra* note 62, at 927 (describing how "segments of the forensic community . . . facilitate the admission of unsound forensic evidence in criminal cases"). Indeed, along with prosecutors, some even attempt to evade judicial reliability assessments altogether. *See* Brett Murphy, *They Called 911 for Help. Police and Prosecutors Used a New Junk Science to Decide They Were Liars.*, PROPUBLICA (Dec. 28, 2022), https://www.propublica.org/article/911-call-analysis-fbi-police-courts (describing how prosecutors worked alongside someone who claims to have expertise in 911 call analysis to avoid having to litigate the scientific foundation of the technique).

[380] *Id.* This author has critiqued such research previously. Sinha, *Automated Gunshot Detection*, *supra* note 34, at 82.

[381] *Id.*; Sinha, *Radically Reimagining*, *supra* note 62, at 927.

[382] *Harris*, 568 U.S. at 244–47.

[383] *Id.* at 244–45.

under relevant circumstances; meta-analyses of studies; or other vetting of a given technology.

Evidence of reliability could be entered at a hearing through expert witnesses or simply provided as part of motions and their responses.[384] Where the parties are made responsible for providing such evidence, it would be up to courts determine if the evidence supplied meets the described criteria. In particular, courts would need to pay close attention to whether evidence supplied by parties is truly independent and disinterested.

Alternatively, courts could seek and obtain such evidence through their own research or by appointment of their own expert.[385] In recognition of the parties' need to fully develop and litigate the reliability of the technology at issue, courts that choose this route should then give the parties an opportunity to bolster or rebut such evidence by providing their own evidence, calling or cross-examining witnesses, or in argument.[386]

Where courts determine that no external, independent, and disinterested evidence exists to indicate a technology's reliability, courts should not simply bypass reliability analysis as many of the courts in this Article's study were revealed to have done. Instead, they should apply an existing framework for assessing the reliability of technology. Adopting a known method for evaluating policing technology reliability when deciding the legality of searches and seizures will assist in resolving the substantial inconsistency in courts' current ad-hoc approaches to evaluating reliability.[387] The most suitable such framework is the reliability test designed for assessing reliability of technology that is applicable at the trial stage, the test the Court laid out in *Daubert*.[388]

The two-pronged framework for addressing policing technology reliability—whereby courts first look for independent, external, and disinterested evidence of reliability and, if no such evidence is available, turn to the *Daubert* test—is designed to be flexible enough to work within a totality-of-circumstances analytical framework, but also sufficiently robust to minimize superficial reliability analysis. On its own, application of the *Daubert* test at

[384] For example, the parties could attach relevant studies to motions to suppress and opposition motions.

[385] *See* FED. R. EVID. 706 (permitting courts to appoint experts of their choosing under certain circumstances.)

[386] Where courts appoint their own experts, they will generally be required to give the parties such opportunities. *Id.*

[387] *See supra* Part III.B.5.

[388] Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 592–94 (1993).

suppression hearings is not a perfect solution. In the thirty years since the case was decided, substantial critique finds that it often fails to filter unreliable scientific and technological evidence in criminal trials.[389]

Using the *Daubert* test as a backstop for evaluating policing technology reliability where no external, independent, and disinterested evidence is available to establish it, however, has the potential to resolve a number of the problems with courts' current approaches revealed here. As a threshold matter, when applied properly, the *Daubert* test *is* an effective method for evaluating reliability of technology.[390] It also has the benefit that judges are familiar with its requirements and application.[391]

The two-pronged framework also has substantial potential to resolve both the doctrinal and non-doctrinal harms identified previously. First, it is a straightforward mechanism for bringing courts into compliance with the Supreme Court's Fourth Amendment reliability precedent. Second, by sending law enforcement the message that its use of technology must survive judicial scrutiny, it also has the potential to mitigate the corollary harm of police overreliance on technology to justify otherwise unwarranted privacy intrusions.

C. *Policy Interventions*

While a doctrinal shift is necessary to slow erosion of the remaining protections of the Fourth Amendment, it is an incomplete solution. In many cases, the harms that flow from police reliance on technology and changes in police behavior resulting from such reliance will occur without ever being subjected to judicial scrutiny. Much technology use will not result in intrusions and thus will not be subject to litigation; intrusions may not result in arrest or prosecution; cases may be dismissed early on in prosecution; or, plea agreements may be made prior to Fourth Amendment litigation commencing. Even where the types of reliability challenges contemplated here are litigated and the recommended approach is applied by courts, the influence of court decisions may be limited by the factual circumstances or technologies at issue. Ensuring reliability of each new and different technology that pops up will prove challenging notwithstanding the doctrinal reform already suggested. Non-doctrinal strategies are, thus, also necessary.

[389] *See supra* notes 275–80 and accompanying text.

[390] *See* Sinha, *Junk Science*, *supra* note 62, at 101 (briefly summarizing critiques of the *Daubert* test and explaining that despite flaws, the test can be used to effectively evaluate reliability).

[391] *Id.*

The abuse and harassment that use of policing technology often results in, coupled with courts' indifference to questions about the reliability of policing technology, suggests that police reliance on technology must be drastically curtailed. Elimination of police reliance on technology, however, is not likely as a comprehensive solution in the near future. Thus, this section proposes intermediary interventions aimed at curtailing use of and mitigating the harms caused by increasing reliance on policing technology for those who currently suffer them.[392]

Much of the preceding discussion emphasizes that, because of their significant impact on communities, policing technologies must be scientifically valid as a precondition to their use. The previous section focuses on a doctrinal reform that may be used to strengthen courts' role in ensuring validity of policing technology. While courts are necessary to scrutinize reliability and hold law enforcement responsible where policing technology is not reliable, they are not necessary to establishing policing technology reliability in the first instance. Rather, legislative and administrative solutions may be designed to incentivize ensuring the reliability of such technology. Legislation can, for example, incorporate *Daubert*'s core precepts emphasizing the importance of testing by requiring that policing technology be validated appropriately and have been demonstrated to operate properly under the circumstances and conditions it will be deployed in prior to use in the community.

Legislative or administrative prescriptions may also help to establish baseline guardrails for preventing and minimizing misuse of policing technology. For example, legislation can mandate that probable cause or reasonable suspicion cannot be premised on the output of a policing technology alone.[393] Lawmakers can require police to submit detailed proposals outlining how a technology will be used and how abuse, misuse, and privacy intrusions will be prevented and mitigated, and that such proposals be approved before a policing technology can be deployed. They can likewise mandate regular, independent audits of policing technology that address reliability, impact on privacy, influence on policing practices, and harm to communities and make such audits available for public review before policing technology use can be

[392] *Cf.* Daniel S. Harawa, *Lemonade: A Racial Justice Reframing of the Roberts Court's Criminal Jurisprudence*, 110 CALIF. L. REV. 681, 688 (2022) (arguing that while "big-picture rethinking" is necessary, those who suffer criminal legal harms "need some solutions *now*" (emphasis in original)).

[393] Some jurisdictions have already put in place such measures. *See* GARVIE, *supra* note 76, at 5 (describing policies in New York and Orlando that disallow police from using facial recognition as sole basis for probable cause).

renewed.[394] Such interventions may have the corollary benefit of encouraging companies that produce policing technology to think about effects on privacy in the development phase.[395]

Greater transparency around police use of technology may also be achieved through prescriptive policymaking. For example, developers and police departments might be required to release testing, use, and error data publicly as a precondition for use. Police can be required to report instances where the output of a technology supported an intrusion and what other information supported the intrusion. Police departments can be required to disclose the geographic locations where policing technology is deployed and document instances where use of policing technology was connected to the use of force or civilian complaints against police for harassment or other misconduct.

Such guardrails can also help to re-align power imbalances between police who leverage technology and communities subjected to its use. Agencies charged with regulating policing technology use can promote education about and community engagement with policing technology. They can encourage collection of public input, including through the audit mechanisms proposed above, to increase civilians' ability to voice concerns about technology use in their communities.[396] Finally, legislation can attempt to balance the scales for the accused by mandating discovery, including information that is frequently protected by the assertion of trade secret privileges,[397] be provided to the accused.

[394] Determining who is properly suited to conduct assessments of legislatively mandated validation testing and audit reports presents a challenge. Legislators and community members may not have the expertise to evaluate technical aspects of policing technology and developers and police departments have built-in conflicts of interest. The ideal choice would involve disinterested, neutral evaluators, perhaps hired by elected legislators.

[395] *See* Matthew Tokson & Ari Ezra Waldman, *Social Norms in Fourth Amendment Law*, 120 MICH. L. REV. 265, 311–12 (2021) (arguing that legislation aimed at governing surveillance, while incomplete, can "influenc[e] corporate behavior")

[396] The ACLU has recommended community advisory boards to counsel legislative bodies on the use of policing technologies. ACLU, AN ACT TO PROMOTE TRANSPARENCY AND PROTECT CIVIL RIGHTS AND CIVIL LIBERTIES WITH RESPECT TO SURVEILLANCE TECHNOLOGY § 8 (2021), https://www.aclu.org/sites/default/files/field_document/aclu_ccops_model_bill_april_2021.pdf (suggesting that community advisory committees can "help guide decisions about if and how surveillance technologies should be used").

[397] Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1378–79 (2018).

It is imperative that such interventions are designed to be more than superficial.[398] Where they are robust, however, they may serve to significantly mitigate the harms stemming from police reliance on technology even where law enforcement conduct does not result in criminal litigation.

These interventions—both doctrinal and non—have the potential to bring about important secondary benefits. The law influences collectively held beliefs and perceptions of what constitutes acceptable conduct.[399] It plays a key role in both generating and modifying social norms.[400] Even where they are relatively stable or long held, the law can shift societal views little by little, over time.[401]

The law especially influences norm development under novel or unexpected circumstances or when norms are unsettled or in the early stages of development.[402] Unsurprisingly then, law influences norms around new technologies and their uses.[403]

The prescriptions described here have the potential to influence collective views regarding several important and interrelated strands of modern policing and privacy. As critical policing decisions are increasingly outsourced to technology, the described interventions can influence perceptions about what degree of technological intervention in police practice is acceptable and, critically, what demands we make of police departments that rely on technology and such technology's reliability.[404] Rather than concretize acceptance of constant deployment of police technologies in communities, doctrinal and policy prescriptions like those posited here can encourage shifting and unsettling increasingly entrenched, but faulty, beliefs about techno-supremacy, including

[398] *See* Sinha, *Automated Gunshot Detection*, *supra* note 34, at 108, 114 (noting that legal reforms can entrench harm and have effects contrary to aims); Sinha, *Radically Reimagining*, *supra* note 62, at 939–43 (laying out framework for non-reformist forensic reform).

[399] Tokson & Waldman, *supra* note 395, at 280.

[400] *See* Janice Nadler, *Expressive Law, Social Norms, and Social Groups*, 42 LAW & SOC. INQUIRY 60, 63–64 (2017) (describing mechanisms by which expressive law can alter views and behaviors).

[401] *See* Tokson & Waldman, *supra* note 395, at 281 (explaining how, over time, legal restrictions on smoking advertising and cigarette altered societal views on smoking).

[402] *See* Roberto Galbiati, Emeric Henry, Nicolas Jacquemet & Max Lobeck, *How Laws Affect the Perception of Norms: Empirical Evidence from the Lockdown*, 16 PLOS ONE 1, 11 (2021) (using empirical evidence related to lockdown restrictions in the United Kingdom to demonstrate the law's effect on social perceptions during the COVID-19 pandemic).

[403] Tokson & Waldman, *supra* note 395, at 281, 296.

[404] *Cf.* Danielle Keats Citron, *Law's Expressive Value in Combating Cyber Gender Harassment*, 108 MICH. L. REV. 373, 407–08 (2009) (describing how court decisions regarding sexual harassment changed perceptions about equality in the workplace); *id.* at 410 (arguing that a "civil rights agenda has the potential to change the social meaning of cyber gender harassment").

perceptions that policing technology is neutral, has inherent value, and is insulated from social and political influence.[405]

CONCLUSION

As the opening vignettes to this Article make clear, the consequences of law enforcement reliance on technology are neither abstract nor theoretical. Even so, criminal procedure doctrine has failed to keep pace with this trend and courts neglect to subject policing technology to the same critical examination applied to traditional information with alarming frequency. The original case analysis presented in this Article is an important window into how policing technology is used to develop reasonable suspicion and probable cause and how courts analyze the reasonableness of intrusions justified by such technology. It also makes clear, though, that more research into these features of the criminal process is necessary. In the meantime, the doctrinal and policy interventions this Article makes are one step toward stemming erosion of privacy rights and remedying the significant real-world harms caused by law enforcement's ever-increasing and un-scrutinized reliance on technology.

[405] *See* Tokson & Waldman, *supra* note 395, at 270 (explaining that "judicial nonintervention" can "normalize new surveillance" practices).