Santa Clara University School of Law

Santa Clara High Technology Law Journal

Volume 40 | Issue 2 Article 1

2024

RELYING ON UNRELIABLE TECH: UNCHECKED POLICE USE OF ALGORITHMIC TECHNOLOGIES

Fraerman, Ali

RELYING ON UNRELIABLE TECH: UNCHECKED POLICE USE OF ALGORITHMIC TECHNOLOGIES

Ali Fraerman[*]

In the past two decades, police forces have come to rely on algorithm-based technologies for investigative leads. Several of these technologies are unreliable. They are prone to error, misidentifying suspects, and crimes. When relied upon, they lead to false arrests and unnecessary stop-and-frisks. Yet, there is no coercive mechanism, either regulatory or judicial, that meaningfully governs the use of these algorithmic technologies in law enforcement. As a result, law enforcement agencies are free to disregard potential errors and deploy emerging technologies against communities with little recourse.

This Article looks closely at three technologies—ShotSpotter gunshot detection, facial recognition technology, and rapid DNA machines—to illuminate reliability issues common to privately-held algorithmic technologies and exacerbated by police misuse. Law enforcement agencies fail to screen technologies before using them to support individualized suspicion for searches and seizures. Thus, the police end up targeting criminal defendants based on unreliable information. But the Fourth Amendment does not meaningfully provide defendants with an avenue to challenge the reliability of technologies used to develop probable cause and reasonable suspicion. Extrajudicial regulation is needed to ensure that the technologies used by law enforcement are reliable. If law enforcement agencies continue to deploy unreliable technologies, courts should suppress evidence stemming from their use.

[*] J.D. 2023, Yale Law School; B.A. 2018, Barnard College. Special thanks to Tracey Meares for supervising and supporting this project; and to Jonathan Manes for involving me in the important advocacy that served as its inspiration. I am grateful to the editors of the Santa Clara High Tech Law Journal, especially Eamon Condon and Blair Huang, for their edits. I am indebted to the mentors who guided me through law school and my nascent career, and, of course, to my family.

CONTENTS

I. INTRODUCTION

When Brandon Johnson was recently elected mayor of Chicago, his election had one unexpected effect. It tanked the stock of private surveillance company SoundThinking, Inc.[1] When campaigning, Johnson vowed to end Chicago's use of SoundThinking's gunshot detection product, ShotSpotter, for which the city pays around nine million dollars per year.[2] ShotSpotter claims to detect gunshots through a system of sensors, enabling faster deployment of police officers to the scenes of gun violence.[3] Johnson acknowledged that the technology did not work and was not worth the cost to the city.

Underlying the debate over Chicago's retention of ShotSpotter is a familiar trade-off between individual liberty and public safety. The police believe that if one life is saved by a faster response to gunfire, the technology is worth the price tag and the tax of increased surveillance on communities.[4] Activists disagree. This debate also has a dimension less explored in the literature because ShotSpotter is not

---

[1] *See Why ShotSpotter Stock Crashed This Week*, THE MOTLEY FOOL (Apr. 7, 2023), https://www.fool.com/investing/2023/04/07/why-shotspotter-stock-crashed-this-week/; *see also Shotspotter Changes Corporate Name to Soundthinking and Launches SafetySmart Platform for Safer Neighborhoods*, SOUNDTHINKING, INC.: PRESS RELEASES (Apr. 10, 2023), https://www.soundthinking.com/press-releases/shotspotter-changes-corporate-name-to-soundthinking-and-launches-safetysmart-platform-for-safer-neighborhoods/. On April 10, 2023, ShotSpotter, Inc. rebranded as SoundThinking, Inc. ShotSpotter is still the name of SoundThinking's gunshot detection product. *See id.* This Article refers to the corporation as SoundThinking and the product as ShotSpotter. Sources cited in this paper from prior to 2022 refer to the corporation as ShotSpotter.

[2] *See* Quinn Myers, *Will ShotSpotter Stay? Mayor-Elect Says 'Better Ways' to Spend Money But Stops Short of Pledging to Dump It*, BLOCK CLUB CHI. (Apr. 12, 2023), https://blockclubchicago.org/2023/04/12/will-shotspotter-stay-mayor-elect-brandon-johnson-says-better-ways-to-spend-money-but-stops-short-of-pledging-to-dump-it/; MacArthur Justice Center, *The Burden on Communities of Color*, END POLICE SURVEILLANCE, https://endpolicesurveillance.com/burden-on-communities-of-color (last visited Oct. 17, 2023).

[3] *See Gunshot Detection Technology*, SOUNDTHINKING INC., https://www.soundthinking.com/law-enforcement/gunshot-detection-technology/ (last visited Oct. 17, 2023).

[4] *See* Fran Spielman, *ShotSpotter Contract Comes Under Heavy Fire*, CHI. SUN-TIMES (NOV. 12, 2021), *https://chicago.suntimes.com/city-hall/2021/11/12/22778971/shotspotter-contract-police-districts-city-council-gunfire-violence-crime.*

accurate. It frequently alerts to sounds that are not gunshots, wasting resources and intensifying a hostile police presence in communities of color.[5]

Evidence of ShotSpotter's unreliability garnered public outcry for the first time in the spring and summer of 2021, at the nexus of two events that exposed the pitfalls in the technology. In March 2021, a thirteen-year-old named Adam Toledo was shot by an officer chasing down a ShotSpotter alert.[6] In April 2021, Chicago prosecutors dropped charges against Michael Williams, a grandfather falsely detained for eleven months for a murder, the only evidence against him an inaccurate ShotSpotter alert.[7] Subsequently, the MacArthur Justice Center (MJC) and the city Office of the Inspector General (OIG) released a report indicating that ShotSpotter was not credibly deploying officers to the scenes of gun crimes.[8] The revelations in the reports prompted city council hearings, but city officials took no further action.[9] Unbeknownst to mayor-elect Johnson, his predecessor quietly renewed the ShotSpotter contract in October 2022.[10] A class-action lawsuit against the city's use of the technology is pending.[11]

Chicago's ShotSpotter debate is not unique to Chicago or to ShotSpotter. The Chicago Police Department (CPD) has been using

---

[5] *See* MacArthur Justice Center, *Shotspotter Study Findings*, END POLICE SURVEILLANCE, https://endpolicesurveillance.com/ (last visited Oct. 17, 2023) (describing analyses that show ShotSpotter falsely alerts to gunfire); *see also* MacArthur Justice Center, *The Burden on Communities of Color*, *supra* note 2 (describing discriminatory deployment of ShotSpotter).

[6] *See* Jon Seidel, *New Lawsuit Aims to Halt Chicago's Use of ShotSpotter*, CHI. SUN-TIMES (July 21, 2022), https://chicago.suntimes.com/news/2022/7/21/23273332/shotspotter-lawsuit-chicago-police-toledo-shooting-michael-williams-arrest-charges-dropped.

[7] *See id*.

[8] *See id*.; *see generally* OFF. OF THE INSPECTOR GEN., CITY OF CHI., THE CHICAGO POLICE DEPARTMENT'S USE OF SHOTSPOTTER TECHNOLOGY, https://igchicago.org/wp-content/uploads/2021/08/Chicago-Police-Departments-Use-of-ShotSpotter-Technology.pdf.

[9] *See ShotSpotter, Chicago Police Defend Gunshot Detection Technology as Chicago Residents Sound Off at City Council Committee Hearing*, CBS NEWS (Nov. 12, 2021), https://www.cbsnews.com/chicago/news/shotspotter-chicago-police-city-council-hearing/.

[10] *See* Tom Schuba, *Lightfoot Administration Quietly Renewed Shotspotter Contract that Johnson has Vowed to Cancel*, CHI. SUN-TIMES (Apr. 7, 2023), https://chicago.suntimes.com/city-hall/2023/4/7/23674434/shotspotter-contract-extended-lori-lightfoot-brandon-johnson-policing.

[11] *See generally* First Amended Complaint, *Williams v. City of Chicago*, No. 1:22-cv-03773 (N.D. Ill. Nov. 14, 2022).

ShotSpotter technology consistently since 2017, but it took high-profile violence and a rare technology audit to bring the system's unreliability to light.[12] Still, CPD continues to deploy ShotSpotter, ignoring errors. And they have SoundThinking on their side: the company has paid for "independent" studies to refute claims of unreliability to save business.[13] They are getting away with it too, because there is no coercive mechanism to vet ShotSpotter's reliability, or the reliability of similar technologies that are privately owned and function through understudied algorithmic mechanisms. Since cities often fail to consult the public when they purchase these technologies from private companies, it takes a huge error, like a false arrest, to draw attention to reliability issues. But instances of error do not mean that a city must stop using a technology: no jurisprudential doctrine or law says that this is the case. Defendants can combat algorithmic technologies piecemeal when they contribute to arrests, but they are ill-equipped to do so because law enforcement agencies and private companies do not share information about the technologies' reliability, mechanisms, or performance.

This Article comes at a meaningful time, when lawyers and activists have begun to challenge algorithmic technologies for their unreliability, after nearly two decades of unexamined use. In the past two years, national civil rights organizations have mounted federal lawsuits challenging facial recognition and gunshot detection software used by law enforcement. The lawsuits contend that these technologies are not fit to provide police cause for a stop or arrest because of their unreliability.[14] Despite reliability concerns garnering nationwide news coverage, major urban police departments, like the CPD, continue to deploy controversial algorithmic technologies. Many agencies have increased their expenditures on technologies like ShotSpotter, despite knowledge of reliability issues, while the federal government foots the bill through initiatives like the American Rescue Plan Act.[15]

---

[12] *See* Michael Wasney, *The Shots Heard Round the City*, SOUTH SIDE WEEKLY (Dec. 19, 2017), https://southsideweekly.com/shots-heard-round-city-shotspotter-chicago-police/.

[13] *See Independent Analysis of the MacArthur Justice Center Study on ShotSpotter in Chicago*, EDGEWORTH ECON., https://www.edgeworthe conomics.com/experience-independent-analysis-of-the-mjc-study-on-shotspotter-in-chicago.

[14] *See* First Amended Complaint, *Williams v. City of Chicago*, No. 1:22-cv-03773 at 15–19, 30–36; *see also* Complaint, *Williams v. City of Detroit*, No. 2:21 Civ. 10827 at 11, 26–27, 52–60 (E.D. Mich. Apr. 13, 2013).

[15] *See, e.g.*, Mark Zaretsky, *New Haven to Spend $12 million for 500 New Surveillance Cameras, ShotSpotter Expansion*, NEW HAVEN REG. (Dec. 22,

The Article proceeds as follows: In Part II, I discuss the value of reliability as an approximation for evidentiary truth. I introduce the philosophical concept of reliabilism, the principle that reasoning from unreliable evidence leads to unreliable outcomes. I also outline how Fourth Amendment probable cause and reasonable suspicion inquiries have sidelined reliability concerns.

In Part III, I introduce a class of emerging police technologies—ShotSpotter gunshot detection, facial recognition technology (FRT), and rapid DNA machines—that rely on algorithms to produce investigative leads. I discuss how programming flaws and use by law enforcement render these technologies unreliable, introducing the potential for false arrests.

In Part IV, I discuss how the reliability of ShotSpotter, facial recognition, and rapid DNA have evaded review. Law enforcement agencies do not screen for reliability when they acquire technologies. Pre-trial inquiries into the legality of searches and seizures ignore reliability when other "circumstances" support suspicion. At trial, technologies that establish investigative leads may not be introduced at all. If they are, it is unclear that courts properly employ evidentiary gatekeeping mechanisms, namely the *Daubert* test, to assess error.

In Part V, I propose solutions to strengthen reliability screening both within and outside the confines of Fourth Amendment jurisprudence. I propose that courts must find algorithmic technologies to be reliable, in a rigorous screening, to allow a lead generated by the technology to contribute to the quantum of suspicion. If a technology is not reliable, evidence generated from its use should be excluded. To regulate the technologies that law enforcement agencies deploy, lawmakers should introduce comprehensive testing for validity and require algorithms to meet certain standards before they are sold. The public should have a continuing say, via democratic processes, in which technologies their cities adopt.

II. RELIABILITY AS A VALUE IN ITSELF

A. *Defining Reliability*

Evidentiary reliability as discussed in this Article[16] is a measure of trustworthiness.[17] It is an approximation for truth. In this

2021), https://www.nhregister.com/news/article/New-Haven-to-spend-12-million-for-500-new-16722043.php#.

[16] Referred to throughout as simply "reliability."

[17] *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 590 n.9 (1993). This is different than scientific reliability, which asks whether the "application of the principle produces consistent results." *Id.*

way, reliability in the legal context is akin to scientific validity, which asks whether "the [scientific] principle shows what it purports to show."[18] For a police technology to be valid, it must produce the output it purports to produce. A perfectly valid gunshot detection system would detect all fired gunshots, but not other sounds. A perfectly valid facial recognition technology would only match an input photo to the same person. A perfectly valid DNA machine would always generate a correct, complete profile. Reliable algorithmic technologies produce trustworthy information.

Reliability requirements are rooted in the truth-seeking function that underlies the United States constitutional justice system.[19] The voluntariness rule, which bars the admission into evidence of confessions given involuntarily as violative of due process, was originally premised on ideas of reliability.[20] Initial Supreme Court rulings on voluntariness excluded confessions, in part, on the principle that admissions infected by measures of coercion are untrustworthy.[21]

When the Court in *Colorado v. Connelly* ultimately rejected the premise that the admission of unreliable evidence violated due process, Justice Brennan dissented from the ruling, writing that "[a] concern for reliability is inherent in our criminal justice system."[22] The Justice explained that the use of an accusatorial system of justice, one in which the government must prove guilt by presenting extrinsic evidence, necessarily requires that such evidence be reliable.[23] Justice Brennan found reliability to be of paramount importance for confessions, because of the confession's "decisive impact on the adversarial process," its persuasive effect of overshadowing all other evidence presented.[24] As early voluntariness cases were concerned with coercion indelibly infecting a confession's trustworthiness, Justice Brennan was concerned with an untrustworthy confession

---

[18] *Id.*

[19] *See* Richard A. Leo, Steven A. Drizin, Peter J. Neufeld, Bradley R. Hall, & Amy Vatner, *Bringing Reliability Back In: False Confessions and Legal Safeguards in the Twenty-First Century*, 2006 WIS. L. REV. 479, 489, 492–94. *Cf. id.* at 499 (describing the end of the Court's use of the reliability rationale to exclude coerced confessions).

[20] *See id.* at 494–95.

[21] *See id.* at 490–94; Scott A. McCreight, Comment, *Colorado v. Connelly: Due Process Challenges to Confessions and Evidentiary Reliability Interests*, 73 IOWA L. REV. 207, 210–11, 210 n. 4 (1987).

[22] *See Colorado v. Connelly*, 479 U.S. 157, 167 (1986); *Connelly*, 479 U.S. at 181 (Brennan, J., dissenting).

[23] *See id.* at 182 (Brennan, J., dissenting).

[24] *Id.* at 182–83 (Brennan, J., dissenting).

indelibly infecting the criminal process. A conviction on such a confession could not stand.[25]

The philosophy of "reliabilism" encapsulates this idea that unreliability in one step of the adversarial process can compound to render the outcome unreliable as well. Put simply, reliabilism is the principle that a belief is justified only if it is obtained through a reliable process.[26] The justificational status of the belief is a function of the process that produces it.[27] The same idea logically works for outcomes: unreliable processes produce unreliable outcomes. Thus, reliabilism supports the notion that the evidence the government relies upon to support searches, seizures, and ultimately, convictions, must be gathered by reliable means.

Reliabilism does not support the recursive principle that a justified belief confers reliability on the process that produced the belief.[28] Likewise, an otherwise justifiable arrest does not confer reliability on the unreliable leads the police followed to affect it. Unreliability in a process, therefore, infects an outcome that would otherwise be true or right by other logic.

There are two ways of thinking about reliabilism in the context of error prone police technologies: a narrow and a broad conceptualization. In the narrow case, the technology itself is the unreliable "process," so it produces an unjustified output. This leads to the conclusion that the police should simply not rely on the technology's output. In the broader case, algorithmic technologies can be thought of as one cog in the investigative "process" that leads to a search or seizure. If the technology's output is unreliable, that output infects the whole process, making it unreliable, and the outcome it led to—for example, an arrest—is unjustified.

---

[25] *See id.* at 183 (Brennan, J., dissenting) ("Minimum standards of due process should require that the trial court find substantial indicia of reliability . . . before admitting the confession of a mentally ill person into evidence . . . . To hold otherwise allows the State to imprison and possibly to execute a mentally ill defendant based solely upon an inherently unreliable confession.").

[26] *See* Alvin Goldman, *What Is Justified Belief?*, *in* JUSTIFICATION AND KNOWLEDGE: NEW STUDIES IN EPISTEMOLOGY at 2, 9–10. (George Pappas ed., 1979).

[27] *See id.*

[28] *See id*. at 9 ("One might initially suppose that the following is a good recursive principle: 'If S justifiably believes q at t, and q entails p, and S believes p at t, then S's belief in p at t is justified'. But this principle is unacceptable. S's belief in p doesn't receive justificational status simply from the fact that p is entailed by q and S justifiably believes q.").

The broader conceptualization is, admittedly, a departure from reliabilism as initially formulated. The reliabilist philosophy explains how a person processes information to arrive at a belief—it does not deal with information stemming from multiple sources, as police investigations do. But the broader conceptualization of reliabilism aligns itself well with the concerns Justice Brennan raised about confessions. False confession experts assert that once police obtain a confession, it "creates its own set of confirmatory and cross-contaminating biases" that cast every other piece of evidence in a worse light for the defendant.[29] The same can be said for investigative leads generated from algorithmic technologies. Police are inclined to put great weight on facial recognition matches, DNA hits, and ShotSpotter alerts that they believe indicate guilt. When those leads are unreliable, they cast doubt on the accuracy of the criminal process. Courts are currently ill-equipped to account for that effect, but reliabilism can capture it, encouraging a harder look at the reliability of each piece of evidence in an investigation.

B. *Reliability and the Fourth Amendment*

The Fourth Amendment provides a constitutional right against "unreasonable searches and seizures."[30] When the police seize an individual on suspicion of ordinary criminal activity, courts typically assess the "reasonableness" of the intrusion by asking whether the police had probable cause[31] or reasonable suspicion[32] to justify it.[33] Searches and seizures premised on unreliable information are unreasonable. The Supreme Court says that probable cause and

---

[29] Leo et al., *supra* note 19, at 519.

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

[31] Probable cause is required to arrest an individual and search their person, possessions, vehicle, or home. Probable cause evades precise definition, but the Court has said that it is the fair probability that an individual committed a crime (for arrests) or the "fair probability that contraband or evidence of a crime will be found in a particular place" (for searches). *Illinois v. Gates*, 462 U.S. 213, 238 (1983).

[32] In *Terry v. Ohio*, the Court developed the reasonable suspicion standard, which governs investigative stops and attendant protective frisks for weapons. 392 U.S. 1, 31 (1968). Reasonable suspicion is a less demanding standard than probable cause, requiring that an officer have a reasonable articulable suspicion that criminal activity may be afoot (to stop) and that an individual is armed (to frisk). *See id.*; *see also Alabama v. White*, 496 U.S. 325, 330 (1990).

[33] *See* THOMAS CLANCY, THE FOURTH AMENDMENT: ITS HISTORY AND INTERPRETATION 676–79 (3d ed. 2017).

reasonable suspicion are "dependent upon both the content of information possessed by police and its degree of reliability."[34]

However, jurisprudence under the Supreme Court's current suspicion inquiry has made reliability a requirement too easily overridden. Both probable cause and reasonable suspicion are determined on the "totality of the circumstances."[35] The Court stresses that this is a flexible test, in which a police officer decides on whether the requisite quantum of suspicion exists based on the "whole picture" before them.[36] Under the current regime, the fact that unreliable information contributed to probable cause or reasonable suspicion is not fatal to the legality of a search or seizure. If a piece of information "has a relatively low degree of reliability," additional information can instead "establish the requisite quantum of suspicion."[37] This generally keeps courts from evaluating the reliability of algorithmic technologies that provide investigative leads, because they can determine that suspicion rests on other grounds.

The "totality of the circumstances" standard was formulated as a rejection of the *Aguilar-Spinelli* test that lower courts used to determine whether probable cause was properly derived from anonymous tips.[38] The previous test required separate findings of both (1) an informant's "basis of knowledge" (how she came by her information) and (2) her "veracity" (her credibility) or, alternatively, the "reliability" of the particular informant's report.[39] Justice Brennan opined that in rejecting the parsing required by *Aguilar-Spinelli*, the Court would authorize findings of probable cause with no assurance that the information on which they were based was obtained reliably.[40]

Justice Brennan's fear was prescient. Considering reliability as part of a "totality" rather than on its own terms contravenes the principles of reliabilism. When potentially unreliable information provides an investigative lead, that unreliable information becomes part of the "process" by which the police arrive at a search or seizure.

---

[34] *White*, 496 U.S. at 330.

[35] *Gates*, 462 U.S. at 238.

[36] *See* Kit Kinports, *Probable Cause and Reasonable Suspicion: Totality Tests or Rigid Rules?*, 163 U. PA. L. REV. 75, 75 n.4 (2014).

[37] *White*, 496 U.S. at 330.

[38] *See Gates*, 462 U.S. at 227, 230–31 (citing *Spinelli v. United States*, 393 U.S. 410 (1969)); *see also Aguilar v. Texas*, 378 U.S. 108, 114 (1964) ("[T]he magistrate must be informed of some of the underlying circumstances from which the informant concluded that the narcotics were where he claimed they were, and some of the underlying circumstances from which the officer concluded that the informant . . . was 'credible' or his information 'reliable.'").

[39] *Gates*, 462 U.S. at 228–29.

[40] *See id.* at 283 (Brennan, J., dissenting).

This should undermine any justification for the search or seizure. However, as Justice Brennan pointed out, the "totality" test allows courts to ignore the presence of unreliability in an investigative step when the police gather enough other information to support probable cause or reasonable suspicion. This ignores how unreliable information infects the investigative process.

Illinois v. Gates, the case overruling *Aguilar-Spinelli* in favor of the totality inquiry, exemplifies the harms of the totality mode of thinking. In *Gates*, police based probable cause on an anonymous letter identifying a couple (Susan and Lance Gates) as drug dealers and identifying their prospective movements.[41] The reversed Illinois Supreme Court and the United States Supreme Court concurred—the letter prompting police to investigate the Gates' activities was unreliable.[42] Nonetheless, the United States Supreme Court, under the "totality of the circumstances" standard, ruled that the letter *plus* an affidavit detailing ostensibly corroborating surveillance supplied probable cause to search the Gates' car and house.[43] But the Court ignored how the unreliable letter shaped the alleged "corroboration." There were significant discrepancies between the predictions in the letter and the Gates' surveilled movements,[44] which "tended to cast doubt" on the hypothesis in the letter that the Gates were trafficking drugs.[45] Nonetheless, the police filled in their suspicions about the Gates from the conclusions in the letter.[46] The unreliable letter became the baseline of the officers' reasoning process, but the majority declined to evaluate how it colored the officers' suspicions.

When the police rely on algorithmic technology, it influences their assessment of individualized suspicion even more than analog tools like tips. Tools like facial recognition and ShotSpotter manufacture probable cause and reasonable suspicion with little analog police work.[47] This allows a significant shift in police power, enabling

---

[41] *See id.* at 225.

[42] *See id.* at 227–28.

[43] *See id.* at 244–45.

[44] *See id.* at 227; *see also Gates*, 462 U.S. 291–93 (Stevens, J., dissenting).

[45] *Gates*, 462 U.S. at 293 (Stevens, J., dissenting).

[46] Justice Stevens suggested that it was unreasonable, given that the letter was wrong, to rely on it as the police did. *See id.* (Stevens, J., dissenting) ("[T]he fact that the anonymous letter contained a material mistake undermines the reasonableness of relying on it as a basis for making a forcible entry into a private home.").

[47] *See* Emily Galvin Almanza & Khalid Alexander, *As Trust in Police Wanes, Cops are Replacing Human Witnesses with Robots*, THE APPEAL (Dec. 20, 2022), https://theappeal.org/police-surveillance-tech-trust/ ("[P]olice officers have increasingly relied on facial recognition software, gunshot

the police to affect a higher volume of stops and arrests without necessitating traditional investigative checks—like first-hand observation—that may cast doubt on the accuracy of a technology's output. And since the "totality" inquiry has left courts under equipped to second-guess reliability, algorithmic technologies have gone unchecked.

III. ALGORITHMIC TECHNOLOGIES: UNRELIABLE AND IMPACTFUL

This Article examines three technologies that the police primarily use as investigative aids: ShotSpotter, facial recognition technology, and rapid DNA machines. All three technologies generate investigative leads via hits or matches to suspects or crime. They all use algorithms to arrive at this hit or match. Broadly defined, algorithms are mathematical processes used to accomplish a particular task.[48] In this Article, "algorithm" refers to pattern matching technology embedded in police tools, and "algorithmic" refers to tools that use this pattern matching.

ShotSpotter takes its inputs from the environment. Sensors detect an impulsive sound and work to classify the sound as a gunshot.[49] ShotSpotter relies on mathematical equations to generate a pinpoint location, while a pattern matching algorithm recognizes whether the sound waves sensors hear match waves it recognizes to correspond to gunshots.[50] The software sends police the pinpointed location of the gunshot with the gunshot classification as a single output, conveying precision.

To use facial recognition technology, police submit a probe photo through a computerized system as the input. An algorithm

---

detection technology, and other automated surveillance technologies (such as triggerfish and stingrays) to maintain control, manufacture probable cause, and arm prosecutors with buckets of 'evidence.'"); *see also Williams v. City of Detroit*, No. 2:21 Civ. 10827 at 4–5, 24–25 (describing probable cause warrant issued from only erroneous facial recognition match and identification by non-eyewitness); Elizabeth E. Joh, *The Unexpected Consequences of Automation in Policing*, 75 SMU L. REV. 507, 523-24 (2022) (describing how the volume of ShotSpotter alerts in an area significantly contributes to Chicago Police Department (CPD) assessments of reasonable suspicion).

[48] *See* ANDREW GUTHRIE FERGUSON, THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT 18 (2017).

[49] *See infra* Section III.A.1 (describing how ShotSpotter works).

[50] *See* Robert B. Calhoun et. al., *Precision and Accuracy of Acoustic Gunshot Location in an Urban Environment*, SHOTSPOTTER 2–8 (2020), https://www.shotspotter.com/wp-content/uploads/2021/08/TN-098-Accuracy-of-Acoustic-Gunshot-Location.pdf.

analyzes the features of the face in the probe photo and attempts to match it with features of photos of known people in a database.[51] The system generates a list of possible matching photos as its output, with varying levels of confidence.[52]

Rapid DNA machines analyze swabs of DNA and generate DNA profiles as an initial output.[53] The machines themselves do not work on algorithms, but for investigative use, they are enabled to upload DNA profiles to databases.[54] DNA databases use search algorithms to match the profile the machine generates to profiles implicated in crimes, generating a secondary output—a match—if successful.[55]

ShotSpotter, facial recognition technology, and rapid DNA technology all operate on privately held software, as do many other algorithmic police tools.[56] These technologies are marketed, sold, and programmed by for-profit entities. The police do not own the mechanisms or algorithms that underlie the technologies' function, and, in some cases, law enforcement does not even own the data the technologies generate.[57] Private ownership creates unique obstacles to assessing reliability.

Police reliance on algorithmic technologies creates bias issues. Humans exhibit automation bias: we tend to trust and over-rely on automated decisions, even when presented with evidence of system error.[58] Algorithmic technologies present as highly technical, generating outputs that convey precision and objectivity.[59] This

---

[51] *See infra* Section III.B.1 (describing in detail how facial recognition technology works).

[52] *See id.*

[53] *See infra* Section III.C.1 (describing in detail how rapid DNA machines work).

[54] *See id.*

[55] *See id.*

[56] *See* Elizabeth E. Joh, *The Undue Influence of Surveillance Technology Companies on Policing*, 92 N.Y.U. L. REV. 19–21 (2017).

[57] *See* Elizabeth E. Joh, Comment, *Reckless Automation in Policing*, 2022 BERKELEY TECH. L.J. 117, 122 (2022) ("[P]ublic agencies usually stand in a customer-vendor relationship with private companies and then adopt the tools of algorithmic decision-making as a matter of purchase, lease, or contract."); *see also* Benjamin Goodman, Note, *ShotSpotter—The New Tool to Degrade What is Left of the Fourth Amendment*, 54 UIC L. REV. 797, 802 (2021) ("[C]ities do not own the data accumulated by ShotSpotter sensors.").

[58] *See* Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1271–72 (2008).

[59] *See, e.g.*, Brief for Brighton Park Neighborhood Council, et al. as Amici Curiae Supporting Defendant, *State v. Williams*, 20 CR 0899601 at 14–15 (Ill. Cir. Ct. May 3, 2021) ("[The ShotSpotter] apps present officers with a display

exacerbates automation bias and obscures contrary evidence of unreliability. Automation bias works in a feedback loop with confirmation bias. Confirmation bias is the human tendency to interpret information in a manner consistent with previously established knowledge or expectations.[60] Thus, when the police are inclined to trust algorithmic technologies, and receive an algorithm-generated output, they will conform subsequent information to match that output. For example, an officer will preference information that supports a facial recognition program's designation of a match or ignore information that does not support it.[61]

Both facial recognition and rapid DNA are considered "feature-comparison" forensic methods. A feature-comparison method relies upon comparing the features from an "evidentiary sample" (e.g., from a crime scene) to those from a "source sample" (e.g., a particular person) to determine whether the two are a match.[62]

In 2016, the President's Council of Advisors on Science and Technology (PCAST) convened with legal experts and produced a report on the scientific validity of feature-comparison forensic methods.[63] The Council's comments on validity are the standard by which experts agree feature-comparison forensics should be judged, and are also a useful guide for evaluating ShotSpotter's validity.

PCAST identified two types of validity: foundational validity and validity as applied. Foundational validity refers to a method or technology's validity as empirically tested—that it "be shown, based on empirical studies, to be repeatable, reproducible, and accurate, at levels that have been measured and are appropriate to the intended application."[64] Or in other words, that it is "in principle, reliable."[65]

To establish foundational validity, PCAST deemed it essential that "a method be subject to empirical testing by multiple groups," all

---

that conveys digital objectivity, showing the number of . . . gunshots and a precise location indicated with a single pin on a street-view map . . . .").

[60] *See* CLARE GARVIE, A FORENSIC WITHOUT THE SCIENCE: FACE RECOGNITION IN U.S. CRIMINAL INVESTIGATIONS 30 (Georgetown L. Ctr. on Priv. & Tech., 2022).

[61] *See id.*

[62] *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* 46 (Sep. 2016); *see also* GARVIE, *supra* note 60 at 13–14 (defining facial recognition as a feature comparison method).

[63] *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *supra* note 62, at 1–2.

[64] *Id.* at 4–5, 47–54.

[65] *Id.* at 56.

which have "no stake in the outcome."[66] When a method requires "subjective" judgment—a determination by a human—the empirical testing must measure the validity of the human component.[67]

Validity as applied measures a method's reliability in practice.[68] Regardless of how a technology performs in a controlled study, the environment in which it is applied and the person applying it affect its reliability. When a method involves a subjective element, validity as applied necessitates a measurement of an individual analyst's error rate, as well as the error rate in the relevant law enforcement jurisdiction.[69] Transparency is crucial to validity as applied—in any given criminal case, PCAST recommends that the procedures used in deploying a technology and the results obtained be available for evaluation.[70]

Also central to assessing validity is accuracy, which encompasses two measurements: a false positive rate and sensitivity.[71] For the purposes of the technologies discussed in this Article, the false positive rate is the probability that a technology declares a "hit" when none exists: the probability that ShotSpotter will identify a gunshot when none was fired; that a facial recognition program will declare two photos are of the same person when they are not; that a DNA database will match two profiles that are, in reality, different. A technology's sensitivity is the probability that it declares a correct hit or match. This encompasses the technology's false negative rate, which is the probability that it will miss a correct hit (sensitivity = 100% – % false negative).

False positive and false negative rates are both necessary to assess a technology's reliability, as they both indicate the existence of error.[72] But false positives are especially important, because only false positives will prompt police to act, leading to wrongful searches and seizures.

Since the software for ShotSpotter, facial recognition, and rapid DNA are all privately owned, it is difficult to assess how (or if) the technology has been foundationally validated for use. Known independent testing on ShotSpotter and facial recognition has been

---

[66] *Id.* at 5, 14.

[67] *See id.* at 5–6.

[68] *See id.* at 5.

[69] *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *supra* note 62, at 6.

[70] *See id.*

[71] *See id.* at 48.

[72] *See id.* at 50.

glaringly insufficient. We know even less about error rates in all three technologies' outputs as applied.

A. *ShotSpotter*

1. How it Works

ShotSpotter is a gunshot detection product owned by SoundThinking, Inc. ShotSpotter systems are composed of networks of sensors with embedded microphones in a given geographic area.[73] The sensors listen for impulsive sounds, sounds with certain measured characteristics like power and amplitude.[74] This encompasses many urban noises, anything that goes "'bang,' 'boom,' or 'pop.'"[75] When multiple sensors detect a sound the computer considers to be impulsive[76]—ShotSpotter needs at least three sensors to pick up a noise[77]—algorithms process the noise to locate and classify it before an alert to gunfire is transmitted to the police.[78]

ShotSpotter works with two algorithms. The first algorithm comes up with a location and timestamp for the sound by comparing the amount of time it took for the noise to reach each sensor in a process called multilateration.[79] The second algorithm generates a classification for the noise as a "gunshot," "possible gunshot," or other noise.[80] This algorithm works by creating a visual image of the waveform of the detected noise and other "features of the incident," like the location and number of nearby incidents, into a "mosaic" that

---

[73] *See* Transcript of Testimony of Paul Greene, Senior Forensic Engineer at ShotSpotter, *United States v. Godinez*, No. 18 CR 00278 at 375–76 (N.D. Ill. June 11, 2019), https://pdfhost.io/v/OGwyQ8Hpd_Godinez_trial_transcript.pdf [hereinafter Greene *Godinez* Testimony].

[74] *See id.* at 379–80; *see also* Calhoun et al., *supra* note 50, at 8.

[75] Greene *Godinez* Testimony, *supra* note 73, at 380.

[76] *See* Calhoun et al., *supra* note 50, at 7–8.

[77] *See* N.Y. Police Dep't, *ShotSpotter: Impact and Use Policy*, N.Y.C. 3 (Apr. 11, 2021), https://www.nyc.gov/assets/nypd/downloads/pdf/public_information/post-final/shotspotter-nypd-impact-and-use-policy_4.9.21_final.pdf.

[78] *See* Greene *Godinez* Testimony, *supra* note 73, at 380–82.

[79] *See id.* at 380; *see also* Calhoun et al., *supra* note 50, at 2–8; First Amended Complaint, *Williams v. City of Chicago*, No. 1:22-cv-03773 at 20.

[80] First Amended Complaint, *Williams v. City of Chicago*, No. 1:22-cv-03773 at 20.

a computer then compares to mosaics previously identified as gunshots.[81]

As a supposed check on the algorithm, SoundThinking employs human analysts to review ShotSpotter's initial classifications.[82] After confirming that the noise is a "gunshot" or a "probable gunshot," they send a ShotSpotter alert to the relevant law enforcement agency.[83] The whole process—from noise detection to dispatch—takes sixty seconds.[84]

Alerts appear in emergency dispatch centers, on an application on officers' smartphones and computers, or both.[85] For example, alerts for both probable gunshots and gunshots are displayed to personnel in Chicago through the ShotSpotter application, which is monitored by personnel at CPD Strategic Decision Support Centers (SDSCs).[86] SDSC personnel send alerts through to the Office of Emergency Management (OEMC), which dispatches officers to respond to all probable gunshots.[87] Officers also have twenty-four seven access to the ShotSpotter app on their own mobile devices, and can view and respond to alerts on patrol.[88]

2. Reliability

ShotSpotter is not foundationally valid, because it has never been independently tested for sensitivity or false positives in an independent study.[89] It is unclear that ShotSpotter can do what its parent company claims it can do: accurately detect gunshots versus

---

[81] *See* Calhoun et al., *supra* note 50, at 8–9. ShotSpotter engineers do not disclose which information thresholds lead a mosaic to be classified as a gunshot or potential gunshot. *See id.*

[82] *See Gunshot Detection Technology*, *supra* note 3 ("Acoustic experts, who are located and staffed in ShotSpotter's 24×7 Incident Review Center, ensure and confirm that the events are indeed gunfire.").

[83] *See* First Amended Complaint, *Williams v. City of Chicago*, No. 1:22-cv-03773 at 24.

[84] *See Gunshot Detection Technology*, *supra* note 3 ("This entire process takes less than 60 seconds from the time of the shooting to the digital alert popping onto a screen of a computer in the 911 Call Center or on a patrol officer's smartphone or mobile laptop.").

[85] *See id.*

[86] *See* OFF. OF THE INSPECTOR GEN., *supra* note 8, at 6–7, https://igchicago.org/wp-content/uploads/2021/08/Chicago-Police-Departments-Use-of-ShotSpotter-Technology.pdf.

[87] *See id.*

[88] *See id.*

[89] *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *supra* note 62, at 47–56 (defining foundational validity).

other loud sounds.[90] SoundThinking and contracting law enforcement agencies have conducted live-fire tests and studies that measured whether ShotSpotter detects fired gunshots.[91] However, no test has measured false-positive errors: how often the system picks up noises other than gunshots. SoundThinking has never shared its algorithm with independent experts to facilitate such testing.[92]

Studies that estimate ShotSpotter's error rates from police data show that the technology is likely very unreliable. The Chicago OIG, a municipal oversight agency, recommended that the city end its contract with ShotSpotter after finding that up to 90.9% of alerts dispatched to CPD may be false positives.[93] A similar study from civil rights organization MJC estimated that the CPD saw roughly 31,640 unfounded deployments from ShotSpotter alerts in a given year (measured 2021–2022), equating to eighty-seven fruitless deployments

---

[90] *See* Jillian B. Carr & Jennifer L. Doleac, *The Geography, Incidence, and Underreporting of Gun Violence: New Evidence Using ShotSpotter*, BROOKINGS 5 (Apr. 27, 2016), https://www.brookings.edu/wp-content/uploads/2016/07/Carr_Doleac_gunfire_underreporting.pdf; *see also* Brief for Roderick & Solange MacArthur Justice Center, et al. as Amici Curiae Supporting Defendant-Appellee, *Commonwealth v. Ford*, No. 2:21 Civ. 10827 at 23 (Mass. App. Ct. Sep. 24, 2021).

[91] *See, e.g.*, Lorraine Green Mazerolle, James Frank, Dennis Rogan & Cory Watkins, *A Field Evaluation of the ShotSpotter Gunshot Location System: Final Report on the Redwood City Field Trial*, OFF. JUST. PROGRAMS (Nov. 1999), https://www.ojp.gov/pdffiles1/nij/grants/180112.pdf; Erica Goode, *Shots Fired, Pinpointed and Argued Over*, N.Y. TIMES (May 28, 2012), https://www.nytimes.com/2012/05/29/us/shots-heard-pinpointed-and-argued-over.html; Dori Keren, *ShotSpotter Pilot Assessment*, LAS VEGAS METRO. POLICE DEP'T 18 (Oct. 2018), https://www.shotspotter.com/wp-content/uploads/2019/08/LVMPD-ShotSpotter-Assessment-V102418.pdf; Calhoun et al., *supra* note 50, at 1, 9–10. If a city conducts a live fire test upon sensor installation, the typical test consists of twelve fired shots. Calhoun et al., *supra* note 50, at 9.

[92] *See* Brief for Roderick & Solange MacArthur Justice Center, et al., *Ford*, No. 2020-P-1334 at 17.

[93] *See* OFF. OF THE INSPECTOR GEN., *supra* note 8, at 2–3. That is, only 9.1% of alerts in a given time period (January 2020 to June 2021) resulted in documented evidence of gun crime. *See id.*

each day.[94] Analyses from other cities also show that ShotSpotter alerts rarely lead police to evidence of gun crime.[95]

These troubling statistics are not surprising, since each step of the ShotSpotter process is prone to error. First, ShotSpotter's initial screening does not adequately distinguish between gunshots and other sounds. As company engineers have often repeated, the sensors pick up impulsive noises.[96] Gunshots are acoustically similar to many noises common in urban environments: fireworks and small explosives, construction, dumpster lids slamming, cars backfiring, and motorcycles.[97] Other features of urban environments, like background noise, make discriminating between gunshots and non-gunshot high amplitude sounds more difficult.[98] SoundThinking has acknowledged these limitations in official documents, warning that "ShotSpotter cannot guarantee '100% detection,' due to interference from 'buildings, topography, foliage, periods of increased traffic or construction noise, and other urban acoustic noises.'"[99]

Second, ShotSpotter engineers admit that ShotSpotter's location algorithm is also easily affected by environmental factors like wind.[100] Thus, alerts can be "significantly mislocated" or time-stamped incorrectly because of the effect of environmental error on the localization algorithm.[101]

Third, the classification algorithm is built on a shaky premise. The algorithm learns to distinguish gunshots from mosaic tiles that have already been classified by human reviewers.[102] However, the

---

[94] *See* MacArthur Justice Center, *Research Findings*, END POLICE SURVEILLANCE, https://endpolicesurveillance.com/research-findings (last visited Feb. 16, 2024).

[95] *See* First Amended Complaint, *Williams v. City of Chicago*, No. 1:22-cv-03773 at 15–16 (describing that only 0.5–5% of alerts in recent years in Dayton, Houston, Atlanta, and Minneapolis led to arrests).

[96] *See, e.g.*, Greene *Godinez* Testimony, *supra* note 73 (testifying that the sensors detect any sound that goes "bang," "boom," or "pop.").

[97] *See* Juan R. Aguilar, *Gunshot Detection Systems in Civilian Law Enforcement*, 63 J. AUDIO ENG. SOC'Y 280, 287 (2015).

[98] *See id.* at 281–82 (2015).

[99] *See* Brief for Amici Curiae Roderick & Solange MacArthur Justice Center, et al., *Ford*, No. 2020-P-1334 at 17 (citing *Detailed Forensic Report*, SHOTSPOTTER, https://www.shotspotter.com/wp-content/uploads/2019/05/DFR-Example-.pdf).

[100] *See* Calhoun et al., *supra* note 50, at 6.

[101] Brief for Amici Curiae Roderick & Solange MacArthur Justice Center, et al., *Ford*, No. 2020-P-1334 at 18 (quoting Testimony of Paul Greene, Senior Forensic Engineer at ShotSpotter, *People v. Simmons*, No. 2016-0404 at 113 (N.Y. County Ct. Monroe County Oct. 17, 2017)).

[102] *See* Calhoun et al., *supra* note 50, at 8.

"ground truth" of the vast majority of the training samples is unknown: ShotSpotter engineers have no knowledge of whether these samples were correctly classified and correspond to actual gunshots.[103] Therefore, the engineers admit, "it is to be expected that some training data are misidentified."[104] ShotSpotter also incorporates non-acoustic information in its sound mosaics, like the "location of recent nearby incidents" and "recent incident counters," meaning that the algorithm also learns to classify gunshots based on more than a noise.[105] Instead, the machine uses its own data on previously triggered alerts to make gunshot determinations, so it is more likely to classify a noise as a gunshot where it has triggered alerts before—regardless of whether that alert was a false positive. According to MJC, this creates feedback loops "that falsely inflate the number of ShotSpotter alerts in particular areas."[106]

Fourth, the proficiency of ShotSpotter's human analysts is unknown. SoundThinking relies on analysts to determine whether ShotSpotter sends out alerts. The human analysts are meant to double-check the algorithm's guess, either confirming its gunshot classification or overriding it.[107] When the ShotSpotter algorithm was criticized, SoundThinking CEO Ralph Clark cited that ShotSpotter outputs are ultimately a product of human review.[108] Yet, no one knows if these analysts are qualified to make such consequential decisions. SoundThinking only requires that employees have one year of professional experience, and though they purportedly favor hiring

---

[103] *See id.*

[104] *Id.*

[105] *Id.*

[106] First Amended Complaint, *Williams v. City of Chicago*, No. 1:22-cv-03773 at 21.

[107] Data suggests that the analysts override the algorithm's determination 10% of the time. *See* Garance Burke, *Confidential Document Reveals Key Human Role in Shotspotter Gunfire Detection System*, CHI. SUN-TIMES (Jan. 22, 2023), https://chicago.suntimes.com/2023/1/22/23567077/confidential-document-reveals-key-human-role-in-shotspotter-gunfire-detection-system.

[108] *See SoundThinking's™ Response to Associated Press Article*, SOUNDTHINKING (Aug. 26, 2021), https://www.soundthinking.com/blog/soundthinking-response-to-associated-press-article ("Our real-time alerting and classification process is driven by a human reviewer . . . ."); *see also* Matt Masterson, *Activists Call on Chicago Officials to Dump ShotSpotter Contract*, WTTW (Aug. 19, 2021), https://news.wttw.com/2021/08/19/activists-call-chicago-officials-dump-shotspotter-contract ("ShotSpotter CEO Ralph Clark told the AP: 'The point is anything that ultimately gets produced as a gunshot has to have eyes and ears on it . . . . Human eyes and ears, OK?'").

people with law enforcement training, no audio engineering experience or proficiency with sound differentiation appears to be required.[109]

PCAST standards would require proficiency testing of ShotSpotter's analysts to ensure foundational validity and validity as applied, since the transmission of gunshot alerts relies on their subjective determinations.[110] But ShotSpotter has never shared details on the proficiency or training of its human analysts. Audio experts express doubts about the human ability to distinguish gunfire from other noises, especially in the sixty seconds ShotSpotter allots for the dispatch of alerts.[111]

Though analysts are supposed to correct for errors in the algorithm, it appears that they contribute to ShotSpotter's high rates of false positives. Journalists recently obtained a nonpublic SoundThinking training document instructing reviewers to err toward gunfire classifications if they are on the fence.[112] Other training protocols are nonpublic and unknown.

Every step of the ShotSpotter process compounds the likelihood that police will receive an erroneous alert. The final step in the system—dispatch—increases the likelihood that they will rely on it. The ShotSpotter app conveys a sense of objectivity to officers and dispatch centers like SDSC and OEMC. Rather than disclaim error or warn of subjectivity, the interface points police to a dot on a street-view map, with a corresponding address, number of gunshots, and timestamp down to the second.[113] Some ShotSpotter packages also convey to police whether an automatic weapon was fired and whether there were multiple shooters.[114]

---

[109] *See Incident Review Center Specialist (Fremont, CA) – 6* Openings, SOUNDTHINKING, https://www.simplyhired.com/job/XbVBC6BE2qtkGGq7OhlfLcWGgY6tH-qTKdkxod7dkeydj-AaeV3KHQ (last visited Oct. 22, 2023) (job posting for ShotSpotter analysts).

[110] *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *supra* note 62, at 5–6, 14.

[111] *See* Burke, *supra* note 107.

[112] *See id.*

[113] *See* Brief for Roderick & Solange MacArthur Justice Center, et al., *Ford*, No. 2020-P-1334 at 21–23 (displaying screenshots of ShotSpotter's mobile app).

[114] *See ShotSpotter FAQ*, SHOTSPOTTER (Aug. 2018), https://www.shotspotter.com/system/content-uploads/SST_FAQ_January_2018.pdf.

B. *Facial Recognition*

1. How it Works

In policing, facial recognition technology is used for both face verification and face identification.[115] Face verification is a one-to-one photo comparison that seeks to confirm a person's claimed identity.[116] Face identification is a search to identify an unknown face, performed by comparing a photo of an unknown person against a database of photos with known identities: the comparison of one photo to many.[117] This Article is concerned with face identification, particularly what the Georgetown Center on Privacy Law refers to as the "investigate and identify" function: when the police input an image of a crime suspect into facial recognition software, hoping to identify the suspect.[118]

"Investigate and identify" procedures typically occur in six steps. First, a police officer chooses a "probe photo" to input in the facial recognition search.[119] This "probe" can be a still from a security camera or a smartphone, a social media post, or a surreptitious photograph.[120] For example, an officer investigating a shoplifting incident may use a still of a suspect from surveillance footage of the store as a probe.[121]

Next, the officer may have a choice in which database the probe image is searched against.[122] The database dictates how many images the probe is searched against and also the type of images.[123] Some databases are restricted to mugshots, while others are repositories of driver's licenses or other ID photos.[124] Third, the officer may choose to "preprocess" the photo by editing the image quality or adding in missing or blurred facial features to heighten her chances of getting a match.[125]

---

[115] *See* CLARE GARVIE ET AL., THE PERPETUAL LINEUP: UNREGULATED POLICE FACE RECOGNITION IN AMERICA 10 (Georgetown L. Ctr. on Priv. & Tech., 2016).

[116] *See id*.

[117] *See id*.

[118] *Id*.

[119] *See id*. at 9–10.

[120] *See generally id*. at Background.

[121] *See* GARVIE, *supra* note 60, at 10.

[122] *See id.* ("For example, an analyst with the FBI's Face Analysis, Comparison, and Evaluation Services Division (FACE Services) can run or request face recognition searches on 21 different state driver's license databases in addition to various federal photo repositories.").

[123] *See id.*

[124] *See* GARVIE ET AL., THE PERPETUAL LINEUP, *supra* note 115.

[125] *See* GARVIE, *supra* note 60, at 11.

In the fourth step, the photo is processed through the facial recognition program's algorithm.[126] The algorithm analyzes the features in the probe photo, searching the database for images with matching features.[127] The program looks for features that align in appearance, size, and distance apart; training itself on what constitutes a match.[128] Most systems generate multiple possible matches, or "candidates," listing them in order of how closely features overlap with the probe image, assigning a confidence level to each candidate match.[129] Thus, facial recognition technology is not designed to generate single positive identifications, but rather lists of possible matches.[130]

This necessitates a fifth step, in which an officer or analyst must examine the candidate list to pick out a possible match.[131] Most law enforcement agencies instruct that a potential match, if identified, is still not a positive identification: it is just an investigative lead, and more corroboration is necessary for probable cause.[132] So, as a sixth step, officers must generate additional evidence corroborating that a match is indeed their suspect in order to effect an arrest.[133] Yet, how much corroboration is legally required, and how much corroboration officers seek in practice, is unclear. In several instances, police have effected arrests based on only a facial recognition match and a suggestive confirmatory procedure: officers have generated "confirmations" by comparing a candidate match to a probe photo by themselves, placing a candidate match in a lineup for identification by a non-witness, and texting a candidate match to a witness.[134]

2. Reliability

The National Institute of Standards and Technology (NIST)[135] has conducted incomplete validation testing on some facial recognition

---

[126] *See id.*

[127] *See* Andrew Ferguson, *Facial Recognition and the Fourth Amendment*, 105 MINN. L. REV. 1105, 1110–11 (2021).

[128] *See id.*

[129] *See id.*

[130] *See* GARVIE, *supra* note 60, at 12.

[131] *See id.*

[132] *See id.*

[133] *See id.*

[134] *See id.* at 2, 5, 7–8.

[135] The National Institute of Standards and Technology (NIST) is a federal body concerned with standards in technology and measurement. *See About*, NAT'L INST. OF STANDARDS & TECH., https://www.nist.gov/about-nist (last visited Jan. 14, 2024). It promotes national technological innovation, in part,

algorithms that parent companies submitted to the agency for review.[136] The tests showed that the algorithms were inconsistently reliable. Some algorithms had very low false negative rates—as low as 0.1%—when matching probe and database photos that were "good quality, frontal images such as 'mugshot images collected with an attendant present."[137] Others had false negative rates of up to 50%, meaning the algorithm could not find a match that was in the database half of the time.[138] Since most facial recognition programs work by generating lists of candidates, false negative errors also indicate the potential for false positive errors. For example, an algorithm with 50% measured false negative error will return a candidate list of completely innocent people (non-matches) 50% of the time.[139] But this algorithm is still generating a list of candidates—meaning an officer could still pick out a false "match" themselves.[140]

This NIST testing has not foundationally validated facial recognition algorithms, because it has not tested the algorithms as they are used in law enforcement.[141] For example, NIST testing has not properly examined how facial recognition technology handles the low-quality probe photos frequently used by law enforcement. Even the poorer quality photos NIST scientists included in sampling were posed.[142] In practice, law enforcement officers and analysts often use much lower quality, unposed probe photos to identify suspects, as with stills from surveillance cameras or cell phones.[143] And when police edit

---

by conducting developmental validation studies on technologies used in the United States. *See id.*

[136] *See* PATRICK GROTHER, MEI NGAN & KAYEE HANAOKA, NAT'L INST. OF STANDARDS & TECH., NISTIR 8271 DRAFT SUPPLEMENT, FACE RECOGNITION VENDOR TEST (FRVT) PART 2: IDENTIFICATION, 7 (2022).

[137] *Id.* at 8–9.

[138] *See id.*

[139] *See* GARVIE, *supra* note 60, at 18.

[140] *See id.*

[141] *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *supra* note 62, at 5 ("Foundational validity requires that a method has been subjected to *empirical* testing . . . *under conditions appropriate to its intended use*.) (second emphasis added).

[142] *See* GROTHER ET AL., *supra* note 136, at 8. The NIST tests included side-view images, poorer-quality webcam images, and ATM-style kiosk photos. The use of these lower-quality photos resulted in recognition error rates "in excess of 20%" among the algorithms tested. *Id.*

[143] *See* GARVIE, *supra* note 60, at 10; *see, e.g.*, Complaint at 18, *Williams v. City of Detroit*, No. 2:21 Civ. 10827 (showing probe photo that supported investigative lead and arrest, as low-resolution, poorly illuminated still from surveillance video in which facial features were barely visible and partially obscured with a hat).

a probe photo, they introduce greater potential for error. The New York Police Department has attempted to "correct" the photos they use as probes, by replacing expressive facial features with features from other photos that more closely resemble mugshots.[144] Police will also use composite sketches as probes, or input celebrity approximations.[145] But because facial recognition technology works by probing individual features, inputting anything other than an exact photograph of someone's face is more likely to generate inaccurate candidate matches.

Therefore, the NIST tests' estimates of inaccuracy, while still significant, are conservative. First, the NIST only tested algorithms willingly submitted by private companies, not necessarily the algorithms in use by law enforcement.[146] Second, the tests did not assess how the programs performed on the many types of poor-quality images commonly used by law enforcement.[147] Third, the NIST tests were conducted on smaller databases than those commonly used by law enforcement, when larger databases introduce greater potential for false positives.[148]

Fourth, and perhaps most importantly, the NIST tests also did not assess the proficiency of the human decisionmakers who choose from the candidate matches generated by FRT programs.[149] Since subjectivity is involved in facial recognition's final output—the match officers rely upon involves a human determination—the technology

---

[144] *See* CLARE GARVIE, GARBAGE IN, GARBAGE OUT: FACE RECOGNITION ON FLAWED DATA (Georgetown L. Ctr. on Priv. & Tech, 2019).

[145] *See id.*

[146] *See* Barry Friedman, Jacob D. Fuchsberg, Written Testimony Before the House Committee on the Judiciary Subcommittee on Crime, Terrorism, and Homeland Security, Facial Recognition Technology: Examining Its Use by Law Enforcement at 8–9 (July 13, 2021). While the most recent round of NIST testing in 2019 assessed newer algorithms, which showed dramatic increases in accuracy, this testing may not reflect changes in the accuracy of law enforcement FRT, because some agencies use outdated technology. *See* GARVIE, *supra* note 60, at 18 ("Current cases also aren't insulated from old algorithms. Public records suggest that in 2016 the Maricopa County Sheriff's Office in Arizona was using an algorithm that was last updated four or more years prior. Up until 2019, the Utah DMV system was similarly operating on an algorithm from a company that appears to have ceased existing in 2012.").

[147] *See* Friedman, *supra* note 146, at 8–9; *see also* GROTHER ET AL., *supra* note 136, at 8–9 (describing that algorithms were tested on "frontal mugshots, profile view mugshots, desktop webcam photos, visa-like immigration application photos, immigration lane photos, and registered traveler kiosk photos," all head on photographs).

[148] *See* Friedman, *supra* note 146, at 8–9.

[149] *See id.* at 9.

must be validated through black-box studies of analysts and officers' performance in selecting matches from candidate lists.[150] No adequate validation study has been conducted on United States law enforcement officers, but other studies suggest that humans—including law enforcement officers—are not very good at recognizing faces. One 2014 study of Australian passport issuance officers found that the officers made identification errors about 50% of the time when selecting from simulated candidate lists, despite using facial recognition software regularly and having an average of 8.5 years of experience in conducting identity comparisons.[151]

Circumstances unique to facial recognition technology as it is used in criminal investigations further complicate recognition.[152] A generated candidate list may contain dozens, or hundreds, of similar looking matches. The subject may not be among them. Probe photographs may be low quality, making comparison difficult. Or, the database photograph may be so old that changes in weight, hairstyle, or other characteristics make matching to a present-day probe photo harder. Specialized training in human facial anatomy and morphology would improve human facial recognition, but there are few indications that United States law enforcement agencies provide this training to officers who use facial recognition software.[153]

The unreliability of facial recognition technology is exacerbated when police seek to identify non-white faces. NIST testing revealed that, across the board, algorithms are more likely to produce false matches when searching for minorities, particularly women of color.[154] False positives were highest when algorithms were searching for people of African and East Asian descent, and lowest when they searched for Eastern European individuals.[155] False positives in searches for women were two-to-five times higher than in searches for

---

[150] *See* GARVIE , *supra* note 60, at 14.

[151] *See* David White et al., *Error Rates in the Use of Automatic Face Recognition Software*, PLOS ONE Vol. 10, Issue 10, 2, 10–11 (Oct. 14, 2015).

[152] *See* GARVIE, *supra* note 60, at 22.

[153] *See id.* at 23–26 ("While some departments have a dedicated, trained face recognition unit, others have allowed most or all law enforcement officers in a given municipality or state to run searches with minimal training.").

[154] *See* PATRICK GROTHER ET AL., NAT'L INST. STANDARDS & TECH., FACE RECOGNITION VENDOR TEST (FRVT) PART 3: DEMOGRAPHIC EFFECTS, 63 (2019).

[155] *See id*. at 7–8.

males.[156] This is the result of developers training facial recognition programs on data sets with low demographic diversity.[157]

C. *Rapid DNA*

1. How it Works

Rapid DNA machines fully automate the process of forensic DNA analysis, which traditionally requires that a trained forensic scientist perform physical chemical tests to extract DNA from a source sample and generate a profile.[158] A profile is determined by counting the amount of times a DNA sequence, called an allele, repeats at a specific section of the human genome, called a locus.[159] Alleles come in pairs—you get one DNA fragment length from each parent at a given locus.[160] So if a person shows a 5,8 allele pattern at the D3 locus, this means that the person the suspect has five repeats of a known sequence on one of the chromosomes at locus D3, and eight repeats on its partner chromosomes at locus D3.[161] Profiles for law enforcement use typically count alleles at either thirteen or twenty specified loci.[162]

Rapid DNA machines are designed to generate DNA profiles from samples without any human intervention. The person running the machine simply needs to place a sample into a cartridge and press a button. The machine's subsystems work to take the sample and spit out a profile in roughly ninety minutes.[163] Machines are connected to

---

[156] *See id.*

[157] *See* Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 PROC. OF MACH. LEARNING RSCH. 1–2 (2018).

[158] *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *supra* note 62, at 69.

[159] *See* ERIN E. MURPHY, INSIDE THE CELL: THE DARK SIDE OF FORENSIC DNA 7–8 (2015) (describing the original 13 core loci); *see also DNA Amplification for Forensic Analysts: Other STR Loci*, NAT'L INST. OF JUST. (July 31, 2013), https://nij.ojp.gov/nij-hosted-online-training-courses/dna-amplification/locus-selection/other-str-loci (describing the expansion of the core loci for profiles uploaded to FBI databases to 20).

[160] *See* MURPHY, *supra* note 159, at 7–8.

[161] *See id.*

[162] *See id*.

[163] *See* Allen Slater, *Policing Project Five Minute Primers: Rapid DNA*, POLICING PROJECT (Jan. 24, 2020), https://www.policingproject.org/news-main/2020/1/23/policing-project-five-minute-primers-rapid-dna; *see also Guide to All Things Rapid DNA, Version 1.1*, FED. BUREAU OF INVESTIGATION, 12 (Feb. 6, 2023), https://le.fbi.gov/file-repository/rapid-dna-guide-january-2022.pdf/view (describing rapid DNA as "developing a DNA

computer software to display the results of the analysis and upload the profile to connected databases.[164]

Rapid DNA machines can automatically upload profiles to DNA databases, and search algorithms will compare the uploaded profile to profiles already in the system by analyzing concordance in the core loci to see if there is a match. The Federal Bureau of Investigation (FBI) maintains a national system called the Combined DNA Index System (CODIS), which is the software program used to search DNA profiles contained at three database levels: the National Database Index System, the State DNA Index System, and the Local DNA Index System.[165] The databases contain samples from known contributors, as well as unidentified profiles generated from crime scenes.[166]

Under the Rapid DNA Act of 2017, rapid DNA machines can connect to CODIS if the instrument is one "approved by the Director of the Federal Bureau of Investigation" and is used "in compliance with the standards and procedures issued by the Director."[167] Currently, the FBI approves two rapid DNA machines for use: the ANDE 6C Series G and the ThermoFisher RapidHIT™ ID DNA Booking System v1.0.[168] The Bureau only allows rapid DNA profiles from these machines to be uploaded to CODIS in two cases: (1) when a single-source mouth swab is processed from an arrestee at an approved booking station and (2) when an accredited forensic laboratory generates a profile from a single-source mouth swab.[169] In order to upload these samples, laboratories and approved booking stations need to follow quality assurance standards.[170] The FBI has promulgated

---

profile . . . without the need for a DNA laboratory and without any human interpretation.").

[164] *See* ThermoFisher Scientific, *Rapid, Rapid, Rapid! What makes Rapid DNA Technology So Rapid? | Uninhibited with Peterjon and Nic*k 4:15, YOUTUBE (Sep. 22, 2021), https://www.youtube.com/watch?v=KXAUn_Z4fak.

[165] *See Frequently Asked Questions on CODIS and NDIS*, FED. BUREAU OF INVESTIGATION, https://www.fbi.gov/how-we-can-help-you/dna-fingerprint-act-of-2005-expungement-policy/codis-and-ndis-fact-sheet (last visited Oct. 22, 2023); *see also* MURPHY, *supra* note 159, at 14.

[166] *See* MURPHY, *supra* note 159, at 14–15.

[167] Rapid DNA Act of 2017, Pub. L. No. 115-50, 131 Stat. 1001 (2017).

[168] *See* MURPHY, *supra* note 159, at 13.

[169] *See Rapid DNA*, FED. BUREAU OF INVESTIGATION, https://le.fbi.gov/science-and-lab/biometrics-and-fingerprints/codis/rapid-dna (last visited Oct. 22, 2023).

[170] *See generally Standards for the Operation of Rapid DNA Booking Systems by Law Enforcement Booking Agencies*, FED. BUREAU OF

standards for rapid DNA use that law enforcement agencies are required to follow, such as machine performance checks and auditing, to ensure that samples are processed properly and can be uploaded to CODIS.[171]

The FBI does not allow crime scene samples processed on rapid DNA machines to be uploaded to CODIS, recognizing that "crime scene samples can present challenges for current Rapid DNA Technology."[172] And "[c]rime scene samples can vary widely, from age, exposure, and characteristics regarding the amount and quality of DNA. . . . [C]rime scene samples often contain mixtures of DNA from more than one person which requires interpretation by a trained scientist."[173]

However, law enforcement agencies are free to process samples, including complex crime scene samples, under less restrictive—and in some cases, nonexistent—protocols, so long as they do not upload the profile to CODIS. Instead, agencies upload profiles for search against state or local databases that need not comply with federal rules.[174] For example, the Connecticut Division of Scientific Services, a statewide agency, allows law enforcement officers to access a RapidHIT DNA kiosk to process crime scene samples; profiles generated from the agency's machine are uploaded and searched against the local "SmallPond" database.[175]

INVESTIGATION (Sep. 1, 2020), https://le.fbi.gov/file-repository/standards-for-operation-of-rapid-dna-booking-systems-by-law-enforcement-booking-agencies-eff-090120.pdf/view; *see also Rapid DNA*, *supra* note 169, at 12.

[171] *See Rapid DNA*, *supra* note 169, at 5–11.

[172] *Id.*

[173] *Id.*

[174] *See, e.g.*, Andrea Roth, *"Spit and Acquit": Prosecutors as Surveillance Entrepreneurs*, 107 CALIF. L. REV. 405, 408, 423–26 (2019) (describing the Orange County District Attorney's Office 150,000 profile DNA database, which operates without most of the privacy and quality assurance safeguards of CODIS).

[175] *See Blood, Sweat, and New Leads: Connecticut's Rapid DNA Program Generates Investigative Leads and Helps Solve Crimes Faster*, THERMOFISHER SCI. (2022), https://assets.thermofisher.com/TFS-Assets/GSD/Reference-Materials/hid-rapidinvestigativelead-casestudy.pdf. SmallPond is a privately owned software product that helps agencies generate their own DNA databases free from the federal regulations of CODIS. *See* Jason Kreag, *Going Local: The Fragmentation of Genetic Surveillance*, 95 B.U. L. REV. 1491, 1503 (2015).

2. Reliability

According to a NIST test of the three industry-standard processors—the ANDE 6C System, the ThermoFisher RapidHIT 200, and the ThermoFisher RapidHIT ID—rapid DNA machines are generally reliable at generating accurate DNA profiles from single-source, high quality samples.[176] But the machines are not free from error: the NIST results indicated that the rapid DNA processors were prone to consuming DNA samples before generating full profiles.[177] One of the machines failed to generate a full profile 55% of the time, due to failure of its allelic analysis subsystem.[178] These errors are dangerous: the preservation of DNA can be essential to solving crimes or exculpating the wrongfully accused.

The machines tested in the NIST study were stationed in controlled environments in law enforcement agencies.[179] But the storage of rapid DNA machines outside of controlled environments and the collection of samples at crime scenes increases the risk of contaminating samples.[180] And when police officers, rather than trained forensic analysts, are the personnel tasked with collecting and processing samples on rapid DNA machines, there is a heightened risk that best practices in DNA collection, preservation, and processing will not be followed.[181] The FBI requires operators of rapid DNA machines at booking stations to complete training and follow protocols that mitigate contamination risks.[182] But law enforcement agencies that

---

[176] *See* Erica L. Romsos et al., *Results of the 2018 Rapid DNA Maturity Assessment*, 65 J. FORENSIC SCI. 953, 956 (2020) (testing the machines in generating profiles with the 20 core CODIS loci). The NIST study does not foundationally validate the machines, however, as it fails on the PCAST requirement of independence. *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *supra* note 62, at 53. One of the authors worked for ANDE. *See id.* at 6.

[177] *See id.* (indicating that of 240 samples in the trial, the machines lost 25 without generating profiles).

[178] *See id.* at 955.

[179] *See id.* at 954.

[180] *See* Slater, *supra* note 163.

[181] *See* Peter Stout, *Caution Is Necessary When Expanding Field Testing Capabilities*, NAT'L INST. OF JUST. (Jan. 5, 2021), https://nij.ojp.gov/topics/articles/caution-necessary-when-expanding-field-testing-capabilities (noting difficulty deploying protocols for field collection of DNA to be rapid-tested).

[182] *See Standards for Operation*, *supra* note 170, at 5–11. The protocols also require that the DNA machines be stationary, eliminating the effects of shocks, vibrations, and temperature fluctuations on its hardware, which can pose real risks to sample processing. *See id.* at 5; *see also* Erik Dalin et al.,

process samples outside of the FBI's purview do not have to follow these standards. For those agencies, training and policy governing the use of rapid DNA machines is scant and unstandardized.[183]

Further, when rapid DNA machines process samples that are not from single-source buccal swabs, they are shockingly error-prone. In one study, machines correctly processed simulated crime scene samples at rates as low as 5%, indicating that profiles generated from such samples are likely to be false, or that DNA is likely to be consumed in the process.[184]

The profile generated by the machine is just the tip of the iceberg. Genetic profile matching entails its own set of reliability concerns, which are beyond the scope of this Article.

D. *Consequences of Police Use of Algorithmic Technologies*

Police use of pseudo-scientific algorithmic technologies increases false arrests and pretextual stops. The outputs of unreliable technologies do not lead police to credible suspects for serious crimes; instead, they provide false pretext to surveil and detain residents of already marginalized communities.[185]

In Chicago, the pervasive use of ShotSpotter justifies a practice of stop-and-frisks with many collateral consequences: the city's programmatic deployment of officers in response to alerts increases feelings of victimization and surveillance in Black and Latinx neighborhoods, heightening community hostility toward the CPD.[186] These stop-and-frisks likely result in increased arrests for non-gun-related minor crimes, impacting housing and employment opportunities for community members, with negligible benefit to public safety.[187]

Rapid DNA: A Summary of Available Rapid DNA Systems, SWEDISH NAT'L FORENSIC CTR. 15, 24 (2022) (describing machine sensitivity).

[183] *See* Slater, *supra* note 163.

[184] *See* Dalin et al., *supra* note 182, at 14.

[185] For example, the Chicago police conducted a conservative 2,400 investigative stops because of ShotSpotter alerts between January 2020 and May 2021. These stops rarely yielded a gun or a gun violence-related arrest. *See* OFF. OF THE INSPECTOR GEN., *supra* note 86, at 3, 16–21. The stops did, however, precipitate frisks that led to arrests for nonviolent crimes. *See id.* at 16, 18.

[186] *See* MacArthur Justice Center, *The Burden on Communities of Color*, *supra* note 2.

[187] *See Stop and Frisk: The Human Impact*, CTR. FOR CONST. RTS. 7–9 (July 2012), https://ccrjustice.org/sites/default/files/attach/2015/08/the-

Wholly pretextual stops are an important consequence of algorithmic technologies, but far from the only consequence. Though we think of technology as amplifying the police ability to solve target crimes like shootings or robberies, the opposite may be true. Anecdotal evidence is clear that when police have access to algorithmic investigative aids, they over-rely on the algorithms' ability to identify both crimes and suspects, forgoing other investigations. The result, thus far, is several bungled investigations that used either ShotSpotter or facial recognition technology to falsely arrest people of color. The mistakes made with each technology warn of potential pitfalls in future and unexamined past cases.

Cases in which police rely on ShotSpotter and facial recognition follow similar patterns. In one notable case, CPD officers arrested sixty-three-year-old Michael Williams on murder charges based solely on a ShotSpotter alert.[188] Williams was driving a young man from his neighborhood home when the man was shot through his car window.[189] After Williams took the young man to the local hospital, police fingered him for the man's killing—all because a corresponding ShotSpotter alert pinned the shot as fired from inside Williams's car.[190] Relying on the alert, the police declined to further investigate the murder, ignoring evidence that tended to show the shot did not come from inside the car, and failing to pursue credible leads to other suspects.[191] They also ignored disclaimers from SoundThinking itself: namely, that ShotSpotter is poor at locating indoor gunfire, and that the system's localization is "accurate" within a *radius*: the pinpoint the system sends to law enforcement is merely a suggestion.[192] Prosecutors dropped the charges against Williams when they were due to respond to reliability challenges against ShotSpotter.[193] By that time, Williams had languished in pretrial detention for eleven months, at the height of the COVID-19 pandemic.[194]

---

human-impact-report.pdf (collecting stories of people impacted after arrests pursuant to stop-and-frisks in New York City).

[188] *See* First Amended Complaint, *Williams v. City of Chicago*, No. 1:22-cv-03773 at 3–4.

[189] *See id.*

[190] *See id.*

[191] *See id.* at 3–4, 51–62.

[192] *See id.* at 3–4.

[193] *See id.* at 61–62.

[194] *See* First Amended Complaint, *supra* note 188 at 65. ShotSpotter is especially dangerous because it has the potential to lead police to mistarget not only suspects, but also crime. For example, CPD officers "stopped, frisked, handcuffed, interrogated, and ultimately arrested" Daniel Ortiz based on a ShotSpotter alert, even though the officers found no evidence that a gun

Michael Williams's ordeal is eerily like that of Robert Julian-Borchak Williams, who was also detained on a tip from algorithmic technology: this time, a facial recognition match. Robert Williams was arrested by the Detroit Police Department (DPD) in 2019, when his old driver's license showed up as a match to a shoddy probe photo—an obscured still taken from a surveillance video of a shoplifter in a Shinola store.[195] The algorithm used by the DPD to match the still to Robert Williams's license exhibited significant error in identifying Black people.[196] DPD Chief James Craig later said that "[i]f we were just to use the technology by itself, to identify someone . . . 96 percent of the time it would misidentify."[197] To that end, the Michigan State Police form that identified Robert Williams as a purported match to the probe photo disclaimed that it was "NOT A POSITIVE IDENTIFICATION," only an investigative lead, and that further investigation was needed to develop probable cause.[198] Still, just as CPD detectives ignored SoundThinking's disclaimers, the DPD detectives ignored their own protocols, and pursued Robert Williams for the shoplifting charge with negligible follow-up investigation.[199] It did not take long for the case to fall apart: when Robert Williams was interrogated in custody, he told the DPD that he was not the man in their surveillance photo.[200] The officers agreed, and charges against Williams were dismissed a few weeks later.[201] Still, the DPD made the same mistakes in two other cases, those of Michael Oliver,[202] arrested months before Robert Williams, and Porcha Woodruff, arrested over

---

was even fired in the vicinity of the alert, let alone that Ortiz was the shooter. *Id.* at 5, 68–74.

[195] *See* Complaint, *Williams v. City of Detroit*, No. 2:21 Civ. 10827 at 16–20.

[196] *See* Joh, *Reckless Automation*, supra note 57, at 128.

[197] Complaint, *Williams v. City of Detroit*, No. 2:21 Civ. 10827 at 3.

[198] *Id.* at 20.

[199] After the DPD received the facial recognition match to Robert Williams, they set out to "confirm" the match. However, they only did so by showing a non-witness Shinola store employee a photo array with Mr. Williams's license photo. The employee in question had never seen the shoplifter in the store, she made her identification by comparing Williams's expired license photo to the surveillance footage she had reviewed, just as the facial recognition program had. *See id.* at 23–24.

[200] *See id.* at 36–37.

[201] *See id.*

[202] *See* 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/.

three years after Robert Williams and Oliver.[203] Other facial recognition cases in Woodbridge, New Jersey;[204] Jefferson Parish, Louisiana;[205] and Baltimore County, Maryland[206] bear similar hallmarks: people arrested on faulty FRT matches easily dispelled by analog investigation. In all these cases, there were easily observable differences between the arrestees and suspects on surveillance footage.[207] All of the falsely arrested victims are Black.

These false arrests are emblematic of how law enforcement's blind reliance on unreliable technology causes violence to members of over-policed communities. Far from an individual effect, use of algorithmic technology influences the way that agencies approach policing in the aggregate. For example, ShotSpotter technology has "has changed the way CPD members perceive and interact with individuals present in areas where ShotSpotter alerts are frequent."[208] Officers have justified stop-and-frisks on the number of *aggregate* past alerts in an area,[209] meaning even false alerts can increase the perceived criminality of a neighborhood. Alert data feeds back into the ShotSpotter algorithm[210] and CPD's predictive policing tools,[211] making neighborhoods with ShotSpotter sensors appear more

---

[203] *See* Kashmir Hill, *Eight Months Pregnant and Arrested After False Facial Recognition Match*, N.Y. TIMES (Aug. 6, 2023), https://www.nytimes.com/2023/08/06/business/facial-recognition-false-arrest.html.

[204] *See* Johnson, *Wrongful Arrests*, supra note 202.

[205] *See* Kashmir Hill & Ryan Mac, *'Thousands of Dollars for Something I Didn't Do'*, N.Y. TIMES (Apr. 6, 2023), https://www.nytimes.com/2023/03/31/technology/facial-recognition-false-arrests.html.

[206] *See* Khari Johnson, *Face Recognition Software Led to His Arrest. It Was Dead Wrong*, WIRED (Feb. 28, 2023), https://www.wired.com/story/face-recognition-software-led-to-his-arrest-it-was-dead-wrong/.

[207] *See, e.g.*, Johnson, *Wrongful Arrests*, *supra* note 202 ("Oliver has visible tattoos, the suspect in the surveillance footage from which he was identified did not.").

[208] OFF. OF THE INSPECTOR GEN., *supra* note 8, at 3.

[209] *See id.* at 19–21.

[210] *See* Calhoun et al., *supra* note 50, at 8 (describing how the ShotSpotter algorithm accounts for past alerts).

[211] *See* Special Order S03-19, *ShotSpotter Flex Program Directive*, CHICAGO POLICE DEP'T (July 5, 2017), http://directives.chicagopolice.org/#directive/public/6138 (directing districts to include ShotSpotter alerts in their Compstat statistical summaries).

dangerous, and exacerbating discriminatory policing in the majority Black and Latinx neighborhoods where the sensors are located.[212]

Crucially, law enforcement agencies like the CPD show a propensity to rely on ShotSpotter as a crime-fighting and data-collection tool with little knowledge of its true reliability. Rapid DNA, though nascent, is poised to similarly influence street policing. Rapid DNA is most consequential for its ability to expand DNA testing and genetic profile storage. With the advent of this faster, automated, DNA processing, law enforcement agencies will adjust their practices to collect and store DNA profiles from everyone they encounter. Professor Erin Murphy posits that officers will take advantage of laws that authorize DNA collection from arrestees, using their discretion to adjust charges to meet statutory requirements as to allow collection.[213] The availability of rapid DNA has already led officers to collect DNA swabs from people they do not arrest, at traffic stops and in stop-and-frisks.[214] These swabs are collected with "consent," but the

---

[212] *See* Brief for Brighton Park Neighborhood Council, *State v. Williams*, 20 CR 0899601 at 13–16. This increase in police presence leads to violence. When officers respond to ShotSpotter alerts, they admit to behaving "tactical[ly]," expecting gunfire. *See* Nick Selby et al., *ShotSpotter Gunshot Location System Efficacy Study*, CSG ANALYSIS at 21 (2011), https://njdc.info/wp-content/uploads/2017/10/Shot-Spotter-Gunshot-Location-System-Efficacy-Study.pdf. Arguably, this tactical response led to the murder of 13-year-old Adam Toledo, who was pursued by a CPD officer following a ShotSpotter alert in March 2021. *See* Rachel Treisman et al., *Chicago Releases Video Showing Fatal Police Shooting Of 13-Year-Old Adam Toledo*, NPR (Apr. 15, 2021), https://www.npr.org/2021/04/15/987718420/chicago-releases-video-showing-fatal-police-shooting-of-13-year-old-adam-toledo.

[213] *See* MURPHY, *supra* note 159, at 156–57, 162 (noting that most states have post-arrest DNA collection statutes that allow law enforcement officers to collect samples from those arrested for violent offenses, while some states allow collection from those arrested for lesser felonies or misdemeanors); *see also Maryland v. King*, 569 U.S. 435, 465–66 (2013) (holding that the Fourth Amendment does not prohibit law enforcement agencies from taking and storing the DNA of individuals arrested for serious offenses).

[214] For example, in Bensalem, Pennsylvania, officers have a practice of gathering DNA swabs from arrested individuals with the object of randomly selecting a few swabs each week to run through a rapid DNA machine. Since getting a rapid DNA machine, the agency has also adopted a practice of detaining "suspicious subject[s]" to swab and run their DNA. *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.

voluntariness of this consent is questionable—as Professor Erin Murphy notes, individuals already stopped by police may feel pressured to give a sample simply to "avoid the hassle of arrest for [things like] trivial [traffic] charges."[215] Some law enforcement agencies will condition charges and plea deals on individuals' willingness to give DNA samples.[216]

The effect of this sampling—and how it relates to reliability—lies in *why* law enforcement agencies seek to build large DNA databases. Put simply: to solve crimes.[217] The government has every incentive to take DNA when it can, and to seek to match identified database profiles to unknown suspects and crime scene samples. A growing bank of profiles means more potential for match error, and faulty matches have attendant effects.[218] District attorneys intend to use rapid DNA results to facilitate plea negotiations, using matches to force pleas.[219] Defendants may be encouraged to take pleas based off faulty profiles, without interrogating rapid DNA's reliability.[220] More DNA collection also just means that the government has your DNA, facilitating the "genetic panopticon" Justice Scalia feared when he dissented in *Maryland v. King*.[221] Solving crime is good, but the list of uses for DNA is long: it can be used in paternity claims, to extract health information, and to trace people, including law-abiding citizens.[222] There are reasons to be uneasy about ever-expanding DNA databases.

The availability of rapid processing will lead to more pretextual stops and more DNA collection—with or without meaningful consent. Law enforcement agencies seek to collect more DNA for the ostensibly the same reasons they seek to deploy more ShotSpotter sensors and more FRT. In theory, more data keeps us safer. But algorithmic technologies are unreliable. Their use is targeted

---

[215] MURPHY, *supra* note 159, at 165–66.

[216] *See* Roth, *supra* note 174, at 416.

[217] *See* MURPHY, *supra* note 159, at 159.

[218] *See* Erin Murphy, *DNA in the Criminal Justice System: A Congressional Research Service Report* (*From the Future)*, 64 UCLA L. REV. DISC. 340, 352–53 (2016). The risk of wrongful arrests and convictions from false matches is especially high when law enforcement agencies upload profiles derived from crime scene data.

[219] *See* Emily Blume, *Rapid DNA Testing to Solve More Local Crime Coming Soon*, KXLY (Jan. 13, 2023), https://www.kxly.com/news/crime/rapid-dna-testing-to-solve-more-local-crime-coming-soon/article_8fe21331-9a27-5cef-832e-db839f11dea3.html.

[220] *See* MURPHY, *supra* note 159, at 166–67.

[221] *See King*, 569 U.S. 480 (Scalia, J., dissenting).

[222] *See* MURPHY, *supra* note 159, at 173.

toward general surveillance rather than competent criminal investigation. So, they do more harm than good to public safety.[223]

Courts have yet to consider the proper use of rapid DNA in criminal investigations, as they "rarely, if ever" consider the role of facial recognition technology in developing probable cause.[224] ShotSpotter, facial recognition, and rapid DNA are not meaningfully screened for reliability at any point: not before deployment, not at probable cause hearings, and not at trial.

IV. THE PROBLEM: ALGORITHMIC TECHNOLOGIES ARE NOT WELL SCREENED FOR RELIABILITY AS INVESTIGATIVE TOOLS

A. *The Front-End Problem*

The failures of an algorithmic technology may lie dormant to the public until a faulty arrest is made and the technology is challenged in court—a realization of "back-end accountability"—remedying a wrong already done.[225] Yet, the problem of unreliable technology starts when police acquire it. Such acquisitions involve little external oversight or democratic input, what Professor Barry Friedman has named as a lack of "front-end accountability" to the policed public.[226] Law enforcement agencies fail to validate acquired technologies, set public policies for their use, and audit whether the technologies are accurate in investigations.[227] These front-end failures foreclose opportunities to reduce error or jettison faulty technologies before unleashing them on the public.

Instead, the police disregard reliability deficits in the technology they acquire. Agencies purchase technologies because of

---

[223] *See, e.g.*, Mitchell L. Doucette et al., *Impact of ShotSpotter Technology on Firearm Homicides and Arrests Among Large Metropolitan Counties: A Longitudinal Analysis*, 1999–2016, 98 J. URB HEALTH, 609, 617–18 (2021) (longitudinal analysis on ShotSpotter's effect on gun violence finding that that presence of the sensors had no effect); *see also* MacArthur Justice Center, *The Burden on Communities of Color*, *supra* note 2 (finding that ShotSpotter deployments decreased 911 call response times in Chicago neighborhoods).

[224] T.J. Benedict, Note, *The Computer Got It Wrong: Facial Recognition Technology and Establishing Probable Cause to Arrest*, 79 WASH. & LEE L. REV. 849, 866 (2022).

[225] *See* Friedman, *supra* note 146, at 2.

[226] *Id.*

[227] In a nationwide survey, The Georgetown Center on Privacy Law found only two law-enforcement agencies that conditioned the purchase of facial recognition programs on the technology's accuracy. *See* GARVIE ET AL., *supra* note 115, at Scorecard, Appendix: Methodology, Accuracy.

their *perceived* ability to root out crime more efficiently.[228] This may be why, when technologies are tested, agencies only assess false negatives rather than false positives.[229] Law enforcement may even prefer that a technology has a higher false positive rate, because the trade-off is that the technology has a lower false negative rate, casting a too-wide net in searching for criminal activity.[230] In service of that aim, Chicago's contract with SoundThinking incentivizes the overreporting of sounds as gunfire.[231] The contract promises ShotSpotter will alert to at least 90% of outdoor, unsuppressed gunshots fired from greater than .25 caliber weapons inside the coverage area, but makes no corresponding guarantee to keep false alerts below a certain threshold.[232]

Absent independent testing, agencies may not have adequate information about the accuracy of algorithmic technologies. Private companies market the technologies on overstated, misleading, or false accuracy rates. For example, SoundThinking claims ShotSpotter has an accuracy rate of 97% and a false positive rate of 0.5%, but this is based on anecdotal customer reports.[233] SoundThinking assesses a false positive only when a customer reports an error, which does not reflect actual false positive rates.[234] Similarly, facial recognition marketing

---

[228] *See City Defends Quiet Contract Extension for ShotSpotter Gunfire Detection System as Residents Complain*, CBS NEWS (Oct. 4, 2021), https://www.cbsnews.com/chicago/news/city-defends-shotspotter-contract-extension (describing that Chicago set efficiency benchmarks for ShotSpotter performance without addressing false positives).

[229] *See supra* Section III.A.2 (describing how ShotSpotter has only been tested for its ability to recognize live fire).

[230] Amidst outcry about false positives, Chicago Police Superintendent David Brown defended ShotSpotter, stating "[i]f one life is saved, we should keep that tool in our toolbox." *Id*. Former Superintendent Brown's response to reliability concerns exemplifies law enforcement's willingness to accept false positives as an expense of less false negatives.

[231] *See* Brief for Roderick & Solange MacArthur Justice Center, et al., *Ford*, No. 2020-P-1334 at 21.

[232] *See id.* (citing Chicago Police Dep't, City Contract No. 71366, Area Acoustic Gun Shot Detection Subscription Service at 95–96, 99 (Aug. 22, 2018), http://ecm.chicago.gov/eSMARTContracts/service/dpsweb/ViewDPSWeb.zul.).

[233] *See Independent Audit of the ShotSpotter Accuracy*, EDGEWORTH ECON. 2 (Mar. 28, 2022), https://www.edgeworthe conomics.com/assets/htmldocuments/Shotspotter-2022-Accuracy-Study.pdf ("[I]nformation on potential errors relies on clients reporting those potential errors to ShotSpotter.").

[234] *See id*.

materials include statements about accuracy that are hard to verify and may not be based on real-world use.[235]

Other marketing tactics also disguise reliability issues. In the wake of backlash, SoundThinking directed its law enforcement clients to endorse ShotSpotter technology and its "positive impact" in the media, hoping to obscure negative press.[236] SoundThinking uses a variety of strategies to capture agencies' loyalty, including providing police departments with assistance in applying for federal grants to pay for its technology.[237] Several police technology companies use grant assistance as a marketing tactic, including ThermoFisher.[238] Activists say federal grants alter cost-benefit analyses and make it easier for agencies to ignore flaws in subsidized technology, because if the agency does not have to pay for it, "[i]t doesn't matter if the stuff works or not."[239]

Private ownership allows technology companies to hide data, obscuring efforts to interrogate reliability at the front end. Law enforcement agencies must contract with private companies to buy algorithmic technologies, and through these purchaser/seller contracts, companies can place restrictions on the data to which the law-enforcement agency is entitled.[240] Companies that peddle police technologies keep the specifics of their algorithms secret from the public and contracting agencies, citing trade secret privileges.[241]

The lack of transparency attendant to private ownership of police technologies renders it imperative that law enforcement

---

[235] *See* GARVIE, *supra* note 60, at 20.

[236] *See* Jon Schuppe & Joshua Eaton, *How Shotspotter Fights Criticism and Leverages Federal Cash to Win Police Contracts*, NBC NEWS (Feb. 10, 2022), https://www.nbcnews.com/news/us-news/shotspotter-police-gunshot-technology-federal-grants-rcna13815. Other companies also employ a "success story" approach. *See, e.g.*, THERMOFISHER SCI., *supra* note 175 (promotional material from ThermoFisher describing the Connecticut state crime laboratory's use of rapid DNA technology to generate investigative leads).

[237] *See* Schuppe & Eaton, *supra* note 236.

[238] *See, e.g.*, *Rapid DNA Solutions for Crime Laboratories*, THERMOFISHER SCI. (2023), https://www.thermofisher.com/us/en/home/industrial/forensics/human-identification/forensic-dna-analysis/dna-analysis/rapidhit-id-system-human-identification/rapidhit-id-system-crime-labs.html (ThermoFisher rapid DNA funding assistance page) (last visited Oct. 22, 2023).

[239] Schuppe & Eaton, *supra* note 236.

[240] *See* Joh, *supra* note 57, at 122; *see also* Goodman, *supra* note 57, at 802 (2021).

[241] *See* Joh, *supra* note 57, at 122.

agencies act independently to understand the reliability of systems they purchase. However, many agencies do not audit their use of algorithmic technologies,[242] and they generally fail to collect or publicize basic data on the use of investigative tactics and technological aids.[243] The Chicago OIG report represents a rare case in which data on a technology's use was examined, publicly and in detail.[244] However, in the wake of the report, which recommended that the city ditch ShotSpotter, city officials doubled down on their support of the system.[245] When MJC came out with findings on ShotSpotter false positives that echoed the OIG report, Chicago Mayor Lori Lightfoot questioned whether the research was "actually accurate,"[246] but there is no evidence that the CPD conducted follow-up testing of their own. Before the reports were published, Lightfoot extended the ShotSpotter contract without public notice or vetting.[247]

Despite known reliability issues, whether due to lack of knowledge or lack of care, law enforcement agencies fail to give officers clear policy guidance on the limitations of algorithmic technologies. For example, the Chicago ShotSpotter policy does not address ShotSpotter's reliability, or how an alert should be corroborated as an investigative lead.[248] It does not specify what circumstances, if any, an officer needs in addition to an alert for probable cause or reasonable suspicion.[249] Similarly, facial recognition and rapid DNA policies often lack guidance on how to corroborate leads gleaned from the technologies—if law enforcement agencies

---

[242] *See, e.g.*, GARVIE ET AL., THE PERPETUAL LINEUP, *supra* note 115 (documenting major police departments' failure to audit use of facial recognition technology).

[243] *See* Barry Friedman & Maria Ponomarenko, *Democratic Policing*, 90 N.Y.U. L. Rev. 1827, 1849 (2015) (describing agency refusals to disclose data on the use of SWAT teams and Stingray electronic surveillance technology).

[244] *See generally* OFF. OF THE INSPECTOR GEN., *supra* note 8.

[245] *See, e.g.*, Don Babwin & Garance Burke, *Chicago Watchdog Harshly Criticizes Shotspotter System*, AP NEWS (Aug. 24, 2021) https://apnews.com/article/technology-business-chicago-1d62906b0c4b4dc67886da89596b1f12 ("Lightfoot has weighed in as well, calling the technology . . . 'a lifesaver.'").

[246] Editorial Board, *If Shotspotter Constantly Misfires, What's Chicago Getting for its $33 Million?*, CHI. SUN-TIMES (May 4, 2021), https://chicago.suntimes.com/2021/5/4/22417660/shotspotter-analysis-macarthur-justice-center-chicago-police-chicago-gun-violence-editorial.

[247] *See* Schuba, *supra* note 10.

[248] *See generally* Special Order S03-19, ShotSpotter Flex Program Directive, CHICAGO POLICE DEP'T (July 5, 2017), http://directives.chicagopolice.org/#directive/public/6138.

[249] *See generally id.*

even create policies addressing the technologies in the first place.[250] In a stark example, the Bensalem police used rapid DNA machines—accompanied by a practice of collecting samples in random traffic stops—without any established policy at all.[251] Controverting FBI and scientific guidance, Bensalem police processed crime scene samples via rapid DNA machines.[252] These policies—or lack thereof—are typical of the sparse regulations that govern policing.[253] Loose policies, leave unfettered discretion in the hands of officers, and make it so investigative technologies have unexpected and detrimental consequences.[254]

The policed public is most affected by these front-end failures. In a phenomenon she calls "reckless automation," Professor Elizabeth Joh identifies how emerging technologies are deployed as "technological experiments" against populations that are already overpoliced: low-income Black and Brown communities.[255] These new technologies impact communities in tangible ways: they increase surveillance, stops, and arrests.[256] But, due to a lack of transparency as to how, or even if, these technologies are being deployed, the affected public is unable to scrutinize the technologies and challenge their use, leaving the police unaccountable.[257]

Acquisition of police technologies is typically an insular process, free from democratic input from both lawmakers and the public. The New York Police Department has purchased and deployed

---

[250] *See* GARVIE, *supra* note 60, at 6.

[251] *See* Murphy, *Coming Soon to a Police Station Near You*, *supra* note 214.

[252] *See id.*

[253] *See* Friedman & Ponomarenko, *supra* note 243, at 1831–32, 1843, 1845 ("[Policing] manuals are often silent on critical aspects of policing . . . [there may be no rules] on informants, drones, consent searches, or other investigative tactics.").

[254] *See* Joh, *Unexpected Consequences*, supra note 47, at 523–24 (describing how officers in Chicago used ShotSpotter contrary to its marketing by using "aggregate alerts" in a given area as a pretext for investigative stops).

[255] *See* Joh, *Reckless Automation*, supra note 57, at 118 (identifying the development of "reckless automation in policing"—the procurement of experimental technologies that "impact communities through increased but invisible surveillance, and with mistakes that impose real-life consequences in police civilian interactions.").

[256] *See id.*

[257] *See id.* at 126–28 (citing Chicago's "Heat Risk" gun violence prediction program and Detroit's facial recognition system as examples of experimental automated decision-making technologies that, due to unforeseen unreliability issues, imposed unnecessary harms like harsher sentencing, harsher charging, and false arrests).

several algorithmic technologies, including ShotSpotter and facial recognition, "while attempting to keep the public and the City Council in the dark," according to a report from the Brennan Center.[258] In Chicago, despite voiced reliability concerns, the ShotSpotter contract has been extended twice, secretly and unilaterally, by former Mayor Lightfoot.[259] Professors Friedman and Ponomarenko write that this "shroud of secrecy" around decision-making is common in policing.[260] And it is detrimental to the policing enterprise and to the public. There is no apparent justification for shielding the acquisition of technologies and data on their use from the public, other than to avoid public backlash—public disclosure should not affect police efficacy.[261] But the lack of transparency around policing decisions renders the police unaccountable on the front end. Secrecy and disregard for public opinion sows distrust in the police.[262] By contrast, when police hold themselves accountable and listen to public feedback when making decisions, they end up with more cogent and responsive policy.[263]

Law enforcement agencies fail to ensure that the algorithmic technologies they use are reliable. They do not sufficiently test technologies for false positives, nor audit their use. They fail to disseminate policies that could curb the consequences of relying on unreliable technologies. They are untransparent about technology acquisition, foreclosing or disregarding public input about whether a technology is fit for use. Law enforcement agencies may leave reliability screening in the hands of the courts, but as I discuss in Sections IV.B and IV.C, the courts are also unequipped to properly evaluate algorithmic technologies.

---

[258] Angel Diaz, *New York City Police Department Surveillance Technology*, BRENNAN CTR. FOR JUST. 1–2, 10 (Oct. 7, 2019), https://www.brennancenter.org/our-work/research-reports/new-york-city-police-department-surveillance-technology.

[259] *See* Schuba, *supra* note 10.

[260] *See* Friedman & Ponomarenko, *supra* note 243, at 1848.

[261] Friedman and Ponomarenko argue that secrecy is only justified when it pertains "both to specific investigations and investigative techniques that, if public, would encourage circumvention." *Id.* at 1884–85. While the disclosure of specific ShotSpotter sensors or surveillance cameras may encourage circumvention, broad disclosure that the police deploy such technologies would not. Data disclosure also has little impact on policing, other than to incur greater public scrutiny of police conduct. *See id.* at 1886.

[262] *See infra* Section V.B for further discussion on how police reliance on unreliable technologies sows distrust in law enforcement.

[263] *See* Friedman & Ponomarenko, *supra* note 243, at 1848, 1852–53, 1879–81 (citing public-influenced police practice and policy changes that resulted in better crime control and enhanced public trust).

B. *The Pre-Trial Problem*

A criminal defendant can seek to exclude evidence from the government's case-in-chief, arguing that such evidence was obtained in violation of the Fourth Amendment, via the filing of a pre-trial suppression motion.[264] Seeking the suppression of evidence recovered based on an antecedent seizure will often implicate an analysis of probable cause or reasonable suspicion. This, in turn, will implicate an analysis of algorithmic technologies when those technologies have contributed to the quantum of suspicion.

As I described in Section II.B, courts evaluate probable cause and reasonable suspicion on the "totality of the circumstances." The "totality of the circumstances" inquiry allows courts to ignore the import of algorithmic technology and bypass meaningfully assessing its reliability.

On motions to suppress grounded in a lack of probable cause or reasonable suspicion, courts will decline to assess the reliability of technology when adequate alternative grounds, or "circumstances" exist to support a stop. When evaluating the traffic stop of a criminal defendant, Terrill Rickmon, a Seventh Circuit panel expressed doubt about the reliability of ShotSpotter, but declined to evaluate its reliability.[265] Rickmon was stopped because he was the only driver in the vicinity of a ShotSpotter alert sent to the Peoria, Illinois police, five minutes prior.[266] The Circuit reviewed evidence that ShotSpotter was not always accurate, but it did not have information demonstrating the unreliability of the specific system in Peoria.[267] What it did have was information about the individual stop, including a concurrent 911 call, "the stop's temporal and physical proximity to the shots, the light traffic late at night, and the officer's experience with gun violence in that area."[268] Though the *Rickmon* court recognized that, "in isolation, any one of those circumstances might not be sufficient," as a whole, the panel majority concluded the circumstances amounted to reasonable suspicion, ostensibly independent of the ShotSpotter

---

[264] This is a simplified explanation of the well-known "exclusionary rule" and the mechanism to invoke it. *See* CLANCY, supra note 33, at 853–54, 863–65.

[265] *See United States v. Rickmon*, 952 F.3d 876, 881 n.2 (7th Cir. 2020).

[266] *See id.* at 879; *see also id.* at 885–86 (Wood, J., dissenting).

[267] *See id.* at 881 n.2 ("[T]he record here does not demonstrate how often the Peoria Police Department received incorrect ShotSpotter reports or anything else attesting to the reliability of the system.").

[268] *Id.* at 884.

alert.[269] Yet, by denying Rickmon's motion, approving of his seizure, and declining to pronounce ShotSpotter unreliable, the court implicitly blessed the officer's reliance on the alert.

That is because, contravening reliabilist theories, the *Rickmon* decision was ignorant of the role algorithmic technologies play in *developing* suspicion. The ShotSpotter alert, reliable or not, clouded the officer's judgment in *Rickmon*. As Judge Diane Wood noted in dissent, the alert was the only precipitating event that caused police to view Rickmon with suspicion.[270] After receiving the alert, the officer who stopped Rickmon "would have stopped literally any car he saw" in the corresponding address block.[271] Though there were independent "circumstances" governing the stop, they rose or fell on the information the officer received from the ShotSpotter alert. It was the alert that caused the officer to color everything else about his stop of Rickmon as suspicious.

The nature of the individualized suspicion inquiry renders a judge unable to see the forest for the trees. Suppression motions often present judges with cases in which a technology has succeeded on one measure: a defendant has been caught with contraband, or in otherwise incriminating circumstances (Rickmon was found with a gun, and convicted as a felon unlawfully in possession of a firearm).[272] With the suppression motion, the judge is presented with only one case, not the universe of false positives that may suggest a technology is unreliable.[273] Much of the vital information in these single cases is provided via officer testimony. In "totality of the circumstances" inquiries, courts often rely on officer assessments of a technology's reliability rather than conduct far-reaching inquiries on accuracy.[274] This envelopes an officer's automation bias into the totality. Bias issues are symptomatic of the way an unreliable technology may infect an entire seizure. Automation bias towards the output of the technology

---

[269] *Id.* at 884–85. Though the Circuit explicitly declined to pass judgment on ShotSpotter's reliability, at least two state court decisions cite *Rickmon* for the proposition that catching a person at the scene of a ShotSpotter alert, even after five-and-a-half minutes have passed, supports reasonable suspicion. *See State v. Nimmer*, 975 N.W.2d 598, 605-06 (Wis. 2022); *State v. Carter*, 183 N.E.3d 611, 629 (Ohio Ct. App. 2022).

[270] *See Rickmon*, 952 F.3d at 887 (Wood, J., dissenting).

[271] *Id.*

[272] *See Rickmon*, 952 F.3d at 879–80.

[273] *See* Friedman & Ponomarenko, *supra* note 243, at 1866.

[274] *See, e.g.*, *Nimmer*, 975 N.W.2d at 600, 605 (finding reasonable suspicion based, in part, on the fact that "ShotSpotter generates reliable reports of gunfire in near real-time" according to the officers who stopped Nimmer near the location of reported gunfire).

predisposes officers to see indicia of criminality that do not exist, thus, relying on an officer's assessment of accuracy imports this bias into a court's assessment of the "circumstances."

Courts are unable to meaningfully interrogate officers' assessments of the reliability of algorithmic technologies; they are impaired from properly assessing algorithmic technologies, because defendants lack the information necessary to present reliability issues. In some cases, a defendant may not even know that an algorithmic technology played a role in her arrest. Take, for example, the case of Randal Reid, a victim of a facial recognition misidentification of the sort discussed in Section III.D. Jefferson Parish police arrested Reid solely on a facial recognition identification, but the warrant for his arrest gave no indication that this was the case. The supporting affidavit cited only a "credible source."[275] Even Reid's lawyers were unable to confirm that facial recognition was used, it was the New York Times that eventually verified that a facial recognition program had identified Reid.[276] Charging documents often bury or disguise the use of facial recognition to affect an identification, and defense attorneys have to dig for evidence indicating that it was even used to mount a reliability challenge.[277] This is one reason why courts have had little occasion to consider facial recognition technology in probable cause inquiries.[278]

Defendants' inability to litigate reliability may also stem from a lack of knowledge about a technology's use in a specific case or jurisdiction. One reason the Seventh Circuit declined to reach the reliability issue in *Rickmon* was because it felt the evidence for the challenge was incomplete, since the record did not include information about ShotSpotter's rate of error in Peoria.[279] The private ownership of algorithmic technologies impedes defendants' ability to gather information about systems used in their cases, as it likely did in *Rickmon*. SoundThinking regards not just its algorithms, but its

---

[275] *See* Hill & Mac, *supra* note 205.

[276] *See id.*

[277] *See* Jennifer Valentino-DeVries, *How the Police Use Facial Recognition, and Where It Falls Short*, N.Y. TIMES (Jan. 12, 2020), https://www.nytimes.com/2020/01/12/technology/facial-recognition-police.html; *see also* Kaitlin Jackson, *Challenging Facial Recognition Software in Criminal Court*, 43 THE CHAMPION 14, 16 (2019).

[278] *See* KELSEY Y. SANTAMARIA, CONG. RSCH. SERV., R46541, FACIAL RECOGNITION TECHNOLOGY AND LAW ENFORCEMENT: SELECT CONSTITUTIONAL CONSIDERATIONS 19 (2020) ("Although investigatory officers have deployed FRT to identify suspects, a survey of case law suggests that courts have rarely considered probable cause challenges to police work that relied on purportedly unreliable FRT matches.").

[279] *See Rickmon*, 952 F.3d at 881 n.2.

gunshot data, as proprietary trade secrets. When cities "subscribe" to its technology by contracting with the company, they lease the data and are permitted to use it for the length of the contract.[280] The law enforcement agency does not own the data. The company's CEO emphasizes to agencies that ShotSpotter data should not be made publicly available, going so far as to issue a "nationwide memo" urging contracting agencies not to disclose data.[281] When a gunshot alert is relevant to a prosecution, contracting law enforcement agencies notify SoundThinking and an analyst from the company produces a report for use in court.[282] These reports serve as a record of the relevant alert or alerts, but the output noted in the report may be edited in "post-processing," when engineers purportedly correct errors in initial alert notifications.[283] Additional data from the jurisdiction is not disclosed.

SoundThinking has fought tooth-and-nail to keep any information extraneous to its post-processing reports out of court. In a 2022 Chicago case, they requested to be held in contempt of court, rather than disclose broader data about their system.[284] Defense counsel had asked for records including: (1) ShotSpotter analysts' qualifications and training materials; (2) any instances in which the company's analysts reclassified alerts or the Chicago police asked ShotSpotter to do so; (3) the methods analysts use to reclassify alerts; and (4) any data on sensors misidentifying gunfire or the location of alerts, as well as data on gunfire ShotSpotter failed to identify.[285] The company has fought to deny discovery in several cases, including those

---

[280] *See* Goodman, *supra* note 57, at 802 (describing ShotSpotter's data-leasing relationship with contracting agencies).

[281] *See* Hannah Bloch-Wehba, *Access to Algorithms*, 88 FORDHAM L. REV. 1265, 1284–85 (2020).

[282] *See* Testimony of Paul Greene, Senior Forensic Engineer at ShotSpotter, *California v. Reed*, No. 16015117 at 12 (Cal. Super. Ct. July 5, 2017).

[283] *See id.* at 12–13.

[284] *See* Matt Chapman & Jim Daley, *ShotSpotter Held in Contempt of Court*, CHI. READER (July 26, 2022), https://chicagoreader.com/news-politics/shotspotter-held-in-contempt-of-court/.

[285] *See id.*

of Michael Williams in Chicago[286] and Silvon Simmons[287] in Rochester, both of which featured doctored post-processing reports. SoundThinking has designed a system that keeps information relevant to reliability challenges out of defendants' hands.

Prosecutors also impede defendants' access to information about algorithmic technologies. In cases where facial recognition technology was used, prosecutors have denied public defenders the complete inputs and outputs of facial recognition searches that resulted in arrests.[288] This prevents defendants from challenging the circumstances of their identification. But facial recognition identifications are ripe for challenge when the initial match was linked to a low-quality probe photo, or the defendant came up as a low-confidence match on a list of candidates.[289] Defendants need broad discovery to understand algorithmic technologies and how they are used; this information is necessary to mount a good defense in pre-trial proceedings and during trial.

C. *The Trial Problem*

Trials also fail to provide an adequate mechanism to assess the reliability of algorithmic technologies, for three reasons. First, technologies used to establish probable cause and reasonable suspicion may not be presented at trial. Second, algorithmic technologies may have programmatic functions, such as widespread surveillance, that do not result in Fourth Amendment intrusions and are therefore immunized from judicial scrutiny. Third, when prosecutors do seek to admit evidence drawn from algorithmic technologies at trial, courts fail to adequately assess their reliability under existing evidentiary standards.

---

[286] After a protracted battle, ShotSpotter sent one document in response to the multiple subpoenas filed in Michael Williams's case. The 19-page document, a training guide for human analysts, was provided under a protective order in the case, and later obtained by the Associated Press. *See* Burke, *supra* note 107. There are other training materials that the company deems "confidential and trade secret." *Id*.

[287] *See* Reade Levinson & Lisa Girion, *A Black Man Risks All to Clear His Name - and Expose the Police*, REUTERS (Nov. 17, 2020), https://www.reuters.com/article/usa-police-rochester-trial-special-repor-idINKBN27X1VO ("Simmons' defense team had sought . . . original recordings and other ShotSpotter records. But the company 'refused to honor the defense subpoena.'").

[288] *See* Jackson, *supra* note 276, at 16.

[289] *See id.*

1. Algorithmic Technologies are not Presented at Trial

Algorithmic technologies primarily used as investigative leads are often not subject to trial. Trial rates are low: approximately 97% of federal defendants plead guilty, as do comparative percentages of state defendants.[290] Many of these defendants are innocent, but plea-bargaining leaves them in a powerless position: at the pre-trial stage, prosecutors know much more about the strength of their case, and have the ability to dictate the charges and the sentence offered.[291] Defendants take plea bargains to avoid the tax of going through with a trial. Sentencing after trial usually incurs a much higher sentence than that offered with a plea, and even factually innocent defendants feel ill-equipped to tackle the government, especially when unfamiliar technological evidence is involved.[292] For example, Nijeer Parks knew he did not commit the crime he was accused of, but he was aware that the government had facial recognition evidence saying he did.[293] Concerned that he would lose at trial when the facial recognition evidence was presented, he considered taking a seven-year plea deal to avoid getting a ten-plus-year sentence if he lost.[294]

In this plea-bargaining system, prosecutors are incentivized to present the outputs of algorithmic technologies as inculpatory fact to encourage defendants to take guilty pleas and save the government the cost of trial. With the proliferation of rapid DNA processing, prosecutors will be able to present incriminating DNA profile matches in more cases. As compared to cases in which traditionally processed DNA is used, rapid DNA will equip law enforcement to process samples in lower-profile cases, and processing speed will allow prosecutors to confront defendants with matches at earlier pre-trial stages.[295] As a result, prosecutors will encourage more defendants to take pleas based on DNA evidence.[296]

Defendants have an incentive to avoid trial when unfamiliar technology is at stake. In many cases, defendants will have no idea that evidence like a DNA match is susceptible to a reliability challenge. When plea deals are negotiated, a prosecutor might signal to a

---

[290] *See* Jed S. Rakoff, *Why Innocent People Plead Guilty*, N.Y. REV. OF BOOKS (Nov. 20, 2014), https://www.nybooks.com/articles/2014/11/20/why-innocent-people-plead-guilty/.

[291] *See id.*

[292] *See id.*

[293] *See* GARVIE, *supra* note 60, at 7, 49.

[294] *See id.*

[295] *See* Blume, *supra* note 219.

[296] *See id.*

defendant that incriminating evidence exists against her, but may obscure its unreliable origins, as when the matched profile comes from a rapidly processed crime scene sample.[297] Similarly, a prosecutor could hint that a defendant was identified, without specifying facial recognition technology made the identification. Even a prosecutor disclosing the technological source of "incriminating" evidence—ShotSpotter, facial recognition, rapid DNA—may mean little to a defendant. The prosecutor will pass off the technology as ironclad proof of guilt, and many defendants will not readily perceive its reliability issues. For many defendants, accepting a lesser sentence via plea bargain is the right choice, when the alternative is challenging law-enforcement's use of algorithmic technology with little information and few resources.

Unsuccessful suppression hearings may also force guilty pleas. Terrill Rickmon conditionally pleaded guilty upon the failure of his motion to suppress ShotSpotter evidence in district court.[298] The suppression hearing serves as a litmus test for the strength of a case at trial: if a defendant cannot exclude incriminatory evidence, she is far more likely to lose in a guilt phase, and forging ahead is not worth the cost of incurring a longer sentence. This means that courts' inadequate rulings on reliability under the "totality of the circumstances" are often the last word on the subject.

If a case involving unreliable algorithmic technology does make it to trial, prosecutorial discretion may work to immunize unreliable technologies from challenge. Prosecutors may not introduce investigative leads at trial when they can support a conviction with other evidence, and there is no trial mechanism to contest the reliability of evidence not introduced. In cases involving facial recognition, for example, prosecutors have not introduced evidence of facial recognition matches at trial.[299] Though facial recognition may lead police to a suspect and provide probable cause for arrest, by the time of trial, the identification of the defendant is likely supported by other means, like traditional witness identifications.[300]

Because prosecutors do not seek to introduce investigative leads at trial, the amount of discovery defense attorneys are entitled to regarding algorithmic technologies is unclear. At least one state court

---

[297] *See* MURPHY, INSIDE THE CELL, *supra* note 159, at 78 (giving example of a prosecutor letting slip that "We have DNA" without specifying that the DNA match in question is based on a low-copy number test and may be unreliable).

[298] *See Rickmon*, 952 F.3d at 880.

[299] *See* Jackson, *supra* note 276, at 20.

[300] *See id.*

has ruled that defendants are not entitled to the full inputs and outputs of facial recognition technology when the facial recognition match is not presented at trial, even if subsequent, introduced identifications were predicated on the match. In Florida, Willie Allen Lynch was convicted after facial recognition technology identified him from photographs snapped by undercover officers engaging in a drug transaction.[301] A crime analyst generated the match to Lynch, picking him out from a list of candidates, and the officers confirmed the man in the photograph sold them crack cocaine.[302] Lynch unsuccessfully sought to suppress the identification, and the case proceeded to trial, at which the officers testified, but the crime analyst did not.[303] Defense attorneys argued that Lynch was entitled to the other candidate matches the facial recognition program generated under *Brady v. Maryland*, in which the Supreme Court held that defendants have a due process right to exculpatory evidence.[304] A Florida appellate court disagreed.[305]

If other courts rule similarly to the *Lynch* court, a prosecutor's choice to insulate an investigative lead from trial may preclude defendants from important discovery about such leads. To obtain relief under *Brady*, a defendant "must convince [the court] that 'there is a reasonable probability' that the result of the trial would have been different if the suppressed documents had been disclosed to the defense."[306] But this "reasonable probability" is difficult to assess if defendants are unable to obtain complete information on algorithmic technologies in the first place. *Lynch* is a good example of this catch-22. The court ruled that because Lynch could not show that the other candidate photos resembled him, he could not argue that another match was credibly an alternate culprit, and therefore would have changed the jury's assessment.[307] The court missed the point: Lynch could not make that argument because he had never seen the photos. And though Lynch had not called the analyst at trial—another knock to his prejudice argument—the court ignored how her candidate choice biased the proceedings by influencing the officers' identifications. Only broad discovery at the outset of criminal proceedings can bypass the problems rife in the Lynch case, ensuring that defendants are adequately equipped to face reliability challenges; had prosecutors given Lynch

---

[301] *See Lynch v. State*, 260 So.3d 1166, 1168–69 (Fla. Dist. Ct. App. 2018).

[302] *See id.* at 1169.

[303] *See id.*

[304] *See id.* at 1169-70 (citing *Brady v. Maryland*, 373 U.S. 83 (1963)).

[305] *See id.* at 1170.

[306] *Strickler v. Greene*, 527 U.S. 263, 289 (1999).

[307] *See id.*

the candidate photos, he would have been primed to question the identifications of the officers and the analyst.

If they do present evidence of investigative leads at trial, prosecutors may present sanitized outputs that hide the unreliability of the original lead. For example, when ShotSpotter is used to support prosecutions, it is presented at trial by "expert engineers" who conduct back-end reviews of the sensors' audio to verify and/or correct the server's initial determinations.[308] Police officers work with company technicians to review audio from given timeframes and locations, resulting in the reclassification of undetected sounds as new gunshots post-hoc.[309] Post-processing the alert to add multiple gunshots can make it appear more reliable,[310] as can adding features that corroborate officer testimony.[311]

2. Programmatic Uses of Algorithmic Technologies are Not Subject to Judicial Scrutiny

Programmatic uses of algorithmic technologies are also obscured from judicial scrutiny. Most police-citizen encounters facilitated by ShotSpotter do not result in arrest or prosecution. This does not mean the technology is not harmful, but it does mean that its role in facilitating pretextual policing strategies like stop-and-frisks is often cloaked from courts' view.

---

[308] Greene *Godinez* testimony, *supra* note 73, at 382–84.

[309] *See id.* at 382–92; *see also* Todd Feathers, *Police Are Telling ShotSpotter to Alter Evidence from Gunshot-Detecting AI*, VICE (July 26, 2021), https://www.vice.com/en/article/qj8xbq/police-are-telling-shotspotter-to-alter-evidence-from-gunshot-detecting-ai. When Silvon Simmons was shot at four times (and hit three times) by a Rochester police officer, Joseph Ferrigno, ShotSpotter analysts reclassified the initial alert—which detected three shots—in post-processing to show that five shots were fired. The adjustment was made at the direction of the Rochester Police Department to support Ferrigno's testimony that Simmons had shot at him first. In fact, Simmons had not shot first, and he was acquitted of attempted murder of a police officer. *See* Harvey Gee, *"Bang!": ShotSpotter Gunshot Detection Technology, Predictive Policing, and Measuring Terry's Reach*, 55 U. MICH. J. L. REFORM 767, 781–83 (2022).

[310] A judge in Massachusetts stated that a series of ShotSpotter alerts is an "acoustic trail of breadcrumbs," with each lending weight to the reasonable suspicion calculus and rendering the others more reliable. The same principle applies to an alert that perceives multiple gunshots. *Commonwealth v. Ford*, 182 N.E.3d 1013, 1018 (Mass. App. Ct. 2022).

[311] *See supra* note 308 (describing the adjustment of the ShotSpotter alerts in *Simmons* to corroborate officer testimony).

In Chicago, ShotSpotter has become a pretext to stop-and-frisk residents in neighborhoods of color.[312] ShotSpotter alerts rarely result in the discovery of gun-related crime.[313] Thus, ShotSpotter alerts *in the present* rarely result in prosecution, but they do lead to extensive police deployments in the minority communities where sensors are placed.[314] However, Chicago police have further relied on *past* alerts to conduct stop-and-frisks in the neighborhoods where ShotSpotter sensors are concentrated, citing a high instance of past ShotSpotter alerts as contributing to reasonable suspicion that a person may be involved in gun crime.[315] This is a use that SoundThinking implicitly supports: it sells predictive technology that works with its gunshot detection sensors to analyze crime and deploy officers accordingly.[316] Since the Chicago sensor network is exclusively concentrated in majority-Black and Latinx neighborhoods,[317] this reliance on past alerts has the effect of, as one ACLU analyst put it: "distort[ing] gunfire statistics and creat[ing] a circular statistical justification for over-policing in communities of color."[318] Attorneys in New York have criticized the technology for perpetuating patterns of surveillance and criminalization against "New Yorkers of color who have already been heavily subject to discriminatory enforcement by the NYPD for decades through stop-and-frisk and other enforcement practices."[319] Thus, ShotSpotter's primary use has not been to detect violent crime,

---

[312] *See* Joh, *Unexpected Consequences*, supra note 47, at 523–29 (analyzing the OIG report and pulling out instances of ShotSpotter's "misuse" as a pretext for frisks).

[313] *See* OFF. OF THE INSPECTOR GEN., *supra* note 8, at 3.

[314] *See id.* at 2 (citing over 50,000 deployments in 17 months).

[315] *See* Joh, *Unexpected Consequences*, supra note 47, at 523–29 (analyzing the OIG report and pulling out instances of ShotSpotter's "misuse" as a pretext for frisks); *see also* OFF. OF THE INSPECTOR GEN., *supra* note 86, at 19.

[316] *See, e.g.*, *ResourceRouter*, SOUNDTHINKING (2023), https://www.soundthinking.com/law-enforcement/resource-deployment-resourcerouter/ (advertising automated patrol direction based on crime data).

[317] In other jurisdictions, like Kansas City, Missouri, and Atlanta, Georgia, ShotSpotter sensors are also concentrated in predominantly non-white communities. *See* Maneka Sinha, *The Dangers of Automated Gunshot Detection*, 5 J. L. & INNOVATION 63, 87.

[318] Jay Stanley, *Four Problems with the ShotSpotter Gunshot Detection System*, ACLU (Aug. 24, 2021), https://www.aclu.org/news/privacy-technology/four-problems-with-the-shotspotter-gunshot-detection-system.

[319] *Comments on NYPD ShotSpotter Impact and Use Policy*, CTR. FOR CONST. RTS. 1–2 (Feb 25, 2021), https://ccrjustice.org/sites/default/files/attach/2021/02/Shot%20Spotter%20Comments%20CCR%20BLH%202-25-21.pdf.

but to surveil in communities of color. The ShotSpotter evidence relevant to the criminal case of a single individual stopped based on an alert, therefore, would implicate only a small slice of the technology's impact.

The use of rapid DNA, too, evidences larger programmatic purposes. Because of the comparative ease and speed of processing DNA on rapid machines, relative to traditional processes, the adoption of these machines portends increased DNA collection across law enforcement agencies.[320] Rapid DNA, like ShotSpotter, is set to result in countless day-to-day incursions on individual liberty when people are asked, with little choice, to give law enforcement their DNA.[321] And if the police use rapid DNA like experts think they will—to build DNA databases from "stop and spit" encounters—most swabs processed on the new machines will not be subject to a reliability examination in court, because they will not lead to a match and criminal prosecution.[322]

Courts are ill-equipped to evaluate and regulate these programmatic uses of technology. When policing strategies rely on individualized suspicion, courts can assess whether an investigation of a defendant was justified.[323] But policing strategies that rely on mass stops are not amenable to court evaluation. Even if people feel they were wrongfully stopped by the police, they may not know what technology precipitated the interaction—insulating the algorithmic technologies discussed here from review. And well-informed plaintiffs, equipped with the knowledge to challenge this technology via civil

---

[320] *See supra* Section III.C (describing the function of rapid DNA machines).

[321] *See supra* Section III.D (discussing the DNA sampling, by "consent" of people stopped by police).

[322] *See* Erin Murphy, *DNA in the Criminal Justice System*, *supra* note 218, at 356, 358, 369 (predicting increased stop and spit sampling and the attendant growth of DNA databases, but a lag in the processing of samples connected to crime scenes); *see also* Lauren Kirchner, *DNA Dragnet: In some Cities, Police Go From Stop-and-Frisk to Stop-and-Spit*, PROPUBLICA (Sep. 12, 2016), https://www.propublica.org/article/dna-dragnet-in-some-cities-police-go-from-stop-and-frisk-to-stop-and-spit (discussing "stop and spit": the practice of collecting DNA in police stops from people not charged with nor suspected of crimes); Heather Murphy, *Coming Soon to a Police Station Near You*, *supra* note 214 (describing Bensalem, PA law enforcement's use of rapid DNA to process stop-and-spit samples).

[323] *See* Friedman & Ponomarenko, *supra* note 243, at 1872.

lawsuits, nonetheless face limitations on justiciability that mean most stops will not make it to court.[324]

3. Algorithmic Technologies are not Adequately Screened at Trial

When algorithmic technologies are presented at trial, the existing legal standards that govern their admission allow for unreliability. The algorithmic technologies discussed in this Article would be presented through expert testimony at trial. The admissibility of such testimony is governed by *Daubert v. Merrell Dow Pharmaceuticals, Inc*.[325] and *Frye v. United States*.[326] *Daubert* established the trial judge as the gatekeeper of expert testimony, tasked with ensuring that the testimony rests on a "reliable foundation."[327] The *Daubert* Court's chief concern was that scientific testimony be valid.[328] *Daubert* replaced *Frye v. United States* as the standard governing admission of expert testimony in the federal courts.[329] Some states, however, still use a modified *Frye* test to determine admissibility.[330] The *Frye* standard hinges admission on whether expert testimony is

---

[324] *See* Friedman & Ponomarenko, *supra* note 243, at 1874. While criminal defendants can challenge algorithmic technologies with suppression motions, people who are stopped or detained, but not prosecuted, can only challenge the stop through civil litigation. To bring a justiciable civil challenge, the plaintiff must have standing to sue under Article III of the Constitution, which requires that a plaintiff have a redressable injury. *See Clapper v. Amnesty Int'l*, 568 U.S. 398, 408. The Supreme Court has interpreted standing to bar meritorious civil rights claims. *See* Friedman & Ponomarenko, *supra* note 243, at 1868. For example, in *City of Los Angeles v. Lyons*, the Court barred a plaintiff who had been choked by the Los Angeles police from suing for injunctive relief, ruling that the plaintiff could not establish "a real and immediate threat that he would again be stopped . . . by an officer who would illegally choke him into unconsciousness." 461 U.S. 95, 95 (1983). The *Lyons* holding makes it difficult for plaintiffs to seek injunctive relief against widespread police practices that violate constitutional rights.

[325] *See generally Daubert v. Merrell Dow Pharmaceuticals, Inc*., 509 U.S. 579 (1993).

[326] *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923).

[327] *Daubert*, 509 U.S. at 597.

[328] *See id*. at 590 ("Proposed testimony must be supported by appropriate validation—*i.e.*, 'good grounds,' based on what is known.") (emphasis in original).

[329] *See id.* at 587.

[330] *See Principles of Forensic DNA for Officers of the Court: The Frye Test*, NAT'L INST. OF JUST. (June 20, 2023), https://nij.ojp.gov/nij-hosted-online-training-courses/principles-forensic-dna-officers-court/11-pretrial-dna-evidence-issues/expert-testimony-dna-cases/frye-test-cont.

based on a technique "'generally accepted' as reliable in the relevant scientific community."[331] These inquiries are premised on the idea that if an expert is given "wide latitude to offer opinions," meant to weigh on an inquiry as serious as guilt or innocence, those opinions should be based in something trustworthy.[332] Relatedly, courts have distinguished technology suitable for use as "investigative leads" but insufficiently reliable for admission at trial.[333] One New York court named facial recognition software as such a technology.[334]

The *Daubert* Court suggested several factors that courts might consider in determining whether a relevant scientific methodology is valid: (1) whether the theory or technique in question can be and has been tested; (2) whether it has been subjected to peer review and publication; (3) whether a particular technique has a high known or potential error rate; (4) whether there are standards controlling its operation; and (5) whether it enjoys general acceptance within a relevant scientific community.[335] In practice, many courts only loosely rely on these *Daubert* factors, resulting in lackluster reliability assessments. Professor Brandon Garrett and attorney Chris Fabricant observe that many judges will admit expert testimony based on the expert themself, relying on their credentials or prior admissions, rather than conducting an assessment parsing the *Daubert* factors relevant to the expert's methodology.[336]

Facial recognition[337] and rapid DNA[338] are yet to be assessed under *Frye* and *Daubert*, but ShotSpotter has undergone examinations

---

[331] *Daubert*, 509 U.S. at 585 (citing *Frye*, 293 F. at 1014).

[332] *See id.* at 591–92.

[333] *See People v. Collins*, 15 N.Y.S.3d 564, 575 (N.Y. Sup. Ct. 2015) ("[T]he results of some other techniques . . . can aid an investigation, but are not considered sufficiently reliable to be admissible at a trial.").

[334] *See id.* at 576 ("The products of polygraph technology and of facial recognition technology similarly can sometimes have value, but evidence produced by those technologies is not generally accepted as reliable by the relevant scientific communities and so cannot be admitted in trials.").

[335] *See Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 149–50 (1999) (summarizing the *Daubert* factors).

[336] *See* Brandon L. Garrett & M. Chris Fabricant, *The Myth of the Reliability Test*, 86 FORDHAM L. REV. 1559, 1568–69, 1579 (2018).

[337] *See* GARVIE, *supra* note 60, at 44–45.

[338] *See* Andrea Roth, *Admissibility of DNA Evidence in Court*, *in* SILENT WITNESS: FORENSIC DNA EVIDENCE IN CRIMINAL INVESTIGATIONS AND HUMANITARIAN DISASTERS 306–07 (Henry Erlich ed., 2020) (discussing the admissibility of rapid DNA as an issue for courts to consider in the future).

in state and federal courts.[339] Since ShotSpotter has never been tested for false positive errors or human analysts' proficiency, courts that admit the technology do so without knowledge of its true error rate.[340] Still, judges have relied on other courts' determinations of admissibility to deem ShotSpotter evidence sufficiently reliable for trial without conducting their own assessments.[341] ShotSpotter, facial recognition, and rapid DNA testing on crime scene samples should fail the *Daubert* test because the methods are not scientifically valid,[342] and the testimony relying on the technologies should fail *Frye* as well, as experts have not reached a consensus on any of the methods' acceptability and reliability.[343] Certifying these technologies for trial before they are properly validated runs the risk that courts will admit them again and again, convicting people on inaccurate evidence. Knowledge of algorithmic technologies is changing fast, and courts should wait to amass complete information on validity before they act.[344] In the next Part, I discuss regulating law enforcement agencies'

---

[339] *See, e.g.*, *State v. Hill*, 851 N.W.2d 670, 679–81 (Neb. 2014); *United States v. Godinez*, No. 18 CR 278, 2019 WL 4857745 at *3 (N.D. Ill. Oct. 2, 2019).

[340] *See, e.g.*, *Hill*, 851 N.W.2d at 680–81 (affirming a lower court's admission of ShotSpotter evidence under *Daubert* when the only "testing" considered was a live-fire test upon sensor activation).

[341] *See, e.g.*, *United States v. Godinez*, 7 F.4th 628, 635 (7th Cir. 2021) (citing *United States v. Godinez*, 2019 WL 4857745). The District Court in *Godinez* justified a ShotSpotter employee's expertise and the technology's sufficient reliability on the prior admission in *Hill*. *See id.* (citing *Hill*, 851 N.W.2d). The District Court's admission was later declared harmless error by the Seventh Circuit. *See id.*

[342] *See supra* Sections III.A.2, III.B.2, III.C.2 (discussing the unreliability of ShotSpotter, facial recognition, and rapid DNA, respectively).

[343] *See, e.g.*, Tr. of Preliminary Hearing, *California v. Gillard*, No. 05-164044-0 at 4065–67 (Cal. Super. Ct. 2014) (identifying ShotSpotter as a new technique not generally accepted by the scientific community); GARVIE, *supra* note 60, at 45–46 ("[T]here is near universal agreement that a face recognition search does not create . . . evidence to be introduced in court."); *Rapid DNA*, FED. BUREAU OF INVESTIGATION, https://le.fbi.gov/science-and-lab/biometrics-and-fingerprints/codis/rapid-dna. (describing that advancements are needed if law enforcement is to reliably use rapid DNA on crime scene samples).

[344] *See generally* Simon A. Cole, *Changed Science Statutes: Can Courts Accommodate Accelerating Forensic Scientific and Technological Change?*, 57 JURIMETRICS J. 443, 445 (2017) (describing "science lag" in legal verdicts—the difficulty judges have with assessing when science has changed).

validation and adoption of algorithmic technologies, as well as strengthening pre-trial reliability inquiries.

V. SOLUTIONS

A. *Court Regulation*

The "fundamental command of the Fourth Amendment is that searches and seizures be reasonable."[345] The crucial inquiry as to whether a police intrusion is constitutional is whether the intrusion is reasonable. The Supreme Court has developed multiple, sometimes incoherent, models for assessing reasonableness, including balancing tests that weigh governmental interests against individual liberty.[346]

When considering algorithmic technologies, the underlying question should be whether a search or seizure stemming from the technology is reasonable, not just whether the intrusion comports with traditional probable cause and reasonable suspicion justifications. As I discussed in Sections II.B and IV.B, the usual totality of the circumstances analysis, applied without modification, does not allow courts to properly assess algorithmic technologies and how they affect investigations.

The Supreme Court previously considered how the totality of the circumstances inquiry applied to drug-sniffing dogs in *Florida v. Harris*, ruling that dog alerts to drugs presumptively supply probable cause to search when a training organization has certified a dog after testing its reliability.[347] While drug dogs are, in many ways, an analog to the algorithmic technologies discussed in this note, the reliability inquiry developed in *Harris* should further adapt to more rigorously consider the reliability of algorithmic technologies. A more stringent suspicion inquiry would account for gaps in the courts' ability to analyze newer technologies. The Supreme Court developed the doctrines that analyze police officers' determinations of probable cause and reasonable suspicion when police suspicion was primarily based on analogous observations or human sources.[348] Now, technology has greatly enhanced the information available to police and affected how they view observed behaviors: recall that Chicago police officers will consider a high amount of ShotSpotter alerts in the vicinity as a

---

[345] *New Jersey v. T.L.O.*, 469 U.S. 325, 340 (1985).

[346] *See* CLANCY, *supra* note 33, at 679–80 (describing five predominant models for assessing reasonableness).

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

[348] *See* FERGUSON, *supra* note 48, at 55–56.

contributor to reasonable suspicion.[349] Reasonable suspicion and probable cause are much easier to come by, making the police more powerful. The Fourth Amendment interest in curbing police power suggests that an inquiry should further interrogate the tools that enhance it.

Herring v. United States further suggests that police reliance on unreliable technology is unreasonable.[350] If reliance on unreliable technology is unreasonable, it follows that evidence that flows from its use should be excluded from court.

1. Require a Stringent Reliability Standard for Suspicion Inquiries

Courts should be required to more rigorously assess the reliability of algorithmic technologies that contribute to probable cause or reasonable suspicion. Not only does this correspond to the Fourth Amendment interest in regulating police power, but it also conforms to reliabilism. Credited investigative leads, whether generated by algorithmic technologies or not, set in motion a process that results in a search or seizure. Using reliabilist thinking, that outcome is not justified unless the process is reliable, so the lead must be reliable, and generated by reliable means. Declining to meaningfully evaluate the role of algorithmic technology in generating an investigative lead is ignorant of the way its output may infect an investigation through officer biases.

The best way for courts to assess the reliability of algorithmic technologies that contribute to probable cause and reasonable suspicion is by using a modified "totality of the circumstances" analysis. *Florida v. Harris* can serve as a model.

Both scholars[351] and attorneys[352] have suggested that the algorithmic technologies used by police function similarly to drug-sniffing dogs. Like drug-sniffing dogs, algorithmic technologies take in inputs (sounds, probe photos, biological material), and spit out an output that alerts to criminality. And like drug-sniffing dogs, the inner

---

[349] *See id.* at 56 (discussing how big data can distort reasonable suspicion by describing a hypothetical case of a young man walking home, subjected to a stop because, among other information the police knew about him, he was in a high crime area).

[350] *See Herring v. United States*, 555 U.S. 135, 146 (2009).

[351] *See* Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 PA. L. REV. 871, 911–912 (2016); *see also* Benedict, *supra* note 224, at 872–75.

[352] *See* SANTAMARIA, *supra* note 278, at 19; *see also* Brief for Roderick & Solange MacArthur Justice Center, *Ford*, No. 2020-P-1334 at 31–34.

workings of algorithmic technologies may be unknown "black boxes" to the police officers who use them.

Scholars,[353] attorneys,[354] and courts[355] have also analogized algorithmic technologies to anonymous tips, but the analogy is less apt. Like anonymous tipsters, algorithmic technologies provide information from a base of knowledge that may be unknown. However, the jurisprudential response to anonymous tipsters requires corroboration of at least some details from analogous police observation. Corroborating innocent facts from a tip, the Supreme Court has decided, makes it more likely that an informant has "inside information" and that their suggestion of criminality is true.[356] This is not so with algorithmic technologies. In the case of technologies like facial recognition, for which the police control an input, corroborating the accuracy of the fed data—i.e., that the shoplifter in the probe photo is the person the facial recognition software identified—says nothing about the technology's own predictive properties.[357] The technology was already fed suspicion. Moreover, the court has placed special importance on corroboration of anonymous informants' predictions of defendants' "future behavior."[358] The technologies discussed in this Article make no such predictions. Corroboration of a suspect match or a fired gun shot is a different process from corroborating a human informant, because the technologies do not provide hints to extrinsic, verifiable detail. The best opportunity to evaluate an algorithmic technology's reliability is not in corroborating its tip or output, but in assessing the underlying computer program and how it is used.

The decision in *Florida v. Harris* followed this premise by hinging the question of reliability on a measure of the police tool's accuracy, the drug dog, itself. *Harris* adapted the "totality of the circumstances" inquiry to drug-sniffing dogs, establishing a presumption of probable cause upon a dog's alert if "a bona fide organization has certified a dog after testing his reliability in a controlled setting," or "in the absence of formal certification, if the dog has recently and successfully completed a training program that evaluated his proficiency in locating drugs."[359] Absent any challenges, these factors sufficiently establish a dog's reliability, but if a defendant

---

[353] *See* Rich, *supra* note 336, at 908–11.

[354] *See supra* note 352.

[355] *See Rickmon*, 952 F.3d at 882 ("[W]e conclude [ShotSpotter] is analogous to an anonymous tipster.").

[356] *See White*, 496 U.S. at 332.

[357] *See* Rich, *supra* note 351, at 911.

[358] *White*, 496 U.S. at 331–32.

[359] *Harris*, 568 U.S. at 242–43, 246–47.

challenges the adequacy of a dog's training, "evidence of the dog's (or handler's) history in the field . . . may sometimes be relevant."[360]

In creating the presumption in *Harris*, the Court reversed the Florida Supreme Court's requirement that the government needed to produce an "evidentiary checklist" demonstrating a dog's reliability, including the dog's "prior 'hits' and 'misses' in the field."[361] A "checklist" was contrary to the "totality of the circumstances" inquiry, and the Court felt "field performance records" had little import in an ordinary case.[362]

However, critiques of *Harris* argue that the Court's focus on "training and certification" as strong evidence of a drug dog's reliability is flawed, as is its conclusion that field performance is weak evidence of reliability.[363] It is unclear how well training approximates a dog's accuracy, as there are no accepted standards for dog training, and dogs may be certified at varying levels of performance.[364] More importantly, training programs in controlled environments fail to account for the real-world circumstances that may cause a dog to falsely alert.[365] For example, a handler biased to expect evidence of drugs could unconsciously cue a dog.[366] Field performance records would better approximate a dog's reliability in detecting drugs. The Court's argument against considering field error rates was that they would underestimate reliability: cases where a dog alerted to a residual odor would be counted as false positives because no physical drugs were found, when in reality, the dog was a competent detector.[367] This argument misunderstands probable cause, because an alert on a residual odor when no contraband is present is a false positive for the purposes of the Fourth Amendment.[368] The *Harris* court eschewed analysis of false positive errors, when false positive errors strike at the core of the Fourth Amendment: they indicate unreasonable searches and seizures.

These critiques of *Harris* analogize to concerns about algorithmic technologies and suggest that courts should consider field performance in assessing algorithmic technologies' reliability. As with drug-sniffing dogs, controlled assessments of algorithmic

---

[360] *Id.* at 247.
[361] *Id.* at 244–45, 250.
[362] *See id.* at 245.
[363] *See* Rich, *supra* note 351, at 917.
[364] *See id.*
[365] *See id.*
[366] *See id.*
[367] *See Harris*, 568 U.S. at 246.
[368] *See* Rich, *supra* note 351, at 917–18.

technologies, at least to date, are poor measurements of reliability. Some technologies, like ShotSpotter, lack independent validation.[369] Others, like facial recognition, are evaluated under circumstances that fail to approximate their real accuracy.[370] And just as the reliability of drug-sniffing dogs may modulate with different handlers and circumstances, the reliability of algorithmic technologies is also affected by use in different environments and with different human reviewers. Further, even assuming *arguendo* that the Court was correct to worry that field performance statistics may overstate false positives,[371] falsely undermining the reliability of good technology, that worry is not present with many algorithmic technologies. False positive rapid DNA matches or facial recognition hits are unlike a "residual odor": matches are either correct or not. Though it is possible to overmeasure false positives for technologies like ShotSpotter, a comprehensive reliability inquiry should also account for the methodology used to tally error data.

Field performance data may be hard to measure or unavailable. Courts should give weight to the closest approximations of field performance of a particular technology, placing more weight on data specific to the jurisdiction at issue. One example of field data would be the Chicago OIG report on ShotSpotter, which approximated how often the sensors installed in the city accurately alerted gun crime through a rigorous analysis of CPD data.[372] Hypothetical field performance data for facial recognition would track when a match was deemed successful because it was eventually corroborated by extrinsic evidence, when a match was a false positive because further investigation deemed it erroneous, or when the system missed a match because someone eventually convicted was in the searched database. Data on rapid DNA would follow a similar structure and would ideally track matches from single machines or agencies. Since agencies fail to track the use and results of technology like ShotSpotter and facial recognition, this data may be unobtainable for the specific jurisdiction, or otherwise impracticable to estimate. Prosecutors should nonetheless produce such data in discovery when it is available.

Courts should also give weight to other case-specific factors that impact a technology's reliability, for example: the exact inputs used and their condition (for images and DNA); a reviewing officer's history and training; the environment in which a system is situated (for

[369] *See supra* Section III.A.2.
[370] *See supra* Section III.B.2.
[371] *See Harris*, 568 U.S. at 246.
[372] *See* OFF. OF THE INSPECTOR GEN., *supra* note 8, at 2–3.

ShotSpotter, and potentially for DNA machines); and any demographic variation in error rates (for facial recognition).[373]

Review of validity testing, the core mechanism for assessing evidentiary reliability as established by PCAST[374] and *Daubert*,[375] should be a requirement in any pre-trial reliability evaluation. The *Harris* decision also places a high value on a policing technology's performance on validity testing, referred to as testing in a "controlled setting."[376] The testing suggested in *Harris* considers both false negatives and false positives—assessing a dog on whether it alerts "where drugs are hidden and where they are not."[377] The Court felt that such testing approximated "trust" in the dog's alert, and therefore rendered reliance on it reasonable to support probable cause.[378] Validity, as the Court opined in *Daubert*, can approximate reliability.[379] But as criticisms of *Harris* reveal, it is an incomplete measure, and courts should require more if the expectation is that the police rely on accurate technologies to furnish individualized suspicion, as the *Harris* Court assumed.[380]

Thus, the factors courts should consider in a complete pre-trial reliability inquiry include: (1) whether a technology used as a lead has been independently tested, in circumstances that mirror how it was used in the case; (2) the error rates (both false positives and false negatives) assessed by such testing; (3) data on, or approximations of, the technology's error rates as deployed in the field; and (4) other factors bearing on the technology's reliability as used in the field: for example, the inputs used in the case, a reviewing officer's history and

---

[373] *See Harris*, 586 U.S. at 247 ("[E]ven assuming a dog is generally reliable, circumstances surrounding a particular alert may undermine the case for probable cause—if, say, the officer cued the dog (consciously or not), or if the team was working under unfamiliar conditions.").

[374] *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *supra* note 62, at 4–5 ("Foundational validity, then, means that a method can, *in principle*, be reliable . . . . Foundational validity requires that a method has been subjected to empirical testing by multiple groups, under conditions appropriate to its intended use.").

[375] *Daubert*, 509 U.S. at 590 ("Proposed testimony must be supported by appropriate validation—*i.e.*, 'good grounds.'") (emphasis in original).

[376] *See Harris*, 586 U.S. at 246–47.

[377] *Id.* at 246.

[378] *See id.* at 246–47.

[379] *See Daubert*, 509 U.S. at 590, n.9.

[380] *See Harris*, 586 U.S. 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.").

training, circumstances of the environment in which a system is situated, and any demographic variation in error rates.

The *Harris* Court also stressed that a "defendant . . . must have an opportunity to challenge such evidence of a dog's reliability."[381] This suggests that part of a reliability standard should be a requirement of information access. To facilitate reliability assessments and cross-examination, prosecutors should be legally required to disclose to defendants when a technology has been used as an investigative lead, regardless of admissibility at trial. Defendants should receive all attendant data with this disclosure. For example, if facial recognition was used in her case, a defendant should receive all relevant probe photos and candidate matches, not just her matched photo.[382] This would combat information deficiency and transparency problems, ensuring that defendants are well-informed enough informed to contest any technology that may incriminate them. It will also help ensure that the court has all the information to properly assess a technology's reliability, by giving both the prosecution and the defense the tools to present information about algorithmic technologies. A proposed federal law, the Justice in Forensic Algorithms Act, would further facilitate this exchange by repealing the use of any trade secret evidentiary privilege to withhold relevant evidence, allowing defendants access to underlying data that could help them challenge algorithmic technologies.[383]

2. Exclude Evidence Stemming from Unreliable Technologies

Dicta in the Supreme Court case *Herring v. United States* suggest that when a court concludes that a technology is unreliable, a defendant should succeed in excluding the resulting evidence.[384]

In *Herring v. United States*, the Supreme Court considered a case in which police arrested Bennie Dean Herring pursuant to a "negligent" error in a local warrant database: the system logged that there was an outstanding warrant out for Herring's arrest, but the warrant had been recalled.[385] The Court upheld evidence seized from Herring's car in a search incident to arrest.[386] The Court accepted,

---

[381] *Id.*

[382] *See supra* Section IV.C.1 for discussion on the importance of disclosing all photos relevant to a facial recognition match.

[383] *See* Justice in Forensic Algorithms Act, H.R. 2438, 117th Cong. § 2(b) (2021).

[384] *See Herring*, 555 U.S. at 145–47.

[385] *See id.* at 137–38.

[386] *See id.* at 147.

arguendo, that the search was a Fourth Amendment violation, but refused to exclude the evidence on the theory that the database error was a negligent, one-off mistake, and exclusion would therefore have no bearing on deterring future police conduct.[387] However, *Herring* suggested that it would be unreasonable under the Fourth Amendment for the police to rely on a system prone to recurring error, and that in such a case, exclusion could meaningfully deter reckless or deliberate conduct.[388] Justice Roberts, in the majority opinion, wrote that "[i]f the police have been shown to be reckless in maintaining a warrant system, or to have knowingly made false entries to lay the groundwork for future false arrests, exclusion would certainly be justified."[389] Justice Roberts implied that police reliance on systems where "systemic errors are demonstrated," "routine," or "widespread" could constitute reckless misconduct justifying application of the exclusionary rule.[390]

Taking to its logical extent, Roberts's opinion in *Herring* suggests that knowing, intentional, or reckless reliance on unreliable technology would be a Fourth Amendment violation. A defendant would have to show that such reliance rises above "nonrecurring and attenuated negligence."[391] Therefore, for the exclusionary rule to apply, it should be sufficient to show that a law enforcement agency was aware, or had reason to be aware, that an algorithmic technology was beset by a high rate of error, and chose to overlook it. *Herring* also suggests that it is unreasonable for police to rely on a system that routinely leads to false arrests;[392] therefore, a showing that an algorithm has led to false arrests should also support the application of the exclusionary rule. Police reliance on facial recognition and ShotSpotter has led to high-profile false arrests, so arrests made and evidence gleaned because of these technologies should be excluded under the Fourth Amendment.

---

[387] *See id.* at 145 ("To trigger the exclusionary rule, police conduct must be sufficiently deliberate that exclusion can meaningfully deter it, and sufficiently culpable that such deterrence is worth the price paid by the justice system. As laid out in our cases, the exclusionary rule serves to deter deliberate, reckless, or grossly negligent conduct, or in some circumstances recurring or systemic negligence. The error in this case does not rise to that level.").

[388] *See id.* at 146.

[389] *Id*.

[390] *See Herring*, 555 U.S. at 146.

[391] *Id.* at 144.

[392] *See id.* at 146 (quoting *Arizona v. Evans*, 514 U.S. 1, 17 (1995) (O'Connor, J., concurring) ("Surely it would *not* be reasonable for the police to rely . . . on a recordkeeping system . . . that *routinely* leads to false arrests.") (second emphasis added)).

In dissent in *Herring*, Justice Ginsburg argued that even so-called negligent errors—errors that appear to be simple mistakes and isolated occurrences—should give rise to application of the exclusionary rule.[393] Even in cases of careless error, Justice Ginsburg believed application of the rule would serve an important deterrent effect, incentivizing police to audit and properly maintain their records.[394] Justice Ginsburg felt this was of paramount importance when "law enforcement databases are insufficiently monitored and often out of date"[395]—a statement that remains globally true of police technology today. Law enforcement agencies have failed to validate algorithmic technologies and audit their use for error. Even if this oversight is merely negligent, application of the exclusionary rule when technologies operate unreliably can push agencies to improve the integrity of their technologies while removing defendants' burden to amass evidence of deliberate violations.[396]

Beyond deterrence, Justice Ginsburg felt excluding evidence obtained through error vindicated other important objectives of the exclusionary rule, like bolstering public trust by delegitimizing the conduct that produced the evidence, "preser[ving] the judicial process from contamination."[397] The Justice's repeated references to "contamination" speak to the core premise of reliabilism—that evidence is tainted because it is the fruit of an illegitimate process.

B. *Extrajudicial Regulation*

Even under the strictures of a new test and a stronger exclusionary rule, when the question of an algorithmic technology's reliability does reach the courts via a Fourth Amendment challenge to a search or seizure, courts may be ill-equipped to opine on what reliability truly requires.[398] Regardless, a court's declaration in a Fourth Amendment case as to what constitutes a technology as "reliable" for

---

[393] *See id.* at 157 (Ginsburg, J., dissenting) ("Negligent recordkeeping errors by law enforcement threaten individual liberty, are susceptible to deterrence by the exclusionary rule, and cannot be remedied effectively through other means.").

[394] *See id.* at 151–56 (Ginsburg, J., dissenting).

[395] *Id.* at 155 (Ginsburg, J., dissenting).

[396] *See Herring*, 555 U.S. at 157 (Ginsburg, J., dissenting).

[397] *Id.* at 152–53 (Ginsburg, J., dissenting) (quoting *Olmstead*, 277 U.S. 484 (Brandeis, J., dissenting)).

[398] *See generally* Brandon L. Garrett, Brett O. Gardner, Evan Murphy & Patrick Grimes, *Judges and Forensic Science Education: A National Survey*, 321 FORENSIC SCI. INT'L 1 (2021) (finding that many state judges lack understanding of forensic disciplines and were untrained in statistical methods and calculation of error rates).

the purposes of individualized suspicion is only the minimum threshold a technology must meet for constitutional compliance.[399] The federal government, state and local governments, and law enforcement agencies may demand algorithmic technologies meet a higher standard to ensure greater accuracy in policing and to strengthen public trust.

Lawmakers and law enforcement agencies should seek to further regulate algorithmic police technologies for reasons wholly beyond courts' reach. Courts regulate police power through addressing violations of constitutional rights, as when a litigant claims a search or seizure is illegal under the Fourth Amendment.[400] This judicial paradigm leaves much police harm unregulated by the courts.[401] Facial recognition technology,[402] ShotSpotter,[403] and DNA stockpiling[404] have all increased surveillance within communities of color, fostering distrust of police, with little attendant benefit to public safety. These intrusions are beyond the scope of the Fourth Amendment until they lead to a search or seizure. But by instituting laws and regulations to ensure that the technologies law enforcement agencies use are reliable, governmental actors can assuage public distrust and promote accurate and effective policing.

From the public's perspective, the use of unreliable technologies undermines police legitimacy. Members of the public do

---

[399] *See* Rachel A. Harmon, *The Problem of Policing*, 110 MICH. L. REV. 761, 777 (2012) (explaining that constitutional rights establish only minimum standards for law enforcement, a ceiling on government action that is more generous to law enforcement than the actual interests at stake would suggest).

[400] *See id.* at 762–63 (describing the Fourth Amendment as the conventional paradigm scholars consider for regulating police conduct).

[401] *See id.* at 763.

[402] *See Ban the Scan New York City*, AMNESTY INT'L (2022), https://banthescan.amnesty.org/nyc/ (campaigning against the New York Police Department's use of facial recognition technology to surveil communities of color; citing FRT as a tool that "can exacerbate discriminatory policing and prevent the free and safe exercise of peaceful assembly, by acting as a tool of mass surveillance").

[403] *See* MacArthur Justice Center, *The Burden on Communities of Color*, *supra* note 2 (campaigning against the Chicago Police Department's use of ShotSpotter; citing ShotSpotter as a tool that burdens Black and Latinx neighborhoods by increasing surveillance and diverting resources through unnecessary police deployments).

[404] *See* Erin Murphy & Jun H. Tong, *The Racial Composition of Forensic DNA Databases*, 108 CALIF. L. REV. 1847, 1897–99 (Jan. 16, 2020) (establishing that states have disproportionately stockpiled DNA from Black people, exposing a greater share of the Black American population to suspicion in DNA searches).

not rely on legal determinations or constitutional standards to determine whether police actions are appropriate: they base their judgments on perceptions of procedural justice.[405] A person's perceptions of whether they have been afforded procedural justice turns, in part, on their own participation in officer interactions—whether they are afforded the ability to explain themselves—as well as the fairness of officer decision-making.[406] Whether a decision is fair depends on the decision-maker's neutrality and objectivity, as well as consistency and transparency.[407] Untransparent use of unreliable technologies therefore diminishes perceptions of fairness. Police reliance on algorithmic outputs tends to foreclose opportunities for people to explain why the computer may be wrong.[408] A lack of transparency as to the use or existence of algorithmic technologies means that many people may not know why they were stopped. These factors indicate that the public will view police stops undertaken in reliance on algorithmic technologies as illegitimate.

This is already coming to fruition in Chicago. Activists cite ShotSpotter's lack of reliability as a reason to distrust the police.[409] Community organizer Adwoa Adyepong commented that "[the police] claim with ShotSpotter and other technology, we are being kept safe. This is obviously a lie."[410] The lack of transparency shrouding the city's ShotSpotter contract has also undermined public trust. When former Mayor Lightfoot extended the ShotSpotter contract without public notice, activists called the quiet extensions emblematic of the

---

[405] *See* Tracey L. Meares, Tom R. Tyler & Jacob Gardener, *Lawful or Fair? How Cops and Laypeople Perceive Good Policing*, 105 J. CRIM. L. & CRIMINOLOGY 297, 300, 309–11, 323, 327–28, 333 (2015).

[406] *See id.* at 308.

[407] *See id.* at 308.

[408] *See, e.g.*, First Amended Complaint, *Williams v. City of Chicago*, No. 1:22-cv-03773 at 69-71. When CPD officers detained Daniel Ortiz while pursuing a ShotSpotter alert, Ortiz, his friend, and several nearby community members attempted to tell officers that no gunshots were fired in the area. *See id.* at 70. The officers did not listen, cuffing Ortiz and repeatedly questioning him about the shots, relying on the pinpointed alert on their ShotSpotter apps over the small crowd of citizens that told them they were wrong. *See id.* at 70–71.

[409] *See* 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 (referencing Adam Toledo's death, one community organizer remarked "[u]sing untested, unverified technology to send police to our communities—that's horrific.").

[410] Masterson, *supra* note 108.

idea that "[c]ommunities have no input on what public safety is for us."[411]

When the public views police authority as legitimate, they are more likely to voluntarily obey the law and cooperate with the police.[412] Poor perceptions of procedural fairness contribute to a sense of illegitimacy, making it less likely that people will defer to police authority, rendering the police less effective.[413] Because the public view of police conduct is a primary driver of police-community relations, it is crucial that police maintain the popular view that they are acting appropriately.[414] This means that if police are to rely on algorithmic technologies, agencies must strengthen both the technologies' actual reliability, as well as the *perception* that they are reliable. The federal government can ensure that all algorithmic technologies marketed to police are foundationally valid through independent testing, though law enforcement agencies are in the best position to ensure the technologies are valid as applied. Municipal lawmakers can further strengthen the legitimacy of law enforcement agencies by publicly and democratically approving reliable technologies for use after a vetting process, giving the public occasion to participate in the decision.

1. Test Algorithmic Technologies for Validity

The first step in ensuring the reliability of new and existing algorithmic technologies is for the federal government to institute more robust independent testing of all products marketed to law enforcement. Validation testing will help to ensure that the technologies marketed to law enforcement are reliable, and therefore effective, in doing what they are supposed to do: aid police in responding to crimes and catching the correct perpetrators. An attendant requirement that the government and law enforcement agencies disseminate the results of these tests will enhance perceptions of fairness through transparency. Any technology available to a law

---

411 Schuba, *supra* note 10. The extensions were so secretive that Lightfoot's successor, Brandon Johnson, who campaigned on a promise to end the ShotSpotter contract, was not aware of them until days after his election. *See id.*

412 *See* Meares et al., *supra* note 405, at 309.

413 *See id.* at 308.

414 *See id.* at 304 ("[I]t is consequently important for police to focus upon two benchmarks of performance: (1) behaving in ways that are consistent with the law, and (2) acting so as to create and maintain the popular view that they are legitimate.").

enforcement agency should be properly validated in an independent study that approximates its use in law enforcement.

To ensure that more algorithmic technologies marketed to law enforcement agencies meet PCAST foundational validity requirements, the NIST should expand testing of police technologies. The NIST already tests subsets of police technologies, including vendor-submitted facial recognition programs, as discussed in Section III.B.2. The proposed Justice in Forensic Algorithms Act provides for the NIST to establish a "Computational Forensic Algorithm Testing Program" to validate any software "used to process, analyze, or interpret evidence," and if passed, would require that all algorithmic technologies be validated by the NIST in order to be admissible at trial.[415] Passage of the Act would help to ensure that all algorithmic technologies law enforcement agencies deploy are foundationally valid, as the Act also provides that all NIST testing shall follow "Computational Forensic Algorithm Testing Standards" that mirror the PCAST requirements for foundational validity.[416]

Along with expanding the technologies that NIST tests, current NIST testing needs to be improved to better mirror law enforcement use of algorithmic technologies and ensure validity.[417] Improved testing would (1) ensure the technology tested is the technology currently in use by law enforcement; (2) use realistic inputs, such as low-quality probe photos or DNA; (3) test use of the technology with "humans in the loop," including law enforcement officers without prior advanced training; and (4) simulate field usage errors, like altered probe photos or improperly stored DNA.[418] The Justice in Forensic Algorithms Act also provides for this improved testing, requiring that the Computational Forensic Algorithm Testing Program use "realistic sample testing data similar to what would be used by law enforcement

---

[415] Justice in Forensic Algorithms Act, H.R. 2438, 117th Cong. §§ 2(d), (g) (2021).

[416] *See id.* at § 2(a) ("Testing Standards shall address (A) the underlying scientific principles and methods implemented in computational forensic software; and (B) requirements for testing the software including the conditions under which it needs to be tested, types of testing data to be used, testing environments, testing methodologies, and system performance statistics required to be reported including—(i) accuracy, including false positive and false negative error rates; (ii) precision; (iii) reproducibility; (iv) robustness; (v) sensitivity; and (vi) system failure rates."); *see also* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *supra* note 62, at 47–54 (describing requirements for testing foundational validity).

[417] *See* Friedman, *supra* note 146, at 8–9 (describing the failure of NIST testing to properly approximate law enforcement use).

[418] *See id.*

in criminal investigations in performing such testing, including incomplete and contaminated samples."[419]

NIST testing would measure foundational validity. Local law enforcement agencies would need to assume responsibility for testing the validity of their systems as applied.[420] This would require that law enforcement agencies run validity testing known to be affected by built environments. For example, ShotSpotter sensors, as installed in a city's chosen locations, could be tested by examiners shooting gunshots from various locations, by examiners deploying gunshot-like sounds, or by examiners meticulously documenting the normal sounds in a specific coverage area and comparing them to alerts over a set period.

Analyst proficiency is key to measuring validity as applied.[421] Each human involved in interpreting the output of an algorithmic technology needs to be tested, ideally in a blind examination, for proficiency.[422] ShotSpotter analysts should be tested on soundbites for which the ground truth (gunshot or not gunshot) is known and assessed on how often they get it right. Similarly, law enforcement personnel who decide on facial recognition candidate matches should be tested for their accuracy in recognizing faces.[423] Rapid DNA proficiency may be harder to test, as no human is involved in the feature comparison element of matching two profiles. The primary issue with law enforcement users is the quality of the swabs they take and adherence to evidence protocols when processing and preserving DNA.[424] But individual machines should be validated for concordance to traditionally processed profiles and assessed for breaks that can cause contamination.[425] These "as applied" error rates affect reliability and should be provided to defendants when a specific analyst or machine is involved in their case.

The results of all validity testing and any underlying data should be made public by the NIST or other testing agencies, as well

---

[419] Justice in Forensic Algorithms Act, H.R. 2438, 117th Cong. § 2(d)(2).

[420] *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. AND TECH., *supra* note 62, at 56–59 (describing validity as applied).

[421] *See id.* at 57–58 (discussing the need for proficiency testing).

[422] *See id.* at 58 (advancing a preference for 'test-blind' proficiency examinations).

[423] *See* GARVIE, *supra* note 60, at 15 (discussing hypothetical proficiency examinations for "humans in the loop" in facial recognition).

[424] *See* Stout, *supra* note 181.

[425] *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 (identifying concerns about accuracy and contamination in rapid DNA processing).

as law enforcement agencies.[426] Law enforcement agencies should keep data on the use of algorithmic technologies in stops, investigations, arrests, and prosecutions. Increased transparency will allow defendants to mount robust reliability challenges and provide courts with the information they need to consider accuracy, field performance, and impact of algorithmic technologies.

2. Regulate the Process of Acquiring New Police Technology

Friedman and Ponomarenko advocate for "democratic accountability" in policing, arguing that police acquisition and use of technologies should be subject to democratic approval.[427] Allowing the public voice in setting policing policies promotes a sense of legitimacy,[428] which is ordinarily lacking when the police unilaterally deploy unknown technologies that greatly enhance their power. Democratic approval processes also create a meaningful back-and-forth debate that contemplates and records information about a new technology's efficacy versus its costs, enabling better judicial review when the technology is used in criminal cases.[429]

One way to accomplish democratic accountability is to have municipalities structure initial approval processes (for new acquisitions) as well as reviews (for existing technologies) through city councils or like bodies, allowing public comment and subjecting the technology's reliability, efficacy, and consequences to vetting in public fora.

The Policing Project, directed by Professor Friedman, has put forth a model policy to involve lawmakers and the public in regulating police technology at the front-end, before it is used in surveillance and investigations.[430] The policy works well to bolster the reliability of algorithmic technologies used in policing.

---

[426] *See* Justice in Forensic Algorithms Act, H.R. 2438, 117th Cong. § 2(d)(5) (requiring that the NIST publish all results of validity testing).

[427] *See* Friedman & Ponomarenko, *supra* note 243, at 1835.

[428] *See* Tom R. Tyler, Phillip Atiba Goff & Robert J. MacCoun, *The Impact of Psychological Science on Policing in the United States: Procedural Justice, Legitimacy, and Effective Law Enforcement*, 16 PSYCH. SCI. PUB. INT. 75, 76 ("Perceiving policing as legitimate includes having opportunities for voice and participation in designing policing policies.").

[429] *See* Friedman & Ponomarenko, *supra* note 243, at 1846–47.

[430] *See Working Draft: Authorized Policing Technology (APT) Act*, POLICING PROJECT 1 (2020), https://static1.squarespace.com/static/58a33e881b631bc60d4f8b31/t/5df2b10dca514d7fda060b45/1576186125973/APT+Act.pdf.

For one, municipalities should require city council approval of new technology acquisitions, or use of technology in a manner not previously approved.[431] The council or a designated committee should consider "whether the public safety benefits of the use of the policing technology outweigh the economic, social, and community costs, including potential negative impacts on civil liberties and civil rights and potential disparate impacts on particular communities or groups," with attention to whether the technology is reliable.[432] To evaluate use and reliability, law enforcement agencies seeking to use the new technology should be required to submit for public review: (1) a description of the technology that includes all known information about the technology's error rates as previously tested, reliability issues as deployed in other jurisdictions, and plans for validity testing of the technology as installed; (2) detailed use policies, including a delineation of authorized uses of the technology, which personnel will have access, supervisory review procedures, and how the technology will be treated as an investigative lead, or as part of probable cause or reasonable suspicion; (3) policies on the retention of data generated by the technology and access to the data generated; and (4) fiscal impact data.[433]

To promote legitimacy, if the technology is deployed, the public should be given every opportunity to undertake their own review of the proposed technology and participate in acquisition decision-making. Therefore, the city council shall disseminate all received information publicly, and promote and hold public hearings on the proposed acquisition.[434]

---

[431] Facial recognition technology and rapid DNA, for example, have multiple uses that implicate different legal and reliability concerns. FRT can be used for ongoing surveillance, which poses an even greater risk to privacy than one-off identifications, though reliability concerns may differ. *See* Garvie, *supra* note 115, at Risk Framework. Rapid DNA testing can be used for quick identifications of disaster victims against samples from next of kin, which does not require that profiles be entered into a larger database: no databasing eliminates the risk of wrongful arrests and convictions from faulty profiles. *See Snapshot: S&T's Rapid DNA Technology Identified Victims of California Wildfire*, DEP'T OF HOMELAND SEC. (Apr. 23, 2019), https://www.dhs.gov/science-and-technology/news/2019/04/23/snapshot-st-rapid-dna-technology-identified-victims.

[432] *APT Act*, *supra* note 429, at 1–2.

[433] *See id.* at 2. These requirements are adapted from the model policy to include emphasis on reliability factors and use of the technology as an investigative lead.

[434] *See* Tyler et al., *supra* note 427, at 85 (noting that "voice in the development of policies" via participation in "community meetings" or "other

The city council or appropriate municipal body should also undertake review of technologies previously acquired to weigh the benefits and costs of their continued use. After an initial review, technologies should be continuously reviewed and audited to ensure that the cost-benefit analysis holds: that the technology has provided some benefit to public safety, and that the attendant intrusions on individuals remain worthwhile. High-profile instances of error can and should change the calculus, as when the Boston City Council voted to ban the use of facial recognition after Robert Williams's false arrest in Detroit.[435] For Boston Police Commissioner William Gross, Williams's arrest illuminated the unreliability and bias inherent in the technology. He told councilors: "I didn't forget that I'm African American and I can be misidentified as well."[436]

VI. CONCLUSION

This Article seeks to highlight a disturbing disregard for the importance of reliability in police tools. ShotSpotter, facial recognition, and rapid DNA are case studies, showing how law enforcement agencies deploy error-prone technologies without vetting and with impunity. These technologies easily overcome even judicial inquiries that claim regard for reliability. Our Fourth Amendment jurisprudence lacks due consideration as to how unreliable leads infect the investigative process.

As algorithmic technologies advance and policing becomes increasingly automated, questions about reliability will evolve. Activists are already concerned that algorithms like facial recognition, on their own, are supplying suspicion with little corroboration, reducing the "totality of the circumstances" to one circumstance. Soon, the police may come to rely on automated suspicion algorithms: computers that will amalgamate data to predict to officers when and where a crime is occurring, and who is committing it.[437] These algorithms will outsource officers' judgments almost entirely, and they

---

mechanisms of seeking community guidance about what policies and practices are acceptable to the people living in the community" promotes a sense of legitimacy and trust in law enforcement).

[435] *See* Ally Jarmanning, *Boston Lawmakers Vote to Ban Use of Facial Recognition Technology by the City*, NPR (June 24, 2020), https://www.npr.org/sections/live-updates-protests-for-racial-justice/2020/06/24/883107627/boston-lawmakers-vote-to-ban-use-of-facial-recognition-technology-by-the-city.

[436] *Id.*

[437] *See Rich*, *supra* note 351, at 871–78.

will also make mistakes.[438] When it comes to deciding whether these algorithms provide probable cause or reasonable suspicion, judging their reliability will be crucial. And as with any algorithmic technology, law enforcement agencies and the policed public must make judgments about how much unreliability to tolerate in exchange for more efficient policing, and, in turn, how much unreliability affects efficiency. That debate continues over Chicago's ShotSpotter contract, and only time will tell if the public finally gets a say.

[438] *See id.* at 882–85.