

SHADOW FORENSICS: UNCOVERING 911 CALL ANALYSIS

Christopher Lau[†]

Junk forensics have played a central role in wrongful convictions. From bite mark evidence to hair microscopy, from arson investigation to toolmark comparison, courts have repeatedly allowed testimony about techniques and methodologies that have no basis in reliable science. The result has been hundreds of exonerations, and likely thousands more wrongful convictions. One of the newest forensic techniques is 911 Call Analysis. Proponents claim the technique can identify, via 911 callers' words and phrasing—including how someone requests help—whether they are the perpetrator of a crime. Despite study after study debunking the reliability of the technique, police departments and prosecutor offices continue to pay for call analysis trainings, and prosecutors continue to use 911 call analysis to seek convictions.

While admission of the technique into criminal cases has followed the same arc of junk forensics before it, 911 call analysis is particularly pernicious because it operates in the shadows. The underlying methods are never disclosed and the founders of the technique are never proffered as experts. Instead, 911 call operators or police officers testify about the culpability of 911 callers after attending a call analysis training. The effect is that the technique largely escapes Daubert or Frye evidentiary challenge. This Article introduces the concept of "shadow forensics" to understand this problem both in the 911 and broader forensic context. This Article makes three primary contributions. One, it offers the concept of shadow forensics: a set of practices that police and prosecutors use to shield forensic techniques, like call analysis, from evidentiary oversight. Second, it provides the first legal discussion

[†] Clinical Associate Professor, Co-Director, Wisconsin Innocence Project, University of Wisconsin Law School. 2026 Scholarly Papers Competition Finalist. For feedback and engagement, thank you to Valena Beety, Teneille Brown, Edward Cheng, Catherine Crump, Eve Hanan, Daniel Harawa, Randy Hertz, Alexis Hoag-Fordjour, Nathan Rouse, Maneka Sinha, Sonja Starr, and Lisa Washington. Thank you as well to workshop organizers and participants of CrimFest, Decarceration Law Professors Workshop, N.Y.U. Clinical Law Review Writers' Workshop, Chicagoland Junior Scholars Conference, and Vanderbilt University's Evidence Summer Workshop. For excellent research assistance, thank you to Ismail El-Hassan, Demetri Sweat, and Grace Woo.

of 911 Call Analysis and situates it within the broader problem of junk forensics in the criminal legal system by reviewing the empirical studies that have repeatedly debunked the technique. It then builds on the literature that examines the failure of courts to exclude unreliable forensics from the criminal system. Finally, it discusses proposals to stop the spread of the technique, and of other Shadow Forensics that escape evidentiary scrutiny.

INTRODUCTION.....	294
I. SHADOW FORENSICS: ESCAPING RELIABILITY ANALYSIS	302
A. Defining Shadow Forensics.....	303
B. Police Trainings as Entry Point for Shadow Forensics.....	307
C. Obfuscation by Police Expertise	311
D. Shadow Experts	316
II. THE UNRELIABILITY OF 911 CALL ANALYSIS	324
A. The Claims of “COPS”	325
B. Failing <i>Daubert</i> and <i>Frye</i>	332
III. THE BROADER IMPLICATIONS OF CALL ANALYSIS.....	336
A. Manufacturing Credibility Discounts	336
B. Deepening Legal Estrangement.....	340
IV. SHINING A LIGHT ON SHADOW FORENSICS: FORCING DISCLOSURE AND LIMITING ITS SPREAD	343
A. The Limits of <i>Daubert</i> and <i>Frye</i>	344
B. Transparency Litigation.....	345
C. Police Procurement Oversight	347
CONCLUSION.....	350

INTRODUCTION

“911, what is the address of your emergency?”
“I found my friend in the closet—she’s dead!”¹

In 2014, Kathy Carpenter called 911, breathless and in shock, after discovering the body of her friend. After the call, she agreed to speak with the authorities because she wanted to help them investigate the murder.² Instead, Carpenter was

¹ 48 Hours, *911 Call: “I Found my Friend in the Closet—She’s Dead!”*, CBS NEWS (Dec. 17, 2015), <https://www.cbsnews.com/video/911-call-i-found-my-friend-in-the-closet-shes-dead/> [<https://perma.cc/D334-2ASU>].
² Audio: *Carpenter’s Lawyers Argue Innocence*, ASPEN PUB. RADIO (June 22, 2014), <https://www.aspenpublicradio.org/2014-06-22/audio-carpenters-lawyers->

arrested after the Colorado Bureau of Investigation performed “911 Call Analysis” of her call and concluded that she was the murderer. According to the agent who performed the analysis, there were “39 indicators of guilty knowledge and no indicators of innocence” in her call.³ Using that report, prosecutors obtained a warrant for her arrest. Kathy Carpenter was arrested, charged, and held without bond.⁴ The agent was purportedly able to decipher Carpenter’s guilt strictly from her call because of his extensive training in 911 call analysis.⁵ Based on that training, he was able to identify more than three dozen indicators of Carpenter’s guilt, including: using “repetitive phrases,” requesting help for herself, identifying the decedent as her friend, acknowledging that her friend was dead, providing information about the people the decedent had conflict with, and having “unexpected knowledge.”⁶

That the agent could decipher Carpenter’s guilt from the transcript of a two-minute call would be remarkable if it were not completely wrong. Three months after Carpenter’s arrest and incarceration, the actual perpetrator confessed to the murder, and Carpenter was released.⁷

The call analysis report used to justify Kathy Carpenter’s wrongful arrest was based on 911 call analysis techniques used in hundreds of cases across the country.⁸ These techniques were developed by Susan Adams, a retired FBI Agent, and Tracy Harpster, a retired police officer in the Moraine, Ohio

argue-innocence [<https://perma.cc/3RSY-NQHD>]; see also Chad Abraham, *Kathy Carpenter Lawyer Faults CBI’s Role in Her Arrest*, ASPEN DAILY NEWS (June 21, 2014), https://www.aspendailynews.com/kathy-carpenter-lawyer-faults-cbi-s-role-in-her-arrest/article_49a7d9c7-f3c5-56f4-9f0d-90918046c72e.html [<https://perma.cc/L7ZH-2EWE>].

³ KIRBY QUINN LEWIS, COLORADO BUREAU OF INVESTIGATION 2014-42/11, at 1 (2014).

⁴ *Mistaken 911 Transcript Used to Charge Carpenter*, ASPEN PUB. RADIO (June 20, 2014), <https://www.aspenpublicradio.org/2014-06-20/updated-mistaken-911-transcript-used-to-charge-carpenter> [<https://perma.cc/3P8H-7M5Q>].

⁵ At the time of his analysis of Carpenter’s call, the agent had attended three trainings performed by Tracy Harpster, one of the founders of the 911 call analysis technique and had “independently reviewed approximately 60 calls through various Internet sites making 911 calls available online.” LEWIS, *supra* note 3, at 1.

⁶ *Id.* at 1–2.

⁷ *Mistaken 911 Transcript Used to Charge Carpenter*, *supra* note 4.

⁸ See Brett Murphy, *They Called 911 for Help. Police and Prosecutors Used a New Junk Science to Decide They Were Liars.*, PROPUBLICA (Dec. 28, 2022), <https://www.propublica.org/article/911-call-analysis-fbi-police-courts> [<https://perma.cc/YK3V-9UM5>] (finding instances of its use in jurisdictions across twenty-six states).

Police Department.⁹ They have published several studies and a book trumpeting their ability to identify deception and guilt in 911 callers.¹⁰ They developed the Considering Offender Probability Scale (COPS) to help law-enforcement agencies identify deception in 911 callers.¹¹ The scale identifies thirty-eight behaviors that indicate guilt and fifteen that indicate innocence. Some of the factors that the scale codes for are: immediate pleas for help, focus on caller, acceptance of death, extraneous information, inappropriately prioritized order, and initial delays in speaking.¹² According to Harpster's initial research, for example, the use of "extraneous information"¹³ during a call predicted deception in calls over 90% of the time.¹⁴

These claims have never been validated or replicated by any other researchers.¹⁵ Like forensic odontology, toolmark comparison, or firearm examination, there is no scientific

⁹ Susan H. Adams & Tracy Harpster, *911 Homicide Calls and Statement Analysis: Is the Caller the Killer?*, FBI LAW ENF'T BULL., June 2008, at 22-23.

¹⁰ See *id.*; Tracy Harpster, Susan H. Adams & John P. Jarvis, *Analyzing 911 Homicide Calls for Indicators of Guilt or Innocence: An Exploratory Analysis*, 13 HOMICIDE STUDS. 69, 69 (2009); TRACY HARPSTER & SUSAN H. ADAMS, *ANALYZING 911 HOMICIDE CALLS: PRACTICAL ASPECTS AND APPLICATIONS* 5 (2017).

¹¹ See Adams & Harpster, *supra* note 9, at 27.

¹² Daniel E. O'Donnell et al., "My Child is Missing": 911 Calls in Mysterious Disappearances of Children, 67 AGGRESSION & VIOLENT BEHAVIOR 1, 3, 6 (Oct. 13, 2022), <https://www.sciencedirect.com/science/article/pii/S1359178922000763/pdf?md5=2758cf0be8215858393237ab5757c12f&pid=1-s2.0-S1359178922000763-main.pdf> [<https://perma.cc/P2KC-7QK2>].

¹³ Harpster defines "extraneous information" as "information that does not appear to be relevant to the victim's situation." HARPSTER & ADAMS, *supra* note 10, at 51. He provides the example of a mother calling 911 about her son not being able to breathe but volunteering that she was recently in a car accident as "irrelevant" to the immediate situation. *Id.* at 52. Subsequent studies attempting to replicate Harpster's findings have identified the fact that determinations of "extraneous" or "irrelevant" information are subjective and rely on *post hoc* determinations of ground truth. Because of the subjective nature of the determination, the studies cannot consistently even define what information is "extraneous" or not. See O'Donnell et al., *supra* note 12, at 8, 10; Michelle L. Miller et al., "911 What's Your Emergency?": Deception in 911 Homicide and Suicide Staged as Homicide Calls, 25 HOMICIDE STUDS. 256, 268 (2021); Patrick M. Markey et al., *Deception Cues During High-Risk Situations: 911 Homicide Calls*, 33 PSYCH. SCI. 1040, 1042 (2022).

¹⁴ Tracy Harpster, *The Nature of 911 Homicide Calls: Using 911 Homicide Calls to Identify Indicators of Innocence and Guilt* 18 (Oct. 25, 2006) (M.S. thesis, University of Cincinnati). The effect size, if true, would represent the single largest predictor of deception in any study, and "among the largest effect sizes ever discovered in the social sciences." Patrick Markey et al., *The (Mis)utilization of Cues During Deception Detection in 911 Homicide Calls*, 29 HOMICIDE STUD. 185, 188 (2022).

¹⁵ See Markey et al., *supra* note 14, at 188.

basis for the claims of 911 call analysis.¹⁶ Study after study has failed to replicate the initial findings.¹⁷ Not only did the majority of the more than fifty behaviors in the COPS scale lack any statistically significant correlation with guilt or innocence, but those that did correlated with the opposite category.¹⁸ Despite being discredited by numerous studies, prosecutors and police departments continue to deploy the technique to secure arrests and convictions, and Harpster continues to conduct trainings across the country.¹⁹ Based on the seeming success of call analysis, companies are offering automated call analysis services that claim to identify deception in call transcripts, emails, and texts.²⁰

¹⁶ See, e.g., Oakland Dispatch Acad., *911 Homicide: Is the Caller the Killer?*, OAKLAND CMTY. COLL., <https://www.oaklandcc.edu/emergency-services-training/911/911-academy-courses/911-homicide-is-the-caller-the-killer> [<https://perma.cc/4H2R-MKD2>] (offering trainings in March 2026). See generally M. Chris Fabricant & Tucker Carrington, *The Shifted Paradigm: Forensic Science's Overdue Evolution from Magic to Law*, 4 VA. J. CRIM. L. 1 (2016); Erica Beecher-Monas, *Reality Bites: The Illusion of Science in Bite-Mark Evidence*, 30 CARDOZO L. REV. 1369 (2009); Jennifer D. Oliva & Valena E. Beety, *Discovering Forensic Fraud*, 112 NW. U. L. REV. ONLINE 1 (2017), https://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=1249&context=nulr_online [<https://perma.cc/5WV5-8DTD>]; Paul C. Giannelli, *Forensic Science: Daubert's Failure*, 68 CASE W. RESV. L. REV. 869, 875 (2018).

¹⁷ See *infra* Part II.

¹⁸ Miller et al., *supra* note 13, at 268 (“[W]hen we did find a significant result between homicide offenders who called 911 without admitting to the crime and uninvolved 911 callers it was in the opposite direction of what was previously found.”).

¹⁹ The use of faulty forensics or pseudoscience in securing convictions is not limited to call analysis. See Fabricant & Carrington, *supra* note 16, at 53 (describing prosecutors’ use of faulty bitemark evidence to secure convictions, despite obvious issues); Valena E. Beety & Jennifer D. Oliva, *Policing Pregnancy “Crimes,”* 98 N.Y.U. L. REV. ONLINE 29, 30 (2023), <https://nyulawreview.org/wp-content/uploads/2023/03/NYULawReview-Volume98-OlivaBeety.pdf> [<https://perma.cc/797K-D8J9>] (“[P]rosecutors have long relied on debunked or questionable science to secure fetal harm convictions against socioeconomically marginalized and racialized individuals.”).

²⁰ Michael Connor, *St. Pete Startup Has Created AI Software to Detect Deception in Texts, Emails and More*, TAMPA BAY BUS. & WEALTH (Jan. 10, 2024), <https://tbbwmag.com/2024/01/10/st-pete-startup-has-created-ai-software-to-detect-deception-in-texts-emails-and-more/> [<https://perma.cc/WH55-BZXW>] (including claim from founder of Deceptio.AI that his software can detect deception in written and verbal statements); see also Study: *Statement Analysis Study Results-DECEPTIO.AI and the “Huh” Factor*, DECEPTIO.AI <https://www.deceptio.ai/post/statement-analysis-deception-detection> [<https://perma.cc/67MU-MGS7>] (promoting Deceptio.AI’s software and citing to Tracy Harpster’s finding that “not answering a question was a 100% indicator of guilt. In Statement Analysis, this is a 100% indicator the person is withholding information. This information may be that he is guilty or that he has information which would help in the investigation.”). When questioned about the scientific foundation for Deceptio.AI, Mark Carson, its founder, has stated that “This is a bit of ‘Trust me, bro’ science.”

The problem of 911 call analysis does not end with Harpster. Recent investigative reporting by ProPublica has documented the use of the technique in more than one hundred cases across twenty-six states.²¹ The investigation shows that for over a decade now, the technique has quietly made its way into criminal cases without much fanfare.²² It is used in homicide cases where the prosecution otherwise has little direct evidence, and of the cases examined, it has been disproportionately used against female defendants, often mothers or caretakers.²³ Where the technique has been proffered, it has largely been admitted by judges without examination into its lack of reliability.²⁴ As many scholars have documented, criminal courts are notoriously bad at excluding unreliable forensic techniques, despite the courts' gatekeeping responsibilities under *Daubert* and *Frye*.²⁵ Consequently, criminal cases have been tainted by forensic techniques that have more basis in "magic" than in science.²⁶ Because the law privileges precedent over reliability, the more courts admit junk forensic techniques, the more additional courts unquestioningly do, creating a cascade of junk forensics.²⁷ One practical—and dangerous—consequence of this cascade is that the courts' unquestioning admission also shapes and affirms the perceived scientific validity of unreliable techniques like call analysis.²⁸

Christopher Beam, *My Mom Says She Loves Me. AI Says She's Lying.*, ATLANTIC (July 29, 2024), <https://www.theatlantic.com/ideas/archive/2024/07/ai-lie-detection-technology/679201/> [<https://perma.cc/6QA8-SSAN>].

²¹ See Murphy, *supra* note 8 (finding documented cases of the use of 911 call analysis by police departments or prosecutors offices in hundreds of cases).

²² *Id.*

²³ See Andrew Markquart, Jessica M. Salerno & Brett Murphy, The Junk Science of 911 Call Analysis, Presentation to the Annual Meeting of the Innocence Network Conference (Mar. 23, 2024). For a discussion of the gendered and racialized implications of 911 call analysis, see *infra* Part III.

²⁴ See *infra* Part I.

²⁵ See Giannelli, *supra* note 16, at 871, 875 (2018); Valena E. Beety & Jennifer D. Oliva, *Evidence on Fire*, 97 N.C. L. REV. 483, 484 (2019).

²⁶ Fabricant & Carrington, *supra* note 16, at 7.

²⁷ See Beecher-Monas, *supra* note 16, at 1372–73 (finding that in the context of bitemark testimony, "Courts frequently admit bite-mark testimony simply because other courts have done so. They find that it is 'not novel' or they let in this shaky testimony precisely because it is 'not science.'"); see also Samuel D. Hodge, Jr. & Amelia Holjencin, *A Post-Mortem Review of Forensic Hair Analysis – A Technique Whose Current Use in Criminal Investigations is Hanging on by a Hair*, 64 ST. LOUIS U. L.J. 219, 225–26 (2020) (detailing the practice of courts repeatedly admitting microscopic hair analysis without entertaining *Daubert* or *Frye* challenges).

²⁸ See Jennifer L. Mnookin, *Scripting Expertise: The History of Handwriting Identification Evidence and the Judicial Construction of Reliability*, 87 VA. L. REV.

One of the challenges with 911 call analysis is that prosecutors have used it in cases without naming the technique or proffering an expert.²⁹ This omission prevents defendants from challenging its reliability in reliability hearings or through confrontation at trial. Notwithstanding the failure of criminal courts to exclude unreliable evidence via *Daubert* or *Frye* hearings, they at least provide an opportunity for courts to evaluate the reliability of a technique or of an expert. In the case of 911 call analysis, Harpster has never testified in court himself about the technique.³⁰ Emails obtained via public records request show that Harpster purposely avoids being proffered as an expert.³¹ One Iowa County prosecutor emailed a colleague to explain that “[Harpster] knows there will be a great legal hurdle getting the research admitted” and “he doesn’t want a legal precedent prohibiting the admission of his research.”³² In some cases, prosecutors have purposely buried the 911 call analysis technique and instead have put on police officers or call dispatchers to testify generally about why they did not believe a 911 caller.³³ The strategy not only insulates the technique from examination but makes it difficult to even identify in which cases the technique has been deployed.

Like the junk forensics before it, the use of 911 call analysis defies notions of fairness and reliability for evidence in

1723, 1729 (2001). Writing in the context of handwriting evidence, Professor Mnookin observes that “judicial determinations play an enormous role in constructing broader cultural perceptions of reliability. The courtroom does not just import reliable evidence but also creates it.” *Id.*

²⁹ See *infra* Part II. Research on 911 calls in general is hampered by its piecemeal nature. 911 was not codified as the national emergency number until 1999. Since their advent in 1968, 911 calls have been, and continue to be routed through different agencies, jurisdictions, and governments. Currently, there are more than 6,000 independent call centers responsible for 911 calls across the United States. See *Next Generation 9-1-1—The Future for Emergency Communications*, NENA (2009), https://cdn.ymaws.com/www.nena.org/resource/resmgr/ng9-1-1_project/ng9-1-1_overview_030909.ppt [<https://perma.cc/E936-J45S>].

³⁰ See Murphy, *supra* note 8.

³¹ See *id.*

³² See *id.*

³³ *Id.* Notably, in the case of Russ Faria, who was wrongfully convicted and then retried for the murder of his wife after he called 911 to report her death, prosecutors put on a 911 call dispatcher to explain why the call was staged. *Id.* The prosecutor got that testimony admitted by, in her own words, “getting creative . . . without calling it ‘science.’” *Id.* Faria was incarcerated for 3.5 years before his conviction was overturned. *Id.*; see also Jeanette Cooperman, *The Unimaginable, Infamous Case of Pam Hupp*, ST. LOUIS MAG. (Jan. 19, 2017).

the criminal legal system.³⁴ It has contributed to wrongful arrests³⁵ and convictions³⁶ and as the technique spreads, it may have a chilling effect on people calling 911 to obtain help.

This Article makes three primary contributions. On a theoretical level, it introduces the concept of “shadow forensics,”³⁷ a set of practices that police and prosecutors use to shield forensic techniques like 911 call analysis from evidentiary oversight. It provides a framework for the legal analysis of 911 call analysis and other junk forensics. It applies that framework through case studies of trials where call analysis was used to convict someone.³⁸ By doing so, it is the first article in the legal academic literature to examine the issue of call analysis within the larger context of forensic evidence in the criminal legal system. On a descriptive level, it shows how 911 call analysis and its impacts have spread all over the country. Finally, on a prescriptive level, this Article offers targeted interventions

³⁴ This article employs the terms “junk forensics” and not “junk science” because outside of the context of DNA testing, the techniques employed are not scientific. See Boaz Sangero, *Safety from Flawed Forensic Sciences Evidence*, 34 GA. ST. U. L. REV. 1129, 1195, 1198 (2018) (“[T]he majority of the ‘forensic sciences’ branches . . . have been developed especially for the purpose of solving crimes; they were not developed by scientists.”).

³⁵ See, e.g., Chad Abraham, *Former Murder Defendant Moves to Seal Records*, ASPEN DAILY NEWS (Dec. 18, 2017), https://www.aspendailynews.com/former-murder-defendant-moves-to-seal-records/article_4032520c-3af3-5165-942b-87106d6c684f.html [<https://perma.cc/2AER-3BCT>] (describing efforts by Kathy Carpenter to seal records after she was released from prison and all homicide charges were dismissed against her).

³⁶ See Murphy, *supra* note 8.

³⁷ The phrase borrows from the literature on shadow dockets, which find that courts’ decisions on procedural matters, while highly consequential, are “unseen, unsigned, and almost always unexplained.” See STEPHEN VLADECK, *THE SHADOW DOCKET: HOW THE SUPREME COURT USES STEALTH RULINGS TO AMASS POWER AND UNDERMINE THE REPUBLIC* 12–13 (2023) (describing the Supreme Court’s growing use of the shadow docket to shape the law). It applies the unseen aspects of the shadow docket to the ways that the state hides forensic practices, and the resulting implications that has in criminal cases. Cf. Faiza W. Sayed, *The Immigration Shadow Docket*, 117 NW. U. L. REV. 893, 893 (2023) (describing the immigration shadow docket and the tens of thousands of unpublished immigration appeals decisions that are rendered annually).

³⁸ While this Article does not provide a comprehensive empirical analysis, it uses case studies to conceptualize the problem of shadow forensics. The case studies encompass the trial transcripts and appellate records of four cases, in four different states, where call analysis was discovered. In all but one case, the practice remained undiscovered until the appellate stage. See *infra* subparts I.B–C. These cases remain the tip of the iceberg: because prosecutors purposely avoid proffering experts in call analysis, it is difficult to know all the cases where the technique has been used against people.

that surface the practice of shadow forensics at different stages in the legal process where they are utilized by the police and prosecution.

This Article proceeds in four parts. Part I introduces the concept of “shadow forensics,” which weaves together strands of literature on police training, expert testimony, and forensic evidence to explain how junk forensics enter cases without judicial examination. It engages with existing scholarship on the failure of criminal court judges to exercise their gatekeeping function under *Daubert* and *Frye*, and how forensic practices such as 911 call analysis evade detection and evidentiary challenges. Part II provides an account of the surreptitious rise and use of 911 call analysis in criminal trials. It examines the empirical studies that have repeatedly failed to replicate the purported success of 911 call analysis. It concludes that the technique is unreliable and should be inadmissible under a *Daubert* or *Frye* framework. Part III examines the deeper harms that call analysis causes in the criminal legal system: epistemic and dignitary harms that result when the words of 911 callers are interpreted through the lens of call analysis; legal estrangement among vulnerable communities against whom call analysis is most often used; and the resulting policy concerns that arise when 911 calls are weaponized against callers. Finally, Part IV evaluates solutions to stop the spread of 911 call analysis and other shadow forensics. It focuses on ways the technique can be exposed. First, it evaluates how strengthening *Daubert* and *Frye* and enacting broader statutory discovery rules can be used to shine a light on shadow forensics. Acknowledging the limited effect of judicial gatekeeping, even after the 2009 National Academy of Sciences (NAS) and 2016 Presidential Council of Advisors on Science and Technology (PCAST) reports that attempted to ensure the reliability of forensic science,³⁹ Part IV calls for solutions that go beyond strengthening evidentiary rules. Second, it draws on the burgeoning scholarship on transparency litigation and the ways that litigation can make opaque police investigative practices more public.⁴⁰ Third, it discusses the creation of democratic oversight and regulation over police technologies and

³⁹ NAT'L RSCH. COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009); PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., EXEC. OFFICE OF THE PRESIDENT, FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS (2016).

⁴⁰ See Hannah Bloch-Wehba, *Visible Policing: Technology, Transparency, and Democratic Control*, 109 CALIF. L. REV. 917, 922 (2021).

trainings—some of the earliest ways shadow forensics enter the criminal system. It concludes that a multipronged strategy to force disclosure is necessary to combat the ways that shadow forensics evade judicial gatekeeping.

I

SHADOW FORENSICS: ESCAPING RELIABILITY ANALYSIS

Over the last decade, 911 call analysis has been employed by police and prosecutors in more than one hundred cases across the country, despite repeated evidence of its unreliability.⁴¹ This Part identifies shadow forensics as a framework to understand how patently unreliable forensic practices make their way into criminal cases. Using call analysis as a case study, it identifies the entry points at which forensics enter a case sans reliability analysis. Forensic techniques are hidden at various stages throughout a case, from initial investigation to pretrial hearings and through trial. They are hidden through parallel construction by the police,⁴² by trade secret claims of private companies,⁴³ and by prosecutors affirmatively seeking to avoid judicial inquiry.⁴⁴ This Part traces a genealogy of the technique by drawing together the literature on police training,⁴⁵ police expertise,⁴⁶ and forensics. It details how the technique has made its way into criminal cases, how courts have largely failed to exercise their gatekeeping responsibilities,⁴⁷ and most dangerously, how the technique escapes evidentiary challenge by the use of shadow experts.⁴⁸

The wrongful arrests and convictions that have resulted from 911 call analysis begin with an officer who has attended a call analysis training. Based on a few hours of training, the officer concludes that a particular 911 caller, based on their cadence, their words, or their use of specific phrases, is the true culprit. Trial courts repeatedly fail to ensure that the technique is reliable before allowing law enforcement

⁴¹ See Murphy, *supra* note 8; *infra* Part II.

⁴² HUM. RTS. WATCH, DARK SIDE: ORIGINS OF EVIDENCE IN US CRIMINAL CASES 10, 12 (2018).

⁴³ See Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972, 2027–30 (2017).

⁴⁴ See Murphy, *supra* note 8.

⁴⁵ See Ingrid V. Eagly & Joanna C. Schwartz, *Lexipol: The Privatization of Police Policymaking*, 96 TEX. L. REV. 891, 895–96 (2018).

⁴⁶ See generally Anna Lvovsky, *The Judicial Presumption of Police Expertise*, 130 HARV. L. REV. 1995 (2017).

⁴⁷ See *infra* subpart I.A.

⁴⁸ See *infra* subpart I.D.

witnesses to testify that they can identify a person's guilt from their 911 call.

A. Defining Shadow Forensics

Perhaps the most pernicious way that prosecutors introduce 911 call analysis is by evading judicial examination altogether. Prosecutors have secured convictions using the practice without proffering any experts, without providing any notice they would be eliciting call analysis testimony, and without litigating *Daubert* or *Frye* hearings.⁴⁹ Instead, they elide the evidentiary process by having police witnesses testify, ostensibly as mere fact witnesses, about the indicators of guilt in a defendant's 911 call.

As many scholars have documented, the adversarial process is a poor system for determining scientific validity.⁵⁰ Overburdened and under-resourced defense counsel may fail to challenge forensic testimony because they lack the time, money, or wherewithal to consult with experts.⁵¹ The problem becomes even more pronounced when prosecutors purposely hide evidence so that defendants cannot challenge it.⁵² In the Fourth Amendment context, scholars have questioned the use of secretive police surveillance and searches. Professors Patrick Toomey and Brett Kaufman assert that government searches and surveillance kept secret from defendants undermine their fundamental right to test the government's evidence.⁵³ They argue that the right to notice is integral to the Fourth Amendment and, indeed, predates the Constitution.⁵⁴ Similarly, Professor Barry Friedman describes the incompatibility of a functional

⁴⁹ See, e.g., Transcript of Proceedings, *State v. Baldwin*, No. 27-CR-18-23849 (Minn. Dist. Ct. Jan. 15, 2021); Partial Report of Proceedings, *People v. Logan*, No. 2019-CF-1648 (Ill. Cir. Ct. Macon Cty., Jan. 5, 2024) [hereinafter Partial Report: People's Witnesses].

⁵⁰ See, e.g., Fabricant & Carrington, *supra* note 16, at 16; Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CALIF. L. REV. 721, 758 (2007); Oliva & Beety, *supra* note 16, at 4–5; Maneka Sinha, *Radically Reimagining Forensic Evidence*, 73 ALA. L. REV. 879, 883 (2022).

⁵¹ See Anthony C. Thompson, *The Promise of Gideon: Providing High-Quality Public Defense in America*, 31 QUINNIPIAC L. REV. 713, 718–19 (2013).

⁵² See Maneka Sinha, *Challenging Automated Suspicion*, 99 S. CAL. L. REV. (Forthcoming 2026) (reporting survey findings that defense counsel are unable to challenge evidence that “police and prosecutors actively hide”).

⁵³ Patrick Toomey & Brett Max Kaufman, *The Notice Paradox: Secret Surveillance, Criminal Defendants & the Right to Notice*, 54 SANTA CLARA L. REV. 843, 845–46 (2015).

⁵⁴ *Id.* at 850.

democracy wherein the polity cannot learn who, and by what means, the police surveil the community as “policing [] in the shadows.”⁵⁵ When people seek to make police methods more transparent, they are repeatedly met with resistance.⁵⁶

In the same way, deliberately concealed forensic practices pose a fundamental threat to the criminal legal system. Shadow forensics is the process by which the State effectively elides gatekeeping; these forensics are not proffered, are not challenged, and therefore escape evidentiary examination by a judge before being presented as expertise to fact finders in pretrial hearings or, more rarely, at trial. Unchallenged, 911 call testimony can prove powerful in convincing jurors of a defendant's guilt.

This subpart provides an overview of the evidentiary landscape and argues that our evidence rules are premised on ensuring that jurors make decisions based on reliable evidence.⁵⁷ To work in our adversarial system, the rules of evidence contemplate disclosure and the ability for questioning and consultation by both sides. Shadow forensics like call analysis fundamentally subvert that system.

The process of shadow forensics is the latest endeavor in the struggle to overcome rules that ensure reliable trial testimony. At each stage, proponents of junk forensics have sought to evade judicial evidentiary gatekeeping. Starting with the Supreme Court's holding in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, there has been a concerted effort to ensure the reliability of expert evidence, but with every amendment of the Federal Rules of Evidence, proponents of junk forensics have sought ways to evade the rules. The use of shadow forensics is one of the most obvious ways parties dodge the reliability standards.

In 1993, the Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*⁵⁸ seemed to “radically change[] the standard for admissibility of scientific testimony.”⁵⁹ For seventy

⁵⁵ Barry Friedman, *Secret Policing*, 2016 U. CHI. LEGAL F. 99, 108.

⁵⁶ See generally JOANNA SCHWARTZ, *SHIELDED: HOW THE POLICE BECAME UNTOUCHABLE* (2023) (describing widespread refusal of police departments to provide officer misconduct reports); see also Bloch-Wehba, *supra* note 40, at 946 (documenting the continual refusal of the NYPD to provide stop-and-frisk data to the New York City Council). Similarly in the Eighth Amendment context, States increasingly “shroud modern executions with multiple levels of secrecy,” including the ways that they procure the drugs used to kill people. See William W. Berry III & Meghan J. Ryan, *Cruel Techniques, Unusual Secrets*, 78 OHIO ST. L. J. 403, 405–06 (2017).

⁵⁷ See Keith A. Findley, *Judicial Gatekeeping of Suspect Evidence: Due Process and Evidentiary Rules in the Age of Innocence*, 47 GA. L. REV. 723, 725 (2013).

⁵⁸ *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993).

⁵⁹ *United States v. Barnette*, 211 F.3d 803, 815 (4th Cir. 2000).

years, courts had followed the test enumerated by the D.C. Circuit in *Frye v. United States* in 1923;⁶⁰ scientific testimony merely had to meet the “general acceptance” test.⁶¹ By the end of the twentieth century, the nature of litigation had changed; vast increases in the use of expert witnesses became the norm, both in civil torts and in criminal cases. To many critics, *Frye*’s “general acceptance” test became an all-accepting test, regardless of the unscientific and unreliable nature of the expert testimony.⁶² They noted that while *Frye* had become the rule nationally, different state courts and federal circuits had been inconsistently applying its test.⁶³ Furthermore, the general acceptance test resulted in courts taking a *laissez-faire* attitude toward their gatekeeping responsibilities.⁶⁴ One result of the lax judicial gatekeeping was the rise of faulty forensic evidence introduced by expert witnesses.⁶⁵

In *Daubert*, the Supreme Court attempted to address concerns about the admission of junk science. To determine reliability, it laid out four factors to guide courts in exercising their gatekeeping function: (1) empirical testing; (2) error rates; (3) peer publication and review; and (4) general acceptance in the scientific community.⁶⁶ In the aftermath of *Daubert*, parties circumvented judicial efforts at gatekeeping by presenting evidence as non-scientific. Instead of qualifying experts in scientific fields, parties instead proffered “expert[s] in lay witness clothing”⁶⁷ to sidestep any judicial examination. In the context of pattern-matching forensic techniques like bitemarks and hair comparison, although experts were presented as “scientists”

⁶⁰ *Id.* (describing the Supreme Court’s departure from *Frye*).

⁶¹ See *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923) (holding that evidence be admitted evidence so long as it has “gained general acceptance” by the relevant scientific community).

⁶² That the *Frye* test would need updating was perhaps unsurprising given that it comprised “only a total of 669 words and did not make ‘a single reference to case law or precedent, nor any references to scientific literature.’” Jim Hilbert, *The Disappointing History of Science in the Courtroom: Frye, Daubert, and the On-going Crisis of “Junk Science” in Criminal Trials*, 71 OKLA. L. REV. 759, 770 (2019) (quoting Jill Lepore, *On Evidence: Proving Frye as a Matter of Law, Science, and History*, 124 YALE L.J. 1092, 1140 (2015)).

⁶³ See *id.* at 773.

⁶⁴ See *id.* at 773 & n.76.

⁶⁵ Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1, 32 (2009).

⁶⁶ *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579, 593–94 (1993).

⁶⁷ Joëlle Anne Moreno, *What Happens When Dirty Harry Becomes an (Expert) Witness For the Prosecution?*, 79 TUL. L. REV. 1, 12 (2004) (quoting FED. R. EVID. 701 advisory committee’s note to 2000 amendment).

who made elaborate technical presentations to jurors and concluded their opinions with “reasonable scientific certainty,” judges refused to inquire about the reliability of the techniques because ultimately, the pattern-matching could be done by jurors and was therefore “non-scientific.”⁶⁸ The Supreme Court addressed this end-run around reliability checks in *Kumho Tire v. Carmichael* when it extended the *Daubert* test to all expert testimony, whether scientific or not.⁶⁹ In finding that there should not be differing standards for expert testimony based on “scientific knowledge” versus “technical” or “other specialized knowledge,”⁷⁰ the Court sought to eliminate a distinction that allowed parties to evade evidentiary inquiry by simply labeling an expert’s testimony non-scientific.⁷¹ The Court reaffirmed the role of judicial gatekeeping and held that non-scientific testimony must also meet standards of reliability before being admitted.⁷² Further reinforcing the reliability standards, in 2000, Rule 702 of the Federal Rules of Evidence was amended to incorporate the reliability requirements of *Daubert* and *Kumho Tire*.⁷³

The effect of *Daubert*, *Kumho Tire*, and the revised Rule 702 on the reliability of forensic evidence has been inconsistent, at best.⁷⁴ There have been stark disparities between civil and criminal cases. In civil cases, these evidentiary safeguards have raised the frequency and caliber of evidentiary challenges and increased judicial scrutiny,⁷⁵ but in criminal cases where under-resourced public defenders lack the funding to consult and hire experts, there have been far fewer evidentiary challenges.⁷⁶ Consequently, unreliable junk forensics “continue to enter court with remarkably little scientific scrutiny or proof of accuracy and validity.”⁷⁷ Criminal defendants are less likely to challenge forensic testimony; but even where they have, criminal court judges have given their gatekeeping responsibilities

⁶⁸ Fabricant & Carrington, *supra* note 16, at 65, 68–69 (detailing the rationale in cases across the country, which relied on the “eyeball test” originally devised in 1869 by the Supreme Judicial Court of Massachusetts in *Commonwealth v. Dorsey*, 103 Mass. 412 (1869)).

⁶⁹ *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, 147–49 (1999).

⁷⁰ *Id.* at 148.

⁷¹ *See id.*

⁷² *Id.*

⁷³ FED. R. EVID. 702 advisory committee’s note to 2000 amendment.

⁷⁴ *See* Jennifer L. Mnookin, Editorial, *Science, Justice, and Evidence*, 382 SCIENCE 741, 741 (2023).

⁷⁵ *See* Oliva & Beety, *supra* note 16, at 4–5.

⁷⁶ *See id.*

⁷⁷ Mnookin, *supra* note 74, at 741.

short shrift, largely deferring to the State to admit proper forensic evidence.⁷⁸ Despite the largely permissive standard in criminal cases, proponents of unreliable forensic techniques avoid even the possibility of having evidence barred by hiding the use of a forensic practice from defense counsel and effectively preventing any form of judicial gatekeeping.

Although there is disagreement about how courts should approach their gatekeeping responsibilities under Rule 702—whether that entails increased judicial discretion, increased reliance on outside experts, or independent judicial research—all the approaches contemplate judicial scrutiny of the forensic practice in question.⁷⁹ With lay witness testimony, judges can assess, on the basis of the proposed testimony alone, whether it is “rationally based on the witness’s perception,” and if true, whether it would be helpful in the case.⁸⁰ By contrast, to determine whether an expert’s opinion about “scientific, technical, or other specialized knowledge” is “the product of reliable principles and methods”⁸¹ the court must have the requisite foundational information so that it can “critically examine an expert’s methodology and conclusions using an ‘exacting standard.’”⁸² By hiding call analysis in the shadows, prosecutors prevent courts from critically examining call analysis testimony.

Shadow forensics like call analysis infiltrate the criminal legal system through a number of entry points that allow them to evade reliability analysis. This Part identifies three of those entry points as (1) police training; (2) police expertise; and (3) shadow experts. It uses call analysis as an illustration of those points, and identifies the ways other shadow forensics evade reliability analysis.

B. Police Trainings as Entry Point for Shadow Forensics

Call analysis trainings are offered to police departments across the country.⁸³ According to Harpster his “training is

⁷⁸ See Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95 AM J. PUB. HEALTH (SUPP.) S107, S109 (2005). In some jurisdictions, like Georgia, legislatures have mandated that expert testimony in criminal cases is always admissible, making it impossible for defense counsel to challenge forensic evidence. See Oliva & Beety, *supra* note 16, at 4.

⁷⁹ See Edward K. Cheng, *Independent Judicial Research in the Daubert Age*, 56 DUKE L.J. 1263, 1266 (2007). See also generally Findley, *supra* note 57.

⁸⁰ FED. R. EVID. 701; see also Beecher-Monas, *supra* note 16, at 1404.

⁸¹ FED. R. EVID. 702.

⁸² Cheng, *supra* note 79, at 1265 (quoting *Weisgram v. Marley Co.*, 528 U.S. 440, 442 (2000)).

⁸³ Murphy, *supra* note 8.

unique and nobody else is doing it. . . because I'm the only one who has done the research.”⁸⁴ In records obtained from police departments, training flyers claim that “30% of all murderers call in 911 pretending to be innocent.”⁸⁵ The advertisements for the trainings note that Harpster “travels extensively, presenting at international, national, and regional homicide conferences.”⁸⁶ He claims that he has “assisted over 700 homicide investigators by analyzing the 911 call to help solve the crime.”⁸⁷

From: "wahi.homicide@gmail.com@mail186.atl61.mcsv.net" on behalf of "Wisconsin Association of Homicide Investigators"
Sent: Tue, 18 Apr 2017 00:50:34 +0000
To: [REDACTED]
Subject: 911 Is the Caller the Killer?

Wisconsinites!

In the past few years I have assisted each of you (or your department) on a 911 homicide call analysis. I will be teaching a two day course for detectives, officers, dispatchers (the true first responder to a homicide), prosecutors and coroners on November 14 & 15, 2016 at the North central Technical College in Wisconsin.

The link for more information is:

<https://www.ntc.edu/ce/conferences/homicide-caller-killer>.

I have a great class, I have shared my program with the FBI, the RCMP, the International Homicide Conference, the National Homicide Conference (X4), the Wisconsin Association of Homicide Investigators, many APCO/NENA conferences, prosecutor conferences (my class is accredited for continuing ed credits for prosecutors) and coroners. 30% of all murderers call in 911 pretending to be innocent.....sometimes this statement is the ONLY statement that a bad guy ever makes to law enforcement!

Figure 1. Training Flyer “911 Is the Caller the Killer?”

⁸⁴ *Id.*

⁸⁵ See E-mail from Wisconsin Ass’n of Homicide Investigators to Eau Claire Police Dep’t Deputy Chief (Apr. 18, 2017) (on file with author).

⁸⁶ E-mail from Northcentral Tech. Coll. to Eau Claire Police Dep’t Officer (Sep. 1, 2017) (on file with author).

⁸⁷ *Id.*

Police trainings play a much larger role in governing police behavior and practice than court decisions do.⁸⁸ Despite decades-long calls for robust oversight and democratic rule-making, police training remains predominantly outside the control of the democratic process.⁸⁹ Professors Ingrid Eagly and Joanna Schwartz have documented that private companies such as Lexipol, which provides training materials for thousands of police departments nationwide,⁹⁰ have amassed disproportionate influence over policing. In the context of the use of force, Lexipol opposes constraints on its use, and instead encourages “maximum flexibility”.⁹¹ While state and local governments have sought to limit the use of deadly police force, Lexipol has used its trainings, advertisements, and media, including “self-published articles, blog posts, white papers, webinars and public statements” to resist reform.⁹² Similarly, Axon (formerly TASER International), the company that manufactures the eponymous TASER, has spent tens of thousands of dollars to host trainings for law enforcement, distribute materials to police chiefs, and invite medical examiners to conferences to promote the existence of “excited delirium” in order to encourage the use of TASER sales and to fend off lawsuits related to officers killing or seriously injuring people with its weapons.⁹³

Companies like Axon and Lexipol pour money into police trainings and materials because they need to manufacture demand for their products in a monopsonistic market.⁹⁴ With products that have limited to no demand in the private market, these companies have a vested interest in buoying demand by the main buyer in the market—the police. The result is that companies market and manufacture products geared toward the aims and goals of the police. Because there is minimal oversight of the trainings and products that police use, these companies can reap huge profits despite concerns over

⁸⁸ See Ingrid V. Eagly & Joanna C. Schwartz, *Lexipol’s Fight Against Police Reform*, 97 IND. L.J. 1, 26 (2022).

⁸⁹ See Eagly & Schwartz, *supra* note 45, at 895–96.

⁹⁰ *Id.* at 894.

⁹¹ Eagly & Schwartz, *supra* note 88, at 26.

⁹² *Id.*

⁹³ See Rachel Kincaid, “Excited Delirium” Training Encourages Law Enforcement Violence, 99 TUL. L. REV. 49, 59–60, 64–65 (2024) (explaining that the diagnosis of excited delirium is an instance of pseudoscience that is rooted in medical racism, disavowed by national medical organizations, and banned in several states).

⁹⁴ See Christina Koningisor, *Police Secrecy Exceptionalism*, 123 COLUM. L. REV. 615, 671 (2023).

the reliability,⁹⁵ invasiveness,⁹⁶ and cost⁹⁷ of their products. In another example, an audit by the New Jersey State Comptroller unearthed the role that Street Cop Training, a police training company, played in conducting trainings that promoted unconstitutional car stops, glorified violence, and promoted the use of excessive force against civilians.⁹⁸ Between 2019–2022, the company received more than \$300,000 from police departments in New Jersey alone. The Comptroller concluded that police trainings in New Jersey were conducted “without oversight or regulation.”⁹⁹

The gospel of 911 call analysis is spread through unregulated police trainings of the kind hosted by companies like Lexipol and Street Cop. These trainings claim that the COPS checklist can help catch the true perpetrator in 30% of homicide cases.¹⁰⁰ The advertisements rely on claims that are based on self-published studies that lack any rigorous basis.¹⁰¹ Despite the lack of peer-reviewed corroboration, there were more than 130 call analysis trainings for police departments between 2006 and 2017.¹⁰² Police departments pay up to \$3,500 for each training.¹⁰³ Open records requests to police departments have thus far yielded email advertisements for call analysis trainings across multiple states.¹⁰⁴ The trainings have proven popular because they give credence to officers’ suspicions, particularly in cases where there is otherwise no evidence or indication that

⁹⁵ See Kincaid, *supra* note 93, at 64–65; Maneka Sinha, *The Dangers of Automated Gunshot Detection*, 5 J. L. & INNOVATION 63, 99 (2023).

⁹⁶ See generally Chaz Arnett, *Black Lives Monitored*, 69 UCLA L. REV. 1384 (2023).

⁹⁷ Diba Mohtasham, *Chicago Will Drop Controversial ShotSpotter Gunfire Detection System*, NPR (Feb. 15, 2024), <https://www.npr.org/2024/02/15/1231394334/shotspotter-gunfire-detection-chicago-mayor-dropping> [<https://perma.cc/VG2G-9828>]. See generally KEVIN D. WALSH, N.J. OFF. OF THE STATE COMPTROLLER, THE HIGH PRICE OF UNREGULATED PRIVATE POLICE TRAINING TO NEW JERSEY (2023).

⁹⁸ WALSH, *supra* note 97, at 1.

⁹⁹ *Id.* at 2.

¹⁰⁰ See E-mail from Wisconsin Ass’n of Homicide Investigators to Eau Claire Police Dep’t Deputy Chief, *supra* note 85.

¹⁰¹ See *infra* subpart II.B.

¹⁰² Murphy, *supra* note 8.

¹⁰³ Brett Murphy, *How Jessica Logan’s Call for Help Became Evidence Against Her*, PROPUBLICA (Nov. 20, 2022), <https://www.propublica.org/article/911-call-analysis-jessica-logan-evidence> [<https://perma.cc/97NN-4A4H>].

¹⁰⁴ Responses on file with author. Research assistants sent out open records requests to more than 400 police departments, sheriff departments, and call operator offices across multiple states.

a caller is guilty, and help police and prosecutors to introduce those suspicions to jurors through 911 calls.¹⁰⁵

The trainings serve as an early entry point for shadow forensics. Because the trainings are sanctioned and paid for by police departments, officers uncritically absorb the claims of the trainers. With no countervailing data, officers who attend the training for just a few hours come away confident in the principles of call analysis and confident in their ability to detect deception in 911 calls; that confidence infects their investigations and their testimony at trial.¹⁰⁶ However, because the trainers of call analysis are never proffered as experts, the practice remains hidden, and the people whose calls are suspected based on the practice may never learn that call analysis was the basis for their arrest.¹⁰⁷

C. Obfuscation by Police Expertise

Once shadow forensics gain an initial entry point via police trainings, they are cemented by in-court invocations of police expertise. Even where defense counsel challenge the scientific underpinning of 911 call analysis, courts have skirted the issue by admitting expert testimony absent *Daubert* or *Frye* inquiry because the practice was accepted by the police. In this way, courts take what Professor Anna Lvovsky has deemed the virtuous account of police expertise,¹⁰⁸ wherein invocations of

¹⁰⁵ See Brett Murphy, *Prosecutors and Judges Push for Conviction Reviews, Ban on Junk Science of 911 Call Analysis*, PROPUBLICA (Feb. 24, 2023), <https://www.propublica.org/article/911-call-analysis-ban-review-convictions> [<https://perma.cc/28R7-J7F6>].

¹⁰⁶ Cf. Saul M. Kassin & Christina T. Fong, “I’m Innocent!”: *Effects of Training on Judgments of Truth and Deception in the Interrogation Room*, 23 L. & HUM. BEHAV. 499, 510 (1999) (finding, however, that increased confidence does not correlate with increased accuracy); see also Bella M. DePaulo et al., *Cues to Deception*, 129 PSYCH. BULL. 74, 83 (2003).

¹⁰⁷ This dynamic is not unique to call analysis. Steven Rhoads, who has trained police departments for decades in subconscious communication detection, has not testified as an expert in the area since 1988. Trial Transcript, State v. Alexander, No. 12230 (Miss. Cnty. Cir. Ct., June 20, 1988); Jordan Smith, *The Junk Science Cops Use to Decide You’re Lying*, INTERCEPT (Aug. 12, 2020), <https://theintercept.com/2020/08/12/blueleaks-law-enforcement-police-lie-detection/> [<https://perma.cc/4RCL-D7HX>]. Records indicate that Rhoads last testified in a case in 1996, when he testified as a shooting reconstruction expert. See Mootye v. Dotson, 73 Fed. App’x 161, 165 (7th Cir. 2003). Similarly, Reid and Associates, the purveyors of the Reid Technique, an interrogation method, train interrogating officers about how to testify about the technique. See, e.g., THE REID TECHNIQUE TIPS, *Preparing to Testify in Confession Cases Part One* (YouTube, Nov. 12, 2024), <https://www.youtube.com/watch?v=27rRYeZFBSY>.

¹⁰⁸ Anna Lvovsky, *Rethinking Police Expertise*, 131 YALE L.J. 475, 535 (2021).

police expertise “are essentially talismanic: articles of faith urging the courts against imposing their own judgment, as one government attorney warned, on how officers ‘perform the work for which they are specially trained.’”¹⁰⁹ Police expertise thus obscures the use of shadow forensics from judicial examination. Compounding the problem of unregulated police trainings, judges increasingly mark police expertise by the mere mention of training, without evaluating the substance of the trainings or evaluating individual officers’ proficiency in specific subject areas.¹¹⁰ The result has been an ever expanding array of topics that officers have been allowed to provide expert testimony on, without having to provide any basis that their testimony is based on reliable principles or methods.¹¹¹ Courts have admitted officers as experts in the *modus operandi* of drug dealers, drug paraphernalia, drug jargon,¹¹² the mental state of couriers,¹¹³ the meaning and significance of rap lyrics,¹¹⁴ and gang affiliation,¹¹⁵ to name just a small subset.

The use of officer expertise to obscure forensics like call analysis is one evolution of the longstanding trend of prosecutors claiming the banner of officer training and experience to convince courts to admit officer opinion testimony. Beginning in the middle of the twentieth century, prosecutors increasingly introduced police officers to the courtroom as informational authorities, and courts “commonly certified such witnesses, welcoming policemen in the guise of ‘experts’ at trial.”¹¹⁶ Over time, courts increasingly deferred to the police, even as the credentials of those officers depreciated.¹¹⁷ The result is that officers who merely attend a few hours of training can be treated as

¹⁰⁹ *Id.* at 536 (quoting Appellants’ Brief at *37, *Pinter v. City of New York*, 448 F. App’x 99 (2d Cir. 2011) (No. 10-3789-cv).

¹¹⁰ Lvovsky, *supra* note 46, at 2024 (finding that judges credited officers as experts with ever lower showings of training, experience, and insight).

¹¹¹ See Brian R. Gallini, *To Serve and Protect? Officers as Expert Witnesses in Federal Drug Prosecutions*, 19 GEO. MASON. L. REV. 363, 364 (2012).

¹¹² See *id.* at 390.

¹¹³ *Diaz v. United States*, 144 S. Ct. 1727, 1730 (2024).

¹¹⁴ Andrea Dennis, *Poetic (In)Justice? Rap Music Lyrics as Art, Life, and Criminal Evidence*, 31 COLUM. J. L. & ARTS 1, 38 (2007).

¹¹⁵ Christopher McGinnis & Sarah Eisenhart, *Interrogation Is Not Ethnography: The Irrational Admission of Gang Cops as Experts in the Field of Sociology*, 7 HASTINGS RACE & POVERTY L.J. 111, 127, 132 (2010).

¹¹⁶ Anna Lvovsky, *supra* note 46, at 2062.

¹¹⁷ *Id.* at 2024.

“experts” and allowed to “step[] into positions of professional authority in the courtroom.”¹¹⁸

In *People v. Spitler*, 16-year-old Riley Spitler was charged in the accidental shooting death of his brother.¹¹⁹ Spitler always claimed it was an accident and never disputed that he shot his brother. Prosecutors argued that it was intentional murder and introduced the testimony of a detective who they said was an expert in 911 call analysis and linguistic statement analysis.¹²⁰ Over the defendant’s objection, the trial judge qualified the detective as an expert without once referencing *Daubert*.¹²¹ The prosecution provided no studies to demonstrate the reliability of the technique and cited no precedent that would support its admission. Still, the court admitted the practice absent any inquiry into either 911 call analysis or the detective’s proficiency because “it’s recognized by MCLE [The Michigan Commission of Law Enforcement Standards] . . . which is a specific law enforcement agency that’s involved in uniform training of law enforcement officials in Michigan and across the country.”¹²² The fact that the MCLE, the Michigan Commission of Law Enforcement Standards, is not a scientific board that certifies or accredits courses, but merely an agency that approves funding for trainings, did not bother the judge.¹²³ As to the detective’s proficiency the judge concluded “clearly he’s had this kind of—training analysis. He uses it on a regular basis.”¹²⁴ The court held that the detective’s expertise in 911 call analysis and linguistic statement analysis would “assist the trier of fact to understand the evidence or to determine a fact.”¹²⁵

The detective testified that he had special training in 911 call analysis and that the technique “[i]s imperative in an investigation . . . because there is pieces [sic] and parts of the

¹¹⁸ See *id.*

¹¹⁹ *People v. Spitler*, No. 331962, 2017 WL 2664729, at *1 (Mich. Ct. App. June 20, 2017).

¹²⁰ Jury Trial Transcript, vol. III, at 128, *People v. Spitler*, No. 2014-5198-FJ (Mich. Cir. Ct. Jackson Cnty. June 7, 2016).

¹²¹ *Id.* The court erroneously mentioned *Frye* despite Michigan having adopted *Daubert* in 2004 and further erred by shifting the burden to the defense to provide studies that disproved the reliability of 911 call analysis and linguistic analysis without asking the prosecution to produce any evidence of reliability. See *id.*; Motion for New Trial Transcript at 17, *People v. Spitler*, No. 2014-5198-FJ (Mich. Cir. Ct. Jackson Cnty. Sep. 27, 2016); *Spitler*, 2017 WL 2664729, at *4.

¹²² Jury Trial Transcript, vol. III, *supra* note 120, at 136–37.

¹²³ Murphy, *supra* note 8.

¹²⁴ Jury Trial Transcript, vol. III, *supra* note 120, at 136.

¹²⁵ *Id.* at 142–43.

call that you actually are analyzing for to determine whether an innocent, honest person has made that call or whether a guilty or deceptive individual has made that call.”¹²⁶ Using his “expertise,” the detective concluded that Spitler’s 911 call “was not about . . . help for the victim.”¹²⁷ He was able to identify Spitler’s deceit because Spitler “didn’t request for assistance for a victim” but instead said things like “help me” and “hurry.”¹²⁸ The detective also told jurors that Spitler had accepted his brother’s death and was not actually seeking help for him, as demonstrated by statements like “I think I killed my brother. Please help me. My life is over.”¹²⁹ Although the officers who responded to the scene reported that Spitler was extremely distraught, the detective explained that his opinion should carry more weight because the officers “are not trained and do not have the experience I have.”¹³⁰

The jury convicted Spitler of second-degree murder, and the judge sentenced him to 20-40 years in prison.¹³¹ Post-conviction, the judge denied a motion for a new trial and a *Daubert* hearing. In again denying a hearing, the court used particularly strained reasoning, likening 911 call analysis to DNA testing: “[M]odern techniques of criminology are evolving all the time. I mean, there was a point when DNA evidence hadn’t been accepted in courtrooms.”¹³² While the court recognized the evolving standards in science, it overlooked the fact that DNA technology is the most rigorously tested forensic technique and one that has repeatedly been subject to evidentiary challenge since its inception in criminal cases more than thirty years ago.¹³³ In contrast, there are no studies validating the claims of 911 call analysis.¹³⁴ The court then erroneously shifted the burden to the defense by stating “I don’t think at this point [the defense] set forth the requisite showings to have

¹²⁶ *Id.* at 119–20.

¹²⁷ Jury Trial Transcript, vol. IV, at 51, *People v. Spitler*, No. 2014-5198-FJ (Mich. Cir. Ct. Jackson Cnty. June 7, 2016).

¹²⁸ *Id.* at 52.

¹²⁹ *Id.*

¹³⁰ *Id.* at 62; Jury Trial Transcript, vol. III, *supra* note 120, at 6.

¹³¹ Murphy, *supra* note 8.

¹³² Motion for New Trial Transcript at 16, *People v. Spitler*, No. 2014-5198-FJ (Mich. Cir. Ct. Jackson Cnty. Sep. 16, 2016).

¹³³ NAT’L RESEARCH COUNCIL, *supra* note 39, at 87).

¹³⁴ *See supra* Part I.

a full blown *Daubert* hearing on the matter”¹³⁵ and ultimately abdicated its role as evidentiary gatekeeper by stating that it would “defer to what the appellate courts ultimately determine on the matter.”¹³⁶

The Michigan Court of Appeals determined that “the trial court misplaced the responsibility for ensuring that MRE 702,” Michigan’s analogue to Federal Rule of Evidence 702, “is satisfied.”¹³⁷ The decision clarified that “[i]t is the trial court’s obligation to conduct a searching inquiry into the reliability of scientific techniques under MRE 702.”¹³⁸ In overturning Spitler’s murder conviction, the court found that a *Daubert* hearing should have been held, as there was nothing in the case law, or provided by the prosecution, to support the qualification of 911 call analysis or linguistic statement analysis as reliable or accepted techniques.¹³⁹ On remand, Spitler was re-sentenced to manslaughter¹⁴⁰ and released after four years of incarceration.¹⁴¹

Spitler illustrates the dangers of trial courts failing to interrogate the claims of police expertise. The trial court’s decision to admit (and endorse) 911 call analysis testimony absent any hearing exemplifies the authority-based deference that Lvovsky decries.¹⁴² When courts defer to police simply because of their authority, they do so “on principle, or perhaps on faith, making no claims to increase the likelihood of improving the law’s application.”¹⁴³ The result is that courts reify police biases that lead to “excess suspicion, overzealousness . . . and pervasive racial prejudice.”¹⁴⁴ In *Spitler*, the prosecution produced no validation studies and cited no precedent for the admission of 911 call analysis. Instead, prosecutors urged the court to simply trust the expertise of the detective instead of examining his testimony. The trial court abnegated its responsibility, deferred to the police, and Spitler may have spent decades more in prison had the appellate court not properly intervened.

¹³⁵ Motion for New Trial Transcript, *supra* note 132, at 16–17.).

¹³⁶ *Id.* at 17.

¹³⁷ *People v. Spitler*, No. 331962, 2017 WL 2664729, at *4 (Mich. Ct. App. June 20, 2017).

¹³⁸ *Id.*

¹³⁹ *Id.*

¹⁴⁰ *Id.* at 5.

¹⁴¹ *Murphy*, *supra* note 8.

¹⁴² *Anna Lvovsky*, *supra* note 108, at 558.

¹⁴³ *Id.* at 556.

¹⁴⁴ *Lvovsky*, *supra* note 46, at 2069 (footnote omitted).

D. Shadow Experts

Prosecutors also circumvent reliability analysis in cases by introducing tenets of call analysis without any proffer or even mention of the technique. They do so through the use of “shadow experts”: lay witnesses who use the forensic language and concepts of a technique without any mention of reliability or claim to witness expertise. The practice avoids examination at various levels. The use of shadow experts shields the technique from challenge at pretrial reliability hearings and prevents defendants from confronting the technique at trial.¹⁴⁵ Further, with no record of the technique, ex-post examination also becomes difficult.¹⁴⁶ The prosecutors in *State v. Baldwin* circumvented the evidentiary process by introducing 911 call analysis testimony without providing any notice and without proffering their witness, a police detective, as an expert.¹⁴⁷ Jennifer Baldwin serving a twenty-one-year prison term after being convicted of murder in the death of a nine-month-old whom she was babysitting.¹⁴⁸ Although the State’s case was entirely circumstantial, Ms. Baldwin was convicted on the basis of medical testimony that the child could only have died from abuse and from the detective’s testimony that she was guilty.¹⁴⁹ The detective told the jury that Ms. Baldwin was guilty because of the words she used and the order she reported events to the 911 dispatcher.¹⁵⁰ He cited his professional training and presented 911 call analysis to the jury as known scientific fact:

I’ve attended training and I certainly have a great deal of experience interpreting 911 calls, and it is known that when a caller is concerned for somebody’s condition, first and foremost, they ask for help for them. Callers who contributed to

¹⁴⁵ Cf. *Crawford v. Washington*, 541 U.S. 36, 61 (2004) (“[T]he [Confrontation] Clause’s ultimate goal is to ensure reliability of evidence, but it is a procedural rather than a substantive guarantee. It commands, not that evidence be reliable, but that reliability be assessed in a particular manner: by testing in the crucible of cross-examination.”).

¹⁴⁶ See DANIEL S. MEDWED, *BARRED: WHY THE INNOCENT CAN’T GET OUT OF PRISON* 35–36 (2022) (describing the limits of direct appeal remedies because they are limited to record-based errors).

¹⁴⁷ See Transcript of Proceedings, *supra* note 49, at 1508–13.

¹⁴⁸ *Id.* at 1678, 1695–98.

¹⁴⁹ *Id.* at 1508–1509, 1621.

¹⁵⁰ *Id.* at 1508–1509. The child, who had a history of prior illness, became unresponsive while in Baldwin’s care. Baldwin called 911 within minutes of noticing. There was no documented abuse, no eyewitness that testified to abuse, and no confession from Baldwin. See generally *id.*

that poor condition offer explanations which may draw attention elsewhere, distracters, things that may explain what happened before they actually ask for any help or assistance for the injured person.¹⁵¹

He told jurors that three factors demonstrated Ms. Baldwin's guilt: first, that she did not immediately ask for an ambulance,¹⁵² second, that she provided "extraneous information" to the operator, and third, that she called the child's mother before calling 911.¹⁵³

The prosecution in *Baldwin* explicitly hid its use of call analysis by engaging in a form of forensic "parallel construction".¹⁵⁴ As Professors Toomey and Kaufman explain, parallel construction entails "laundering evidence in order to avoid court review of its actual provenance—in other words, making evidence look like it came from source 'B' when, in fact, it originally came from source 'A.'"¹⁵⁵ For example, agents of the Special Operations Division of the Drug Enforcement Agency were officially encouraged to use this tactic in order to hide information obtained via National Security Agency (NSA) tips.¹⁵⁶ Human Rights Watch has documented the myriad ways that government agents use parallel construction as a means to conceal illegal or covert surveillance.¹⁵⁷ In *Baldwin*, the detective similarly led the jurors to believe that his general training and experience—and not his specific training in 911 call analysis—led him to conclude that the defendant was guilty.

During closing arguments, the State repeatedly emphasized the junk forensic claims about Ms. Baldwin's 911 call.¹⁵⁸ Prosecutors played her 911 call for the jury and commented on

¹⁵¹ *Id.* at 1509.

¹⁵² This testimony ignores the fact that her first words to 911 were "Yes, can I have an ambulance here, please." 911 Transcript, Ex. 29A, State v. Baldwin, No. 27-CR-18-23849 (Minn. Dist. Ct. Hennepin Cnty. Aug. 28, 2019).

¹⁵³ See Transcript of Proceedings, *supra* note 147, at 1508–09; Memorandum in Support of Second Petition for Post-Conviction Relief at 45–46, State v. Baldwin, No. 27-CR-18-23849 (Minn. Dist. Ct. Hennepin Cnty. Oct. 12, 2023) (describing the officer's testimony in Exhibit 9).

¹⁵⁴ Toomey & Kaufman, *supra* note 53, at 863.

¹⁵⁵ *Id.*

¹⁵⁶ *Id.* (citing John Shiffman & Kristina Cooke, *Exclusive: U.S. Directs Agents to Cover up Program Used to Investigate Americans*, REUTERS (Aug. 5, 2013), <https://www.reuters.com/article/2013/08/05/us-dea-sod-idUSBRE97409R20130805/> [<https://perma.cc/N8A8-LPT5>]).

¹⁵⁷ HUM. RTS. WATCH, DARK SIDE: ORIGINS OF EVIDENCE IN US CRIMINAL CASES 12 (2018).

¹⁵⁸ See Transcript of Proceedings, *supra* note 147, at 1606–09.

the alleged indicators of her guilt no less than a dozen times during closing arguments.¹⁵⁹ Jurors convicted Ms. Baldwin after less than one day of deliberation.¹⁶⁰

The prosecution in *Baldwin* never proffered the detective as an expert; it never provided a CV and neither mentioned, nor documented, the call analysis trainings the detective attended. Consequently, counsel never objected to the testimony, and no *Frye* hearing was ever held. Neither did the defense know to put on its own expert witness to question the scientific validity of call analysis. And with no knowledge of its use, defense counsel could not properly confront the bases of the detective's opinion.¹⁶¹ Ms. Baldwin only learned that the detective had attended a call analysis training when her post-conviction counsel sent Freedom of Information (FOIA) Requests for the detective's personnel records.¹⁶² In ruling on Baldwin's post-conviction motion, the District Court agreed that call analysis "has been scientifically discredited" and that the Detective gave false testimony, but it denied the motion, because the practice was not explicitly referenced during trial.¹⁶³ The court noted that "the State was not attempting to elicit scientific 911 call analysis from [the Detective] as expert testimony" and that "the 911 call analysis training . . . was not mentioned or described to the jury."¹⁶⁴ Even after the practice was made visible to the court, the State's use of a shadow expert protected its conviction. The Minnesota Court of Appeals denied Ms. Baldwin's appeal based on the false 911 analysis testimony but reversed the district court's second denial of her petition for postconviction relief on other grounds.¹⁶⁵

Shadow forensics exemplifies the whack-a-mole nature of ensuring forensic reliability. As courts, scientific organizations, and the Federal Rules of Evidence Advisory Committee

¹⁵⁹ *Id.*

¹⁶⁰ *Id.* at 1670; Transcript of Proceedings, *supra* note 147, at 1678.

¹⁶¹ See Meghan J. Ryan, *Secret Conviction Programs*, 77 WASH. & LEE L. REV. 269, 274 (2020) ("Without the necessary information about the evidence that is being used to convict him, a defendant is denied the opportunity to properly challenge this evidence.").

¹⁶² See Memorandum in Support of Second Petition for Post-Conviction Relief, *supra* note 153, at 45-46.

¹⁶³ Order Denying Motion for Post-Conviction Relief at 23, *State v. Baldwin*, No. 27-CR-18-23849 (Minn. Dist. Ct. Hennepin Cnty. Feb. 28, 2025).

¹⁶⁴ *Id.* at 23-24.

¹⁶⁵ *Baldwin v. State*, No. A25-0555, 2026 WL 264721, at *1, *11-12 (Minn. Ct. App. Feb. 2, 2026).

seek to strengthen the reliability requirements of forensic evidence, prosecutors and police find ways to introduce forensic techniques that lack reliability, but help secure convictions.¹⁶⁶ While this dynamic has played out since the introduction of forensics,¹⁶⁷ the infiltration of call analysis in the criminal legal process is one of the clearest manifestations of how evidentiary rules are subverted to gain unfair advantage.

In emails, Harpster and prosecutors acknowledge that call analysis does not pass scientific scrutiny and would rightfully be barred if subject to a *Daubert* or *Frye* hearing.¹⁶⁸ However, because the testimony is compelling to jurors, prosecutors hide the practice so that no evidentiary record can be made, no judge can rule against its admission, and therefore no precedent can be set. Harpster coaches prosecutors to shepherd the principles of call analysis into trials without proffering any expert. Instead, officers or call operators are taught to discuss the principles of call analysis under the guise of their “training and experience.”¹⁶⁹

These testifying officers act as shadow experts that bundle unreliable forensic testimony into a trial. Although they do not explicitly reference 911 call analysis, the officer will use the phrases and concepts of call analysis to show jurors that a defendant acted in ways consistent with the COPS guilty indicators and inconsistent with the innocent indicators. Prosecutors cement that impression by highlighting the officers’ training and experience and reinforcing notions that the concept of call analysis is a commonly held fact. By focusing

¹⁶⁶ See Murphy, *supra* note 50, at 745 (describing the “well-documented and long” history of forensic scandals including “tales of fabrication and improper handling, falsification of results and reports, rogue or incompetent analyses, and corrupt or misleading testimony.”).

¹⁶⁷ Professor Binyamin Blum documents the British Empire’s use of untested forensic practices, like forensic dog tracking, as early as 1888. The Empire tested forensic practices against individuals—often non-white, non-British individuals—in the colonies, where their use was met with fewer qualms and less resistance, before introducing those same practices in England. See Binyamin Blum, *The Hounds of Empire: Forensic Dog Tracking in Britain and its Colonies, 1888-1953*, 35 L. & HIST. REV. 621, 622, 632 (2017). Similarly, Professor Mitra Sharafi explores the use of precipitin testing to identify blood, which proliferated in the British-ruled India but was rarely used in the British metropole, a fact which reflected British anxiety about the truthfulness of Indian witnesses. See Mitra Sharafi, *The Imperial Serologist and Punitive Self-Harm: Bloodstains and Legal Pluralism in British India*, in GLOBAL FORENSIC CULTURES 60–61 (Christopher Hamlin & Ian Burney, eds., 2019).

¹⁶⁸ See Murphy, *supra* note 8.

¹⁶⁹ See *id.*

jurors on officers' training and experience, prosecutors get the benefit of perceived expertise without any of the burdens of having to establish reliability through a *Daubert* or *Frye* hearing.

Even more dangerous in the taxonomy of shadow forensics are cases where prosecutors seek to evade mention of 911 call analysis at all, especially in cases like *People v. Logan*, where the technique is central to the investigation and prosecution.¹⁷⁰ Like in *Baldwin*, prosecutors in Illinois evaded judicial scrutiny of 911 call analysis by choosing not to proffer a detective as an expert in the practice.¹⁷¹ However, the Illinois prosecutors went a step further: Not only was the detective not proffered, the prosecutors avoided asking him *any* questions about the practice, even where his conclusion about the defendant's 911 call significantly contributed to her arrest and conviction. The lack of explicit mention of 911 call analysis on direct examination belied the fact that it played a role in every aspect of convicting Jessica Logan for the death of her 19-month-old son.¹⁷²

When Logan, a twenty-two-year-old mother discovered that her son was not breathing, she called her adoptive mother and then called 911 immediately to ask for help.¹⁷³ She sobbed, "I need my baby," and told the operator "I came in my son's room to try to give him a breathing treatment because he needs breathing treatments," and "he's not breathing."¹⁷⁴ According to a local detective who had taken a two-day call analysis training, Logan did not ask for help soon enough into the call and did not phrase her request for help the way an innocent caller should.¹⁷⁵ Using the COPS checklist, he concluded that Logan exhibited fourteen guilty indicators but only two innocent indicators. He noted that she did not immediately tell the operator what was wrong, did not explicitly ask for an ambulance, and attempted to convince the operator that her son had a history of breathing problems.¹⁷⁶ That her son did, in fact, have a

¹⁷⁰ See Murphy, *supra* note 103; Partial Report of Proceedings, *People v. Logan*, No. 2019-CF-1648 (Ill. Cir. Ct. Macon Cnty. Jan. 5, 2024) [hereinafter Partial Report: Closings].

¹⁷¹ See Partial Report: Closings, *supra* note 170, at 3–18.

¹⁷² See Murphy, *supra* note 103.

¹⁷³ Partial Report: People's Witnesses, *supra* note 49.

¹⁷⁴ Murphy, *supra* note 103.

¹⁷⁵ *Id.*

¹⁷⁶ *Id.*

history of hospitalizations for respiratory issues was of no moment for the detective.

The detective's conclusion that Logan was guilty made her a suspect and infected every subsequent part of the case against her. The detective brought his conclusion to the initial responding officer. That officer, who had first documented that Logan was "upset and crying" subsequently changed his opinion in a second report that indicated that he thought her emotions were insincere because "her crying somehow did not seem genuine and felt forced."¹⁷⁷ The detective shared his conclusion at the coroner's inquest. There, he "told the jurors he was trained in basic and advanced 911 call analysis and said he found 'mostly all guilty indicators' in Logan's recording."¹⁷⁸ The detective also conferred with the state's forensic pathologist who suggested that the detective have Logan do a reenactment of the night her son died. After the un-*Mirandized* reenactment, the detective noted that Logan "did not appear to be upset or troubled when manipulating the mannequin and illustrating how she discovered her deceased son in his bed." After the detective's subsequent report, the pathologist concluded that the baby suffocated by smothering.¹⁷⁹

At trial, as in *Baldwin*, the prosecutors in Logan's case did not proffer the detective as an expert in 911 call analysis. Emails obtained by ProPublica indicate that prosecutors purposely avoided proffering him to avoid any *Frye* scrutiny of the practice. Notably, prosecutors were so wary of bringing 911 call analysis to light that they did not ask the detective a single question about the practice. He instead testified as a shadow expert; despite the central role of the detective's conclusion in the investigation, not once did prosecutors ask the detective to discuss the technique, his training, or its implications. Instead, prosecutors incorporated his conclusions into their closing argument. They played Logan's 911 call and argued that her call evinced her guilt.

¹⁷⁷ *Id.*

¹⁷⁸ *Id.*

¹⁷⁹ *Id.* Subsequent pathologists asked to review the case disagreed with the homicide diagnosis. For a general critique of manner of death testimony, see Keith A. Findley & Dean A. Strang, *Ending Manner of Death Testimony and Other Opinion Determinations of Crime*, 60 DUQ. L. REV. 302 (2022).

Jurors convicted Ms. Logan of first-degree murder after deliberating for less than two hours.¹⁸⁰ After the guilty verdict, the detective emailed Harpster to thank him for his help: “[T]he 911 call analysis was a tremendous benefit to both my investigation and the trial.”¹⁸¹

In *State v. Faria*, prosecutors similarly used 911 call analysis to convict Russell Faria of murdering his wife, but without proffering the technique or qualifying any witness as an expert. Prosecutors recognized that the practice “is not ‘recognized’ as a science in our state” and in their own words, purposely disguised the technique by “getting creative . . . without calling it science.”¹⁸² The prosecutor instead used a 911 call operator as a shadow expert. The call operator opined about the defendant’s 911 call using tenets of call analysis, but without any proffer of the technique.¹⁸³ Over defense counsel’s objection, the operator observed that someone who was genuinely hysterical after finding their wife murdered would not be able to answer questions as clearly as Mr. Faria had.¹⁸⁴ The operator implied, based on her training and experience, that Mr. Faria was being deceptive.¹⁸⁵ In closing arguments, prosecutors reinforced that idea and argued that Faria was actively trying to deceive the 911 call operator with his answers.¹⁸⁶ Faria was convicted of murder and sentenced to life in prison.¹⁸⁷

After Faria appealed his conviction and won a new trial,¹⁸⁸ the prosecution again attempted to convict him with the help of 911 call analysis.¹⁸⁹ The same prosecutor tried to elicit tes-

¹⁸⁰ Murphy, *supra* note 103. On appeal to the Illinois Supreme Court, Logan argued that her statements in the reenactment were taken in violation of *Miranda*. Although the Illinois Supreme Court agreed that the statements violated *Miranda*, it concluded that it did not amount to plain error when balanced against evidence of Logan’s guilt, which included evidence of her economic precarity, the fact that she cashed in her son’s life insurance policy, and an internet search for the phrase “how do you suffocate.” *People v. Logan*, 250 N.E. 3d 548, 569–70 (Ill. 2024).

¹⁸¹ Murphy, *supra* note 103.

¹⁸² Murphy, *supra* note 8.

¹⁸³ *Id.* at 396–403.

¹⁸⁴ Transcript on Appeal at 403, *State v. Faria*, No. ED 100694 (Mo. Ct. App. E.D. Jan. 5, 2024).

¹⁸⁵ *Id.*

¹⁸⁶ *Id.* at 1108.

¹⁸⁷ *Id.* at 1154; Murphy, *supra* note 8.

¹⁸⁸ Order, *State v. Faria*, No. ED 100694 (Mo. Ct. App. E.D. Feb. 24, 2015) (ordering new trial based on newly discovered evidence of alternate suspect).

¹⁸⁹ Murphy, *supra* note 8.

timony about the operator's conclusions based on her call analysis training. Faria's counsel at his second trial, now aware of the call analysis testimony, moved in *limine* to exclude any expert testimony about the 911 call. The court limited the operator's testimony. Although the jury heard the call itself and about the operator's attendance at the call analysis training, the operator was not allowed to provide her impressions of the call. Without the shadow expert testimony, Faria was acquitted.

While this Article focuses on the use of shadow forensics to subvert evidentiary rules in order to bring call analysis into criminal trials, it recognizes that the harms of call analysis enter a case far earlier than at the trial stage, assuming the case even gets that far.¹⁹⁰ Like many forensic practices, call analysis may be used at the inception of a case to focus on a suspect, or as in Kathy Carpenter's case, to provide probable cause for arrest.¹⁹¹ Because Rule 702 and *Daubert* do not govern the admissibility of evidence at pretrial hearings or at the investigative stage, and because over 90% of criminal cases do not go to trial, the use of call analysis may go undetected, and therefore unscrutinized, in the vast majority of cases where it is deployed.¹⁹² Because the practice has remained so hidden, few defense attorneys have been aware enough to even raise the issue at pretrial suppression hearings. The result, like with other forensic practices, is that the technique does not get scrutinized, and many people may never learn that it was used against them.

¹⁹⁰ As widely noted, the vast majority of criminal cases resolve via guilty plea. By most accounts, 97% of federal criminal cases and 94% of state court criminal cases resolve via guilty plea. *Missouri v. Frye*, 566 U.S. 134, 135 (2011); see also Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2463, 2466 (2004).

¹⁹¹ See *Transcript Used to Charge Carpenter*, *supra* note 4. At the onset of a case, call analysis, like other junk forensics, gives credence to uninformed hunches that can trigger confirmation bias throughout every step of a police investigation and ultimately lead to wrongful convictions. See Samantha Bean et al., *From Caller to Suspect: Identifying Behaviors That Trigger Suspicion in 911 Calls*, 50 L. & HUM. BEHAV. 182, 193, 195 (2025). See generally Kyle C. Scherr, Allison D. Redlich, & Saul Kassin, *Cumulative Disadvantage: A Psychological Framework for Understanding How Innocence Can Lead to Confession, Wrongful Conviction, and Beyond*, 15 PERSPS. PSYCHOL. SCI. 353 (2020).

¹⁹² See *Sinha*, *supra* note 95, at 103; *Frye*, 566 U.S. at 135.

II

THE UNRELIABILITY OF 911 CALL ANALYSIS

Much has been written about the rise and spread of forensic practices like handwriting analysis,¹⁹³ arson investigation,¹⁹⁴ bitemarks,¹⁹⁵ hair microscopy, and toolmark comparisons.¹⁹⁶ These practices have been long discredited as scientifically invalid and unreliable,¹⁹⁷ and therefore pose a threat to the integrity of the court system. To ensure the reliability of forensic techniques, a rash of safeguards were erected by the judiciary, the scientific community, and Congress. Starting with the Supreme Court's decision in *Daubert v. Merrell Dow* and the adoption of *Daubert's* reliability requirement by the Federal Rules of Evidence, and continuing with the publication of reports by the President's Council of Advisors on Science and Technology (PCAST) and the National Academy of Sciences (NAS), a growing number of voices have sought to limit the admission of unreliable junk forensics in court.¹⁹⁸ Despite these efforts, junk forensics have repeatedly been introduced into criminal trials and have played a pivotal role in wrongful convictions.¹⁹⁹ While others have provided rich histories of the ways that junk

¹⁹³ See Mnookin, *supra* note 28.

¹⁹⁴ See Beety & Oliva, *supra* note 25, at 483.

¹⁹⁵ See generally M. CHRIS FABRICANT, JUNK SCIENCE AND THE AMERICAN CRIMINAL JUSTICE SYSTEM (2022); Fabricant & Carrington, *supra* note 16.

¹⁹⁶ See Maneka Sinha, *Junk Science at Sentencing*, 89 GEO. WASH. L. REV. 52, 60–62 (2021).

¹⁹⁷ See NAT'L RSCH. COUNCIL, *supra* note 39, at 8 (identifying the “notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods”). A follow up study by the President's Council of Advisors on Science and Technology (PCAST) came to similar conclusions regarding “feature-comparison” techniques. See PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., EXEC. OFFICE OF THE PRESIDENT, FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON MODELS 122 (2016).

¹⁹⁸ See NAT'L RSCH. COUNCIL, *supra* note 39, at 8; PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., EXEC. OFFICE OF THE PRESIDENT, FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON MODELS 122 (2016).

¹⁹⁹ According to the Innocence Project, misapplied forensic techniques contributed to nearly a quarter of all wrongful conviction cases since 1989. INNOCENCE PROJECT, MISAPPLICATION OF FORENSIC SCIENCE, <https://innocenceproject.org/misapplication-of-forensic-science/> [<https://perma.cc/6ETE-B83H>]; see also John Morgan, *Wrongful Convictions and Claims of False or Misleading Forensic Evidence*, 68 J. FORENSIC SCIS. 908, 908 (2023); McCrory v. Alabama, 144 S. Ct. 2483, 2483 (2024) (Sotomayor, J., statement on denial of certiorari) (“Hundreds if not thousands of innocent people may currently be incarcerated despite a modern consensus that the central piece of evidence at their trials lacked any scientific basis.”).

forensics are admitted and accepted by courts,²⁰⁰ this Part uses 911 call analysis as a case study to provide an overview of *Frye* and *Daubert* and to show how call analysis should be wholly inadmissible in court. It proceeds by (1) describing the claims that call analysis makes; (2) situating those claims within the standards set forth first by *Frye* and then *Daubert*, and (3) concluding that call analysis should fail under any reliability analysis.

A. The Claims of “COPS”

The founders of 911 call analysis purport to have created a system that can identify the perpetrators of crime via their 911 calls. Police departments and prosecutors have used the evidence to identify and arrest suspects—sometimes when the 911 call audio is the only identifying evidence in a case.²⁰¹ The fundamental assumption of the technique is that 911 calls provide “insightful, uncontaminated verbal and vocal cues” and that trained examiners can differentiate indicators of innocence from indicators of guilt.²⁰² The technique claims that to identify whether a 911 caller “is the killer,” one can broadly examine six dimensions: (1) what the call was about; (2) who the call was about; (3) how the call was made; (4) aggressive demands/passive defenses; (5) cooperation/resistance with the call dispatcher; and (6) additional guilty indicators.²⁰³

The technique uses what its founders named the Considering Offender Probability Statements (COPS) scale. The COPS scale is a checklist of thirty-eight “guilty indicators” and fifteen “innocent indicators” meant to identify deception in 911 callers.²⁰⁴ Examiners listen to 911 calls for indicators like: urgency (a caller who does not express urgency is guilty); the immediacy of the plea for help (a non-immediate plea indicates guilt); extraneous information (indicating guilt); aggressive demands (indicating innocence); and cooperation with the dispatcher (indicating innocence).²⁰⁵

²⁰⁰ See *supra* notes 189–191.

²⁰¹ See *Mistaken 911 Transcript Used to Charge Carpenter*, *supra* note 4.

²⁰² See Adams & Harpster, *supra* note 9, at 22–23. According to Adams and Harpster, a 911 call would be less “contaminated” by “suspects’ attempts to conceal the truth, attorneys’ advice to remain silent, and investigators’ leading questions.” *Id.*

²⁰³ See *id.* at 27.

²⁰⁴ See HARPSTER & ADAMS, *supra* note 10, at 260.

²⁰⁵ *Id.*

911 COPS Scale®

Considering Offender Probability in Statements

Place a check mark for any indicator that applies or circle n/a for not applicable

<u>Innocent Indicators</u>	WHO is the call about?	<u>Guilty Indicators</u>
Immediate Plea for Help for Victim ☐	n/a ☐	No Immediate Plea for Help for Victim ☐
Immediate Assessment of Victim ☐	n/a ☐	No Immediate Assessment of Victim ☐
Focus on Victim ☐	n/a ☐	Focus on Caller ☐
No Acceptance of Death ☐	n/a ☐	Acceptance of Victim's Death ☐
Aid Provided to Victim ☐	n/a ☐	No Aid Provided to Victim ☐
WHAT is the call about?		
Relevant Information ☐	n/a ☐	Extraneous Information ☐
Sensory Details ☐	n/a ☐	Lack of Sensory Details ☐
Prioritized Order ☐	n/a ☐	Inappropriate Order ☐
Bleeding Comments ☐	n/a ☐	Blood/Brains Comments ☐
HOW is the call made?		
Urgency ☐	n/a ☐	No Urgency ☐
Fear for Caller's Safety ☐	n/a ☐	No Fear for Caller's Safety ☐
Proximity to Victim ☐	n/a ☐	No Proximity to Victim ☐
Initial Sounds or Comments ☐	n/a ☐	Initial Delays ☐
Addressing/Coaching the Victim ☐	n/a ☐	Insults or Blames Victim ☐
Aggressive Demands ☐	n/a ☐	Passive Defenses: ☐ Defendant Mentality ☐ Ingratiating Remarks ☐ Mental Miscues ☐ Minimizes
Cooperation with Dispatcher ☐	n/a ☐	Resistance to Dispatcher: ☐ Dispatcher Confusion ☐ Diversion ☐ Equivocation ☐ Evasion ☐ Hangs Up ☐ Only Answers What's Asked ☐ Pauses ☐ Repetition ☐ Self-Interruption ☐ Short Answers ☐ Unintelligible Comments
Additional Guilty Indicators:		
☐ Attempts to Convince	☐ Awkward Phrases	☐ Conflicting Facts
☐ Isolated Plea	☐ Lack of Contractions	☐ No Modulation
		☐ "Huh?" Factor
		☐ "I Don't Know"
		☐ Unexplained Knowledge

Figure 2. 911 COPS Scale Checklist

In Kathy Carpenter's case, for example, a Colorado Bureau of Investigation Agent referenced the COPS checklist and concluded that Carpenter's call "revealed 39 indicators of guilty knowledge and no indicators of innocence."²⁰⁶ According to the agent, Carpenter demonstrated her guilt by, among other things: using repetitive phrases, imploring the operator to help her (instead of her friend), and failing to answer questions about what happened without stammering.²⁰⁷ Other guilty indicators, according to the agent, included accepting that her friend was dead and being confused when asked if there was a defibrillator around.²⁰⁸ Colorado prosecutors used the report to identify

²⁰⁶ LEWIS, *supra* note 3, at 1.

²⁰⁷ *Id.* at 1-2.

²⁰⁸ *Id.* at 2-3.

Carpenter as the main suspect in the murder and attached the report to obtain Carpenter's arrest warrant.²⁰⁹

Foundational to the theory of 911 call analysis is the assumption that particular binary factors are representative of guilt or of innocence.²¹⁰ There is no empirical evidence that this is the case.²¹¹ Notwithstanding its lack of scientific validity, call analysis remains tantalizing to police and prosecutors because of its persuasive power. The practice is particularly persuasive to factfinders because of its demonstrative nature. Like early handwriting experts, witnesses use call analysis to point out aspects for jurors to see and hear for themselves.²¹² In trials, police witnesses identify all the "guilty indicators" in a suspect's 911 call for jurors.²¹³ In this way, they render the "guilty indicators" visible to factfinders; factors like whether a someone called another person before calling 911 or whether they provided "extraneous" medical information about the victim cannot be unheard after jurors are made aware of them.²¹⁴ But call analysis suffers from the same infirmities that plague other lie detection techniques like the polygraph.

Broadly, lie detection techniques purport to allow people to identify deception from people's verbal²¹⁵ and nonverbal behavior.²¹⁶ These techniques claim to be able to subdivide a complex range of human cues—about demeanor, body

²⁰⁹ Abraham, *supra* note 2.

²¹⁰ For example, according to the scale, an immediate plea for help indicates innocence, but a non-immediate plea indicates guilt. Similarly, providing "relevant information" indicates innocence, but providing "extraneous information" indicates guilt. See Harpster, Adams & Jarvis, *supra* note 10, at 74–75, 78; HARPSTER & ADAMS, *supra* note 10, at 260.

²¹¹ See *infra* subpart II.B.

²¹² See Mnookin, *supra* note 28, at 1733.

²¹³ Calling another person before calling 911 and providing medical background information are both "guilty indicators" according to Harpster's analysis. See HARPSTER & ADAMS, *supra* note 10, at 51–53, 198–99.

²¹⁴ In a meta study of call analysis, Patrick Markey found that "the majority of findings . . . suggested that most of the COPS scale items were unrelated to caller deception." Patrick M. Markey et al., *Validity Concerns of the Considering Offender Probability Statements Scale in 911 Homicide Call Analysis: An Empirical Study*, 31 PSYCH. PUB. POL'Y & L. 22, 23 (2025).

²¹⁵ See Verena A. Oberlader et al., *Validity of Content-Based Techniques to Distinguish True and Fabricated Statements: A Meta-Analysis*, 40 L. & HUM. BEHAV. 440, 440–41 (2016).

²¹⁶ See Tim Brennen & Svein Magnussen, *Lie Detection: What Works?*, 32 CURRENT DIRECTIONS PSYCH. SCI. 395, 395–396 (2023); see also Julia Simon-Kerr, *Unmasking Demeanor*, 88 GEO. WASH. L. REV. ARGUENDO 158, 160–61 (2020), <https://www.gwlr.org/wp-content/uploads/2020/09/88-Geo.-Wash.-L.-Rev.-Arguendo-158.pdf> [<https://perma.cc/DB84-FBE5>].

language, gestures, and verbal reactions (or non-reactions)—into discrete indicators of deception or truth.²¹⁷ Call analysis is simply the latest attempt to streamline the complexity of human behavior into a prescriptive checklist that allows the police to divine truth.²¹⁸ The COPs checklist, like other lie detection techniques, ignores the fact that “psychological phenomena are intrinsically noisy”²¹⁹ and presents itself as a key to unlocking what are ostensibly psychic abilities.

Meta-analyses of polygraph tests and micro-expressions show that these techniques do not reliably detect deception.²²⁰ Studies demonstrate that people detect lies based on micro-expressions at scarcely better than chance levels.²²¹ In fact, because of its unreliability, polygraph evidence is effectively barred from admission in court.²²² Despite this, police departments still employ polygraph testing and other detection techniques, and law enforcement agencies still spend thousands of dollars sending officers to trainings where self-proclaimed experts teach them to distinguish lies from truth based on subconscious gestures and innate expertise.²²³ People like Steven Rhoads, an expert used by prosecutor’s offices across the country, teach officers, for example, that “the left side of

²¹⁷ Smith, *supra* note 108 (describing, for example, police officers at a “New Tools for Detecting Deception” training hosted by the Wisconsin High Intensity Drug Trafficking Area learned about “the seven universal facial expressions that all people have all over the world as a good indicator” of lying).

²¹⁸ Law enforcement has, and continues, to employ techniques that have been found to be unreliable, like the polygraph and microexpressions. *Id.*; see United States v. Scheffer, 523 U.S. 303, 313–14 (1998) (holding the polygraph can be made admissible); see also Stephen Porter, Leanne ten Brinke & Brendan Wallace, *Secrets and Lies: Involuntary Leakage in Deceptive Facial Expressions as a Function of Emotion Intensity*, 36 J. NONVERBAL BEHAV. 23, 32 (2012) (finding that individuals can only detect deception via micro-expression at “not significantly above chance” levels).

²¹⁹ Brennen & Magnussen, *supra* note 216, at 400.

²²⁰ See Friedo J.W. Herbig, *The “Psychology” of Polygraph: Endangering Differential Salience – Concerns and Caveats*, J. PSYCH. RSCH., April 2020, at 2; Leonard Saxe, Denis Dougherty & Theodore Cross, *The Validity of Polygraph Testing: Scientific Analysis and Public Controversy*, 40 AM. PSYCHOLOGIST 355 (1985).

²²¹ See Stephen Porter & Leanne ten Brinke, *Reading Between the Lies: Identifying Concealed and Falsified Emotions in Universal Facial Expressions*, 19 PSYCH. SCI. 508, 508 (2008); see also Sarah Jordan et. al., *A Test of Micro-Expressions Training Tool: Does it Improve Lie Detection?*, 16 J. INVESTIGATIVE PSYCH. & OFFENDER PROFILING 222, 222 (2019) (finding that individuals specifically trained to detect deception in micro-expressions detected deception at lower than chance rates).

²²² United States v. Scheffer, 523 U.S. 303, 314 (1998); see also Herbig, *supra* note 219, at 2 (confirming that the polygraph is an unreliable measure of deception).

²²³ See Smith, *supra* note 108 (noting that classes from Renee Ellory, the trainer at the “New Tools for Detecting Deception” cost \$1950 per person).

the body is more apt to reveal known deception than the right side of the body.”²²⁴ Compounding the problem is that while police are no more skilled than civilians at detecting deception,²²⁵ they are far more likely to be overconfident in their assessments.²²⁶ Research demonstrates that officers who take even a single training significantly overrate their ability to detect lies.²²⁷ As is the case with police trainings in any other lie detection technique, 911 call analysis trainings conducted across the country²²⁸ pose a threat because officers leave with higher confidence in their abilities to detect deception, but without any reliable methods to do so.

Beyond the absolute lack of scientific reliability or accuracy, call analysis also suffers from racialized and gendered norms of appropriateness, like other lie detection techniques that make universal claims based on unrepresentative samples. In large meta-analyses of lie-detection techniques, researchers note that the race or ethnicity of respondents was rarely recorded.²²⁹ Lie detection studies fail to account for the ways that respondents across race, culture, and gender display emotions differently.²³⁰ Amanada Carlin has documented the ways that defendants are punished for performances that deviate from

²²⁴ See *id.* (quoting Steven Rhoads, Training Materials - Subconscious Communication for Detecting Danger (2018)).

²²⁵ See Charles F. Bond, Jr. & Bella M. DePaulo, *Accuracy of Deception Judgments*, 10 PERS. & SOC. PSYCHOL. REV. 214, 216, 230 (2006) (finding that officers detected deception at a 56% rate, just above chance); see also Eugenio Garrido, Jame Masip & Carmen Herrero, *Police Officers' Credibility Judgments: Accuracy and Estimated Ability*, 39 INT'L. J. PSYCH.. 254, 254 (2004) (finding that officers were less accurate than laypeople at detecting deception).

²²⁶ See Bella M. DePaulo & Roger L. Pfeifer, *On-the-Job Experience and Skill at Detecting Deception*, 16 J. APPLIED SOC. PSYCH. 249, 251–52 (1986); Garrido, Masip & Herrero, *supra* note 224, at 254; see also Saul M. Kassir, Christian A. Meissner & Rebecca J. Norwick, “I’d Know a False Confession if I Saw One”: A Comparative Study of College Students and Police Investigators, 29 L. & HUM. BEHAV. 211, 216 (2005).

²²⁷ Cf. Kassir & Fong, *supra* note 106, at 505, 510.

²²⁸ Harpster continues to tour the country training police departments in call analysis, including in West Virginia and Michigan. See, e.g., *9-1-1 Homicide: Is the Caller the Killer- Identifying Innocence or Guilt of a Caller Reporting Homicide*, ST. ALBANS POLICE DEPT., <https://saintalbanspolice.com/event/9-1-1-homicide-is-the-caller-the-killer-identifying-the-innocence-or-guilt-of-a-caller-reporting-homicide/2024-04-08/> [<https://perma.cc/KD36-MHMC>] (offering trainings in April 2024); see also Oakland Dispatch Acad., *supra* note 16) (offering trainings in March 2026).

²²⁹ DePaulo et al., *supra* note 106, at 83. Other meta-analyses make no mention of race or ethnicity. See generally Oberlander, *supra* note 215.

²³⁰ See David Matsumoto & Hyisung C. Hwang, *Judgments of Subtle Facial Expressions of Emotion*, 14 EMOTION 349, 355 (2014) (finding that witness reactions

typical white racial responses.²³¹ Lie detection techniques also fail to account for the ways that non-native English speakers respond to verbal and behavioral cues.²³²

The call analysis framework suffers from similar sampling errors. The COPS checklist of guilty and innocent indicators is based on an exploratory study of just one hundred calls.²³³ In addition to the examiners being non-blinded to the ultimate guilt or innocence of the callers, there is no indication that the calls were coded for race, gender, or age.²³⁴ The study was not a representative sample of 911 callers.²³⁵ The result is that lie detection techniques like call analysis exacerbate racial and gender biases by rendering suspect responses that do not accord with the typical responses by white male subjects.²³⁶ A study of cases where call analysis was used to convict defendants demonstrates the dangers that forensic practices can have in reinforcing existing biases. In a study of exonerees wrongfully convicted after calling 911, a disproportionate number were women accused of murder.²³⁷ Women comprise

were affected by a wide array of sociocultural factors including: familial experience of emotional expression and national culture).

²³¹ Amanda Carlin, Comment, *The Courtroom as White Space: Racial Performance as Noncredibility*, 63 UCLA L. REV. 450, 469–72 (2016); see also Note, *Dialectal Due Process*, 136 HARV. L. REV. 1958, 1960 (2023) (noting the ways that failures to account for dialectal differences in Black English or Appalachian English harm people in the legal system).

²³² See Amy-May Leach et al., *Looks Like a Liar? Beliefs About Native and Non-Native Speakers' Deception*, 34 APPLIED COGNITIVE PSYCH. 387, 388 (2019); see also Jasmine B. Gonzales Rose, *Color-Blind but Not Color Deaf*, 44 N.Y.U. REV. L. & SOC. CHANGE 309, 319 (2020) (finding that accents associated with people of color “are perceived as unappealing and unintelligible”).

²³³ Harpster, *supra* note 14, at 65.

²³⁴ See generally *id.*

²³⁵ *Id.* at 65 (Harpster conceded as much in his original thesis: “The present study examined only one hundred homicide calls and the results, should therefore, be interpreted with caution. As this was an exploratory study, additional research will be needed to confirm the reliability.”). Harpster has since stopped using this caveat when presenting his work to police departments. See, e.g., Adams & Harpster, *supra* note 9, at 23 (noting that the number of calls analyzed but not the associated limitations).

²³⁶ See E. Paige Lloyd et al., *Black and White Lies: Race-Based Biases in Deception Judgments*, 28 PSYCH. SCI. 1125, 1134 (2017) (finding that white perceivers implicitly associated Black targets as deceptive, as tracked by unconscious eye gaze, even while explicitly expressing that Black targets were more truthful); see also Gonzales Rose, *supra* note 231, at 319 (finding that “[a]ccents associated with people of color who are perceived as foreign are associated with negative evaluations of the speakers’ intelligence, competence, ambition, education, and social class.”).

²³⁷ See Kristen Gittings et al., *Behind the Call: Exploring the Characteristics of Exonerees Who Called 911*, Presentation at the Annual Conference of American

just 9% of known wrongful convictions,²³⁸ and just 10% of the incarcerated population overall.²³⁹ In these cases, call analysis was used to implicate female callers where there was little other evidence, simply because they did not accord with the ways that police thought that callers should act.²⁴⁰ For example, in the case of Jessica Logan, who was convicted for the death of her son, the police found her suspicious because “[her] crying somehow did not seem genuine” and because she later “did not appear to be upset or troubled” during an interrogation.²⁴¹

The proliferation of lie detection techniques is not only due to overblown claims about their efficacy, but the abilities their founders claim to have as well. Call analysis founder, Tracy Harpster, for example, claims to have the experience and expertise to divine truth from deception. The foundations of call analysis are based on an exploratory study from Harpster’s master’s thesis in criminal justice.²⁴² Before his graduate degree, Harpster was a police officer in the Moraine, Ohio police department, where his personnel file indicates that he had little involvement in homicide investigations or in investigating 911 calls.²⁴³ Steven Rhoads, one of the leading trainers on subconscious communication training, often plays up his experience of being a police chief in Cook County—the implication being that he worked on serious crimes in Chicago. Oft left unmentioned is that Rhoads was fired as police chief of Calumet City, a city in Cook County with a population of approximately 34,000. After his stint as police chief, Rhoads spent some years “Christian clowning” —working as a rodeo clown performing

Psychology-Law Society (Mar. 2024) (finding that women comprised 36% of exonerees who called 911, in contrast with just being 8% of exonerees overall. Of those women, 74% were convicted of homicide).

²³⁸ MEDWED, *supra* note 146, at 119.

²³⁹ See Press Release, Aleks Kajstura & Wendy Sawyer, Prison Policy Initiative, Women’s Mass Incarceration: The Whole Pie 2024 (Mar. 5, 2024), <https://www.prisonpolicy.org/reports/pie2024women.html> [<https://perma.cc/M6LY-JARD>].

²⁴⁰ See *infra* subpart III.A. In particular, police and prosecutors use tropes about good and bad mothering in call analysis cases. More generally, cases against caretakers comprise the largest proportion of wrongful convictions against women. See Elizabeth Webster & Jody Miller, *Gendering and Racing Wrongful Conviction: Intersectionality, “Normal Crimes,” and Women’s Experiences of Miscarriage of Justice*, 78 ALBANY L. REV. 973, 973 n.5 (2014–2015).

²⁴¹ Brett Murphy, *supra* note 103; see also *infra* subpart I.D.

²⁴² See HARPSTER & ADAMS, *supra* note 10, at 339.

²⁴³ See Murphy, *supra* note 8 (noting that Moraine, a town of 6,500, “averages less than one murder a year”).

“Gospel magic tricks and entertain[ing] the crowd”²⁴⁴—before beginning to train police departments in deception detection. Renee Ellory, who trains police officers in detecting subconscious deception, identifies herself as one of “only 50 people who can spot deception with exceptional accuracy” according to “scientists.”²⁴⁵ Like Rhoads, Ellory, and others, Harpster’s training flyers play up the importance of his training to “all criminal justice personnel.”²⁴⁶ They also play up his experience and trainings, claiming that he has “used his research to investigate over 1500 homicides.”²⁴⁷ Police and prosecutors are only too happy to rely on demonstrably faulty forensic techniques promoted by fake experts because the techniques help to justify arrests and secure convictions. These forensic practices give the veneer of science while appealing to ostensibly intuitive notions of behavior. In that way, Harpster, Rhoads, Ellory, and their forensic techniques are part of a larger problem that courts have long struggled to regulate.

B. Failing *Daubert* and *Frye*

This section lays out the claims of 911 analysis, reviews the literature that has examined those claims, and concludes that, like other junk forensic practices, the technique is not reliable scientific evidence that should be admissible in court. A review of the criminological literature demonstrates that, even on its face, 911 call analysis does not meet the *Daubert* or *Frye* standard. No error rate or competency studies have been performed, and high-profile cases have demonstrated that 911 call analyses have led to wrongful arrests and convictions.²⁴⁸

²⁴⁴ These facts emerged after Rhoads, who had been receiving a disability pension from Calumet City, was seen working as a rodeo clown at several rodeos. See *Rhoads v. Bd. of Trs. of City of Calumet City Policemen’s Pension Fund*, 810 N.E.2d 573, 576 (Ill. App. Ct. 1 2004).

²⁴⁵ *New Tools for Detecting Deception Training*, WIS. HIGH INTENSITY DRUG TRAFFICKING AREA, <https://embed.documentcloud.org/documents/7033547-New-Tools-for-Detecting-Deception-Training/?embed=1> [<https://perma.cc/4E2E-FLKC>] (offering trainings in November 2015).

²⁴⁶ E-mail from Waukesha County Technical College to Darrin Wearing (Sep. 5, 2019) (attaching a flyer advertising a training, explaining that “This training is specifically designed to educate *all criminal justice personnel* including Dispatchers, Law Enforcement Officers, Detectives, Supervisors, Medical Examiners and Prosecutor[s]” (emphasis added)).

²⁴⁷ See *id.*

²⁴⁸ See Cooperman, *supra* note 33; see also *Mistaken 911 Transcript Used to Charge Carpenter*, *supra* note 4.

Multiple studies have tested the claims that call analysis can identify deception in 911 calls but none have successfully replicated the original findings.²⁴⁹ With respect to the COPS scale, studies have shown that “only a handful of factors had any discernible effect in discriminating between truthful callers and deceptive callers.”²⁵⁰ In one study, only six of the forty-one variables were found to have any statistically significant effect.²⁵¹ In another, just two.²⁵² What’s more, of the factors that did show some discriminating effect, they had the *opposite* effect to that predicted by the COPS scale.²⁵³

Although the technique has buy-in from police departments that send officers to call analysis trainings and from prosecutors who use the evidence in criminal proceedings, it does not have the support of the scientific community.²⁵⁴ Strikingly, even the Federal Bureau of Investigation (FBI), which first broadcast the technique in its monthly bulletin, has since published studies from its own Behavioral Analysis Unit debunking the claims of 911 call analysis.²⁵⁵ Researchers in the FBI’s Behavioral Analysis Unit found issues with several COPS scale indicators. Indicators like Direct Pleas, Lack of Urgency, and Extraneous Information—which call analysis claims are highly predictive of deception—did not differentiate truthful callers from deceptive ones.²⁵⁶ Instead those indicators were equally present across all callers. Specifically with regard to urgency, while call analysis claims that the expression of urgency indicates truthfulness (and the lack of urgency indicates deception), it appears that conclusion was based upon popular notions of how people *should* behave in emergency situations, and not upon empirical study.²⁵⁷ Studies instead demonstrate

²⁴⁹ See e.g., O’Donnell, *supra* note 12, at 8; Markey et al., *supra* note 13, at 1041; Jon D. Cromer et al., *911 Calls in Homicide Cases: What Does the Verbal Behavior of the Caller Reveal?*, 34 J. POLICE & CRIM. PSYCH. 156, 161 (2019).

²⁵⁰ O’Donnell, *supra* note 12, at 3.

²⁵¹ *Id.* at 7.

²⁵² Cromer, *supra* note 249, at 1042.

²⁵³ Miller, *supra* note 13, at 268.

²⁵⁴ See *id.*

²⁵⁵ Compare Adams & Harpster, *supra* note 9 (FBI overviews 911 call analysis), with O’Donnell, *supra* note 12, at 1, 3 (FBI-affiliated authors discrediting FBI analysis), and Miller, *supra* note 13, at 256, 272 (same).

²⁵⁶ O’Donnell, *supra* note 12, at 8.

²⁵⁷ See Jessica M. Salerno et al., 128 *Failing to Express Emotion on 911 Calls Triggers Suspicion through Violating Expectations and Moral Typecasting*, J. PERSONALITY & SOC. PSYCH. 765, 781 (2025) (finding that law enforcement assumptions about what constitutes “normal behavior” during the reporting of violent crime

that most truthful callers do not express urgency during 911 calls.²⁵⁸

Further work by researchers associated with the FBI Behavioral Analysis Unit found that while Voice Modulation was a distinguishing factor between truthful and deceptive callers, it weighed in the opposite direction as predicted by the COPS scale. Contrary to the claims that callers who modulate their voices are more likely to be innocent, voice modulation was more likely in deceptive callers than in uninvolved ones.²⁵⁹ Similarly, while the COPS scale instructs that acceptance of a victim's death is a sign of guilt, subsequent studies found that callers notifying authorities about a dead body were far more common among uninvolved callers.²⁶⁰ That the 911 call analysis research could not be replicated or validated is not surprising given the flaws in the original study. Subsequent researchers have explained that the original call analysis study was coded, non-blind, and by just one person—Tracy Harpster himself.²⁶¹ There was no double-blind coding and testing, important steps researchers take to mitigate bias in scientific investigation.²⁶² Even more egregious is the fact that Harpster personally contacted each police department to solicit specific 911 calls and therefore knew the outcome of the calls before he coded them. The lack of non-blind coding explains why the 911 call analysis studies produced effect sizes unheard of in social science literature. According to Harpster, he identified indicators that predicted deception 91% of the time. These would be “among the largest effect sizes ever discovered in the social sciences.”²⁶³

While the purported predictive power of the technique is tantalizing to law enforcement, it simply does not pass *Daubert* or *Frye* muster. 911 call analysis not only fails when empirically tested, it fails all the other *Daubert* factors as well. Its error rate when tested has proven no more effective than chance. To date, no peer-reviewed study has been able to replicate its findings. Finally, though employed by prosecutors and police departments, the technique has not been accepted by social scientists who have reviewed the claims of call analysis. Ultimately, multiple research teams, including two from the FBI, have questioned

leads to suspicion about callers, regardless of empirical work that finds no correlation between emotion and truthfulness).

²⁵⁸ O'Donnell et al., *supra* note 12, at 8.

²⁵⁹ *See id.* at 3.

²⁶⁰ *See, e.g., id.*; Cromer, *supra* note 249, at 161.

²⁶¹ *See* Markey et al., *supra* note 13, at 1041.

²⁶² NAT'L RSCH. COUNCIL, *supra* note 39, at 124.

²⁶³ Markey et al., *supra* note 13, at 1041.

the efficacy of the COPs scale,²⁶⁴ have cautioned against its use,²⁶⁵ and more broadly, have cautioned against the use of 911 call analysis as evidence in criminal proceedings.²⁶⁶

Unlike other forensic techniques that have ostensibly created the veneer of scientific acceptance in specialized trade journals—journals like the *Journal of Forensic Science* and *Forensic Odontology* in the case of bite mark evidence, or The Association for Firearms and Tool Mark Examiners (AFTE) *Journal*²⁶⁷ in the case of toolmark comparisons—no journal exists in the field of 911 call analysis. No colleges or universities provide courses in 911 call analysis. And no study published by anyone but its founders has validated its claims.

The lack of scientific validity has led even some District Attorneys to reject the technique.²⁶⁸ This rejection is notable because, historically, prosecutors have opposed attempts to regulate forensic evidence. The Department of Justice, under multiple administrations, has rejected policies that might hinder prosecutions; even those based on faulty forensic methods.²⁶⁹ In

²⁶⁴ See *id.*; O'Donnell, *supra* note 12, at 1, 3; Miller, *supra* note 13, at 256, 272.

²⁶⁵ See Miller et al., *supra* note 13, at 270.

²⁶⁶ See *id.*; see also Markey et al., *supra* note 13, at 1046 (concluding that “it would be a mistake to use [911 call information] to make a definitive conclusion concerning criminal activity.”). Even where researchers who have not questioned the underlying premise of 911 call analysis have cautioned “that it would be a mistake for investigators to draw a firm conclusion about guilt or innocence of a caller based solely on verbal behavior during a 911 call.” Cromer, *supra* note 249, at 163.

²⁶⁷ The Journals function less like peer-reviewed scientific publications and more as trade journals that promote their respective fields. Subsequent study, for example, has debunked studies published by the AFTE as lacking “a statistical foundation supporting inferences of specific source attribution for firearms/toolmarks.” See Clifford Spiegelman & William Tobin, *Analysis of Experiments in Forensic Firearms/Toolmarks Practice Offered as Support for Low Rates of Practice Error and Claims of Inferential Certainty*, 12 L. PROBABILITY AND RISK 115, 131 (2013). In particular, the authors decried the lack of standard operating procedures for validity in the field, which makes it impossible to provide a scientific baseline for studies. *Id.*

²⁶⁸ See Murphy, *supra* note 105.

²⁶⁹ The Department of Justice, under both Democratic and Republican administrations, has resisted efforts to ensure the reliability of forensic evidence. In 2016, under President Obama, the President's Council of Advisors on Science and Technology (PCAST) recommended that government attorneys “ensure expert testimony in court about forensic feature-comparison methods meets the scientific standards for scientific validity.” PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., *supra* note 39, at 18. Despite the findings, Attorney General Loretta Lynch unequivocally refused to adopt any of the council's findings related to forensic science. Gary Fields, *White House Advisory Council Report Is Critical of Forensics Used in Criminal Trials; U.S. Attorney General Says Justice Department Won't Adopt Recommendations*, WALL ST. J. (Sep. 20, 2016), <https://www.wsj.com/articles/white-house-advisory-council-releases-report-critical-of-forensics-used-in-criminal-trials-1474394743> [https://perma.cc/4FHK-DLWW]. Further,

contrast, prosecutors in the Fair and Just Prosecution network have called the practice pseudoscience and have encouraged its member offices to reject its use.²⁷⁰ Despite these warnings, call analysis continues to be used in criminal investigations. The next section explores the deeper harms that result from the admission of shadow forensics like call analysis.

III

THE BROADER IMPLICATIONS OF CALL ANALYSIS

Beyond the harms of unreliable forensics and wrongful convictions, this Part analyzes the deeper implications that call analysis presents for the criminal legal system. The use of call analysis not only subverts evidentiary rules but presents epistemic and dignitary harm. It also contributes to legal estrangement, which presents important policy implications for public safety. Marginalized communities already have reason to be reticent about calling 911,²⁷¹ and the threat of call analysis leading to arrest and prosecution disincentivizes people from seeking help and further alienates vulnerable communities.²⁷²

A. Manufacturing Credibility Discounts

Prosecutors rely on call analysis to argue that a defendant called 911 not to obtain help, but to deceive the police—in other words, to indict the credibility of defendants. As Professor Julia Simon-Kerr writes, many of our evidence rules and traditions exist to highlight the importance of credibility in trials, including hearsay rules, impeachment rules, and jury instructions

in 2017, Attorney General Jeff Sessions decided to unilaterally disband the National Commission on Forensic Science, a commission which was borne out of the original NAS Report. Janet Burns, *Sessions Scraps Federal Commission on Forensic Accuracy, for Some Reason*, FORBES (Apr. 11, 2017), <https://www.forbes.com/sites/janetwburns/2017/04/11/sessions-scraps-federal-commission-on-forensic-accuracy-because-reasons/> [<https://perma.cc/H4B4-NTAQ>].

²⁷⁰ See Murphy, *supra* note 105.

²⁷¹ In a survey of hundreds of unhoused Portland, Oregon residents, 42% reported not feeling safe calling 911, and that number was higher for Black (48%) and Latino (52.9%) residents. CITY OF PORTLAND BUREAU OF FIRE & RESCUE, PORTLAND STREET RESPONSE: YEAR TWO PROGRAM EVALUATION 40–41 (June 2023), https://www.pdx.edu/homelessness/sites/homelessness.web.wdt.pdx.edu/files/2023-07/HRAC%20Portland%20Street%20Response%20Year%20Two%20Evaluation%20Report_FINAL%20FOR%20WEBSITE.pdf [<https://perma.cc/X5N9-KU4R>]; see also Shawn E. Fields, *Weaponized Racial Fear*, 93 TUL. L. REV. 931, 970 (2019).

²⁷² Cf. Monica C. Bell, *Police Reform and the Dismantling of Legal Estrangement*, 126 YALE L.J. 2054, 2085–87 (2017) (discussing legal estrangement).

on credibility and demeanor.²⁷³ Since the foundation of our jurisprudential system relies on jurors to be the ultimate fact-finders, it is jurors who must have the opportunity to assess witness credibility for themselves.²⁷⁴ 911 call analysis not only increases the risk of wrongful convictions but also functions as a layer between jurors and the defendant and risks displacing the jury's role as the arbiter of witness credibility.²⁷⁵ Despite hearing the literal cries for help from callers, officers who testify about call analysis render those cries illegible and twist them into proof of deception rather than desperation.

For this reason, notwithstanding call analysis's lack of reliability, it also raises the possibility of epistemic harm to a witness's ability to tell their own story. Even if the practice were foolproof, its use would still challenge "the dignity and personhood of the witness as the narrator of their own memory and experience."²⁷⁶

The use of call analysis by the state creates an epistemic injustice because it downgrades the credibility of 911 callers based not on reliable evidence, but on the hunches of the police.²⁷⁷ Miranda Fricker defines epistemic injustice as harm done to a person's ability to contribute to knowledge production.²⁷⁸ Professor Deborah Tuerkheimer, drawing on the work of Fricker's epistemic injustice framework, uses the concept of "credibility

²⁷³ Julia Simon-Kerr, *Credibility in an Age of Algorithms*, 74 RUTGERS U. L. REV. 111, 114–15, 126–27, 131 (2021).

²⁷⁴ See *id.* at 118 & n.27, 157.

²⁷⁵ A fundamental premise of our criminal trial system is that "the jury is the lie detector." *United States v. Scheffer*, 523 U.S. 303, 313 (1998) (quoting *United States v. Barnard*, 490 F.2d 907, 912 (1973) (emphasis added)).

²⁷⁶ Simon-Kerr, *supra* note 272, 157 (quoting Emily R.D. Murphy & Jesse Rissman, *Evidence of Memory from Brain Data*, J.L. & BIOSCIENCES, Jan.–June 2020, at 50.) (asserting that in the context of lie-detection algorithms, even fully accurate technologies would violate the dignity of witness testimony), <https://academic.oup.com/jlb/article-pdf/7/1/lsaa078/37356599/lsaa078.pdf> [<https://perma.cc/XC8F-QFLG>].

²⁷⁷ See Samantha Bean et al., *supra* note 191, at 193. Scholars have increasingly recognized the myriad ways that the legal system shapes knowledge production and creates epistemic harms. See generally Deborah Tuerkheimer, *Incredible Women: Sexual Violence and the Credibility Discount*, 166 U. PA. L. REV. 1, 13–15 (2017) (describing the harms suffered by sexual assault survivors in the criminal legal system); S. Lisa Washington, *Survived and Coerced: Epistemic Injustice in the Family Regulation System*, 122 COLUM. L. REV. 1097 (2022) (examining the epistemic harm suffered by survivors and their families in the family regulation system); M. Eve Hanan, *Invisible Prisons*, 54 U.C. DAVIS L. REV. 1185 (2020) (examining the epistemic harms suffered by incarcerated persons whose accounts are silenced in the process of shaping sentencing policy); Anibal Rosario-Lebrón, *Evidence's #MeToo Moment*, 74 U. MIA. L. REV. 1 (2019) (describing the ways that evidentiary rules reinforce the credibility discounting of survivors of sexual and gender-based violence).

²⁷⁸ See MIRANDA FRICKER, *EPISTEMIC INJUSTICE: POWER & THE ETHICS OF KNOWING* 1 (2007).

discounting” to describe the ways in which the criminal legal system discredits the knowledge of sexual assault survivors.²⁷⁹ In the context of call analysis, the police effectively manufacture a credibility discount.²⁸⁰ Officers who use call analysis to create evidence against suspects doubly harm their credibility. Not only does the use of call analysis cast suspicion on their calls for help, but because it focuses police suspicion on one individual, it cuts off further investigation—precluding access to evidence and leads that would corroborate or exonerate people calling for help.²⁸¹ In the case of Kathy Carpenter, had the police taken seriously her 911 call—recall that she provided police with information about the conflict between the victim and the actual perpetrator—instead of casting it aside as “extraneous information,” they would have discovered much sooner that the victim’s tenants were the true perpetrators, and Carpenter would have been exonerated months earlier.²⁸²

The use of call analysis creates credibility discounts at multiple levels. Where police officers use the technique to claim that a defendant’s call for help is deceptive, it erases the meaning of their calls, and, where defendants choose to testify, it diminishes the credibility of their testimony.²⁸³ Thus diminished, defendants are unable to share their knowledge, and as a result are “demoted from subject to object.”²⁸⁴ Call analysis also harms defendants where it effectively deprives them of their opportunity to testify. Tuerkheimer writes about the experience of sexual assault victims and the “‘pre-emptive testimonial injustice’ or ‘testimonial smothering’ [that] keeps a would-be speaker from sharing her account.”²⁸⁵ A similar dynamic plays

²⁷⁹ See *id.* (describing “testimonial injustice” as that which occurs where “prejudice causes a hearer to give a deflated level of credibility to a speaker’s word”); Tuerkheimer, *supra* note 277, at 14.

²⁸⁰ See sources cited *supra* note 279.

²⁸¹ Cf. Tuerkheimer, *supra* note 277, at 11 (discussing how reports of rape are often discounted at the outset).

²⁸² Compare Lewis, *supra* note 3, at 2 (“The caller provides extraneous information and diverts attention from herself to some people that were living in Ms. PFISTER’s house.”), with *Carpenter’s Lawyers Argue Innocence*, *supra* note 2 (describing the conflict with the tenants), and *Mistaken 911 Transcript Used to Charge Carpenter*, *supra* note 4 (providing the update that one of the tenants confessed to the murder and the charges against Carpenter were dropped).

²⁸³ Cf. Tuerkheimer, *supra* note 277, at 14 (discussing “credibility discounting” generally).

²⁸⁴ FRICKER, *supra* note 278, at 132.

²⁸⁵ See Tuerkheimer, *supra* note 277, at 45–46 (footnotes omitted) (quoting Kristie Dotson, *Tracking Epistemic Violence, Tracking Practices of Silencing*, 26 HYPATIA 236, 244 (2011)).

out with call analysis; after hearing from an officer who uses the veneer of scientific expertise to label a caller deceptive, and hearing a defendant's 911 call, jurors are "unwilling or unable to gain the appropriate uptake of proffered testimony."²⁸⁶

In these cases, it is the defendant's own call for help that preemptively silences them. In Jessica Logan's case, jurors heard her 911 call through the guilt-refracting lens of call analysis. The prosecution played her 911 call and elicited an officer's opinion that Logan's exhortation, "I need my baby" instead of "I need *help* for my baby," and her acknowledgment that "I think he's gone," were evidence of guilt rather than grief and trauma from discovering her lifeless baby.²⁸⁷ Only on the last day of trial did the jury hear Ms. Logan testify.²⁸⁸ She testified²⁸⁹ about her background, about her son's history of poor health, and her version of what happened the night of his death.²⁹⁰

Logan's testimony immediately preceded closing argument. In closing, the prosecutor argued that Logan "was the person who killed [her son]" and reminded jurors that Logan did not immediately call 911 and did not ask for help until at least forty-five seconds into her call, as proof that they should not trust her.²⁹¹ Jurors needed less than two hours to convict Logan of first-degree murder.²⁹²

Beyond the testimonial injustice done to defendants like Jessica Logan, the use of call analysis leads to hermeneutical injustice as well. Fricker identifies a hermeneutical injustice as one that obscures someone's social experience "from collective understanding owing to a structural identity prejudice."²⁹³ Where call analysis is disproportionately used against female

²⁸⁶ Dotson, *supra* note 285, at 244 (describing the effect as "testimonial oppression").

²⁸⁷ Murphy, *supra* note 103 (emphasis added).

²⁸⁸ See Partial Report: Closings, *supra* note 49, at 25.

²⁸⁹ While preemptively silenced on an epistemic level, Logan *was* able to testify. Many criminal defendants are not so lucky. Scholars have written about the ways that the criminal legal system literally silences defendants. See Bennet Capers, *Bringing Up the Bodes*, 2022 U. CHI. LEG. FORUM 84, 86 (describing the silencing of defendants throughout the criminal process, from arrest and arraignment, through trial and the appellate process); see also Anna Roberts, *Conviction by Prior Impeachment*, 96 B.U. L. REV. 1977, 2001-04 (2016) (critiquing the practice of impeachment by prior conviction as effectively precluding many defendants from testifying in their own defense).

²⁹⁰ See Partial Report: Closings, *supra* note 49, at 32-39.

²⁹¹ See *id.* at 10-11, 27.

²⁹² Murphy, *supra* note 103.

²⁹³ FRICKER, *supra* note 278, at 155.

caretakers,²⁹⁴ it both draws on and feeds into racialized and gendered notions of what “good mothering” should look like.²⁹⁵ In cases involving the tragic deaths of children, the state can transform illness or accident into crimes of “bad mothering.”²⁹⁶ In those cases, call analysis provides the corroboration that accords with the tropes that mothers who are not white, or who lack resources or time, are criminally bad mothers.²⁹⁷ Defendants who are preemptively silenced by call analysis cannot effectively communicate their experiences and are thus excluded from contributing to the wider collective understanding.

B. Deepening Legal Estrangement

The rise of call analysis also contributes to the legal estrangement of vulnerable communities. The result is that already marginalized groups who may have reason to fear calling the police will be less likely to call, less likely to receive help, and less likely to see the responding police as legitimate. Exacerbating the problem are the racialized and gendered norms that already disfavor women and people of color.²⁹⁸ Call analysis serves as a veneer of scientific expertise that twists the words of callers who are already at risk of discreditation.²⁹⁹

This section uses the framework of legal estrangement to explain some of the broader implications of call analysis. Professor Monica Bell coined the term “legal estrangement” in part to explain the distrust between marginalized groups and legal system actors—including courts and law enforcement.³⁰⁰ She identifies

²⁹⁴ Markquart, Salerno & Murphy, *supra* note 23; *see also* Gittings et al., *supra* note 236.

²⁹⁵ *See* Webster & Miller, *supra* note 240, at 975.

²⁹⁶ Valena E. Beety, “Unfit”: Gender, Ableism and Reproductive Wrongful Convictions, UCLA L. REV. (Forthcoming 2026); Webster & Miller, *supra* note 240, at 1030; *see also* Washington, *supra* note 277.

²⁹⁷ *See* Washington, *supra* note 277; Webster & Miller, *supra* note 240, at 975; Beety, *supra* note 296.

²⁹⁸ *See* Webster & Miller, *supra* note 240, at 1030; Richard A. Fabes & Carol Lynn Martin, *Gender and Age Stereotypes of Emotionality*, 17 PERSONALITY & SOC. PSYCH. BULL. 532, 532 (1991) (finding that expectations of emotionality differ by gender—with women expected to express more emotionality than men); Beety, *supra* note 296 (explaining that “the gender of a female defendant makes her particularly vulnerable to a wrongful conviction because she is perceived as both violating the law and as violating social gendered norms.”).

²⁹⁹ *See* Amy D. Ronner, *The Cassandra Curse: The Stereotype of the Female Liar Resurfaces in Jones v. Clinton*, 31 U.C. DAVIS L. REV. 123, 130 (1997); *see also* Rosemary C. Hunter, *Gender in Evidence: Masculine Norms vs. Feminist Reforms*, 19 HARV. WOMEN’S L.J. 127, 165 (1996).

³⁰⁰ Bell, *supra* note 272, at 2054.

the problem of estrangement as “a cultural orientation about the law that emanates from symbolic and structural exclusion—that is, both subjective and objective factors.”³⁰¹

The power of estrangement theory is that it explains how vulnerable communities can simultaneously view the police as legitimate while also believing that the police cannot be trusted to protect them or their community.³⁰² In this way, legal estrangement complicates notions of mistrust and cynicism. Where some might suggest that in communities where mistrust of the police is high, “that a large subset of this group possesses values that are antithetical to law-abiding behavior,” Bell theorizes how a poor Black resident could both sympathize with the police but also believe that the police are too corrupt to help her.³⁰³ Similarly, in earlier work on legal cynicism, Bell demonstrates that Black women simultaneously score highly on measures of cynicism and also call the police at high rates, building on research that “police receive more calls from the people and neighborhoods highest in cynicism”.³⁰⁴

Estrangement theory applies squarely to the communities against whom call analysis is deployed, particularly given its disproportionate use against economically vulnerable female caretakers.³⁰⁵ Despite having reason to mistrust the police, they reached out to the police—as evinced by their 911 calls for help. But rather than getting help, as callers in affluent neighborhoods do,³⁰⁶ they are often arrested and prosecuted.³⁰⁷ Estrangement theory best maps onto the experiences of vulnerable communities because the issue is estrangement from, rather than noncompliance with, police.³⁰⁸ In call analysis cases, defendants are typically extremely compliant. Not only do they call the police, but in the case of Jessica Logan for

³⁰¹ *Id.* at 2100.

³⁰² *See id.* at 2087.

³⁰³ *See id.* at 2087, 2092.

³⁰⁴ *See* Monica C. Bell, *Situational Trust: How Disadvantaged Mothers Reconceive Legal Cynicism*, 50 L. & Soc. REV. 314, 316, 325, 338 (2016) (defining legal cynicism broadly as distrust of the law).

³⁰⁵ *See* Gittings, *supra* note 237.

³⁰⁶ *See* Bell, *supra* note 272, at 2123 (contrasting the slow response times of police in urban West Baltimore with their speed in responding to calls in more affluent, neighboring, Baltimore County).

³⁰⁷ *Cf.* Fields, *supra* note 271, at 935 (discussing how “bias-motivated civilians weaponize law enforcement to respond to their irrational racial fears”).

³⁰⁸ *See* Bell, *supra* note 272, at 2099 (comparing legal estrangement to legitimacy theory, which would predict noncompliance with police from racially and socially marginalized communities).

example, they may provide multiple statements to the police, agree to multiple interviews, and agree to do a reenactment.³⁰⁹

Weaponizing the cooperation of vulnerable callers deepens their legal estrangement by contributing to the subjective experience of injustice and to the objective structural exclusion of vulnerable communities that call 911 for help.³¹⁰ Defendants who have been convicted with the use of call analysis express a sense of alienation from having their own calls for help used against them. Riley Spitler expressed shock that his trial court did not perform any of its gatekeeping responsibilities in letting in call analysis without even a *Daubert* hearing.³¹¹ Similarly, Russ Faria was “appalled” by the fact that call analysis was used to wrongfully convict him for the death of his wife.³¹² And Jessica Logan has expressed the helplessness of being suspected in her son’s death simply because her words did not accord with the detective’s notions of appropriateness.³¹³

Call analysis also exacerbates legal estrangement by reinforcing the objective structural exclusion of vulnerable groups. Structural exclusion describes the ways that the criminal legal system produces and normalizes vastly disparate treatment based on the race and class of communities.³¹⁴ The system’s policies and structures shut marginalized communities out of the legal process, ensuring that resources only go to those who are already well off.³¹⁵ It leads marginalized communities to be simultaneously overpoliced and under-protected.³¹⁶ The use of call analysis serves as another means

³⁰⁹ Partial Report: Closings, *supra* note 170, at 36 (At a detective’s request, Logan performed a reenactment of the night of her son’s death at her home.).

³¹⁰ See Bell, *supra* note 272, at 2100.

³¹¹ Speaking about the trial judge’s failure to perform any evidentiary reliability checks, Spitler said: “It’s just insane that a judge wouldn’t be the wiser to that. . . . But that’s our system.” Murphy, *supra* note 8.

³¹² See *id.*

³¹³ Murphy, *supra* note 103 (“You shouldn’t feel like somebody’s guilty just because the way their emotions is. Everybody react and have emotions different than everybody else,” she said, a refrain she’s repeated since she was first arrested. “I mean there really nothing I could have said differently or did differently. I shouldn’t had to reword stuff to make it, I guess, to make [the Detective]—I feel like I did nothing wrong on the 911 call.”)

³¹⁴ See Bell, *supra* note 272, at 2114.

³¹⁵ See *id.* at 2114–15.

³¹⁶ See *id.* at 2118; see also SIMON BALTO, OCCUPIED TERRITORY: POLICING BLACK CHICAGO FROM RED SUMMER TO BLACK POWER 2, 6 (Heather Ann Thompson & Rhonda Y. Williams, eds. 2019); Chaz Arnett, *Race, Surveillance, Resistance*, 81 OHIO ST. L.J. 1103, 1104 (2020) (describing how police disproportionately test new technologies and techniques on poor Black communities).

of excluding these communities. While fears of arrest and police violence have long been concomitant dangers associated with calling the police, the advent of call analysis provides cover for officers who seek to police, rather than protect, vulnerable communities. The presence of call analysis heightens the risks of calling 911, particularly as the tenets of call analysis spread through automated systems.³¹⁷ Deepening legal estrangement has important policy implications. Communities already reticent about calling 911 for emergencies may become increasingly less likely to,³¹⁸ furthering estrangement and under-protection.

IV

SHINING A LIGHT ON SHADOW FORENSICS: FORCING DISCLOSURE AND LIMITING ITS SPREAD

The history of forensics tells us that prosecutors will continue to wield faulty forensics against people to secure convictions.³¹⁹ And technicians of those forensic techniques will benefit by creating cottage industries that cater to prosecutions with trainings, publications, and expert testimony.³²⁰ Given this backdrop, the rise and spread of 911 call analysis is not exceptional, but stemming its spread may prove difficult. This Part evaluates solutions that not only map onto the problem of call analysis but provide a framework for limiting shadow forensics more broadly. It focuses on solutions that create transparency and force the state to disclose its use of forensics. Bringing these forensic practices out of the shadows is important because defendants cannot challenge techniques if they are unaware of their use.³²¹

It evaluates in turn: (1) efforts to strengthen *Daubert* and *Frye*; (2) transparency litigation; and (3) police procurement oversight. Ultimately, it concludes that to stem the spread of

³¹⁷ Deceptio.AI's software claims to be able to accurately diagnose deception based merely on transcripts run through its program. Beam, *supra* note 20.

³¹⁸ See Bell, *supra* note 304, at 322 (describing the pre-existing concerns that Black women with children may fear when considering engaging with the police, including: arrest, incarceration, police violence, nonchalance, racial bias, and threats).

³¹⁹ See Fabricant & Carrington *supra* note 16, at 53 (describing prosecutors' use of faulty bitemark evidence to secure convictions, despite obvious issues); see also Beety, *supra* note 19, at 48 ("If history is our guide, prosecutors will proceed to charge pregnant people with crimes due to their pregnancy behaviors either with no scientific evidentiary support or by relying on junk science.").

³²⁰ See FABRICANT, *supra* note 195.

³²¹ See Sinha, *supra* note 52, at 1.

911 call analysis, a multiprong approach that maps onto the entry points of shadow forensics is necessary.³²²

A. The Limits of *Daubert* and *Frye*

Since the inception of *Daubert* in 1993, scholars have reflected on how to strengthen its promise of judicial gatekeeping and evidentiary reliability. Proposals have encouraged independent judicial research,³²³ sought to amend the Federal Rules of Evidence,³²⁴ and advocated for stricter internal guidelines in prosecutor's offices.³²⁵ As forensic testimony is often admitted with, at best, cursory examination, on the basis of precedent, courts must take seriously their gatekeeping responsibilities to establish precedent that denies unreliable forensics. The more courts critically examine evidence, the easier it will be for other courts to do the same.

Compared to criminal cases, federal civil courts have been more successful when applying *Daubert* to deny the admission of unreliable forensic evidence testimony.³²⁶ Professors Oliva and Beety identify civil disclosure rules as the key to this discrepancy. While the Federal Rules of Civil Procedure mandate pretrial disclosure of expert evidence, the Criminal Rules of Procedure have no such requirements.³²⁷ Criminal defendants are entitled to extremely limited pretrial discovery, which often leaves them unaware of the forensic evidence against them until the eve of trial (if at all).³²⁸ To close the civil-criminal *Daubert* gap³²⁹ Oliva and Beety propose the adoption of the Federal Rules of Civil Procedure and disclosure rules into criminal trials.³³⁰

³²² Recognizing the carceral logic of forensic evidence, this strategy accords with scholarship that seeks to reimagine the evidence rules. See Sinha, *supra* note 50, at 880; Bennett Capers, *Evidence Without Rules*, 94 NOTRE DAME L. REV. 867, 867 (2018) (calling for “radical rethinking of Rules of Evidence”).

³²³ Cheng, *supra* note 79, at 1264.

³²⁴ See Oliva & Beety, *supra* note 16, at 1.

³²⁵ Findley, *supra* note 55, at 762, 774.

³²⁶ See Valena E. Beety & Jennifer D. Oliva, *Emancipating Expert Evidence*, in CRITICAL EVIDENCE (Bennett Capers, Jasmine Harris & Julia Simon-Kerr eds., forthcoming 2026).

³²⁷ See *id.* (manuscript at 130).

³²⁸ See *id.* (manuscript at 124).

³²⁹ See Erin Murphy, *Neuroscience and the Civil/Criminal Daubert Divide*, 85 FORDHAM L. REV. 619, 621–22 (2016) (contrasting *Daubert* decisions in civil cases, where *Daubert* was strictly applied, to criminal cases, where similar prosecution evidence was largely admitted); see also D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99, 99 (2000).

³³⁰ See Oliva & Beety, *supra* note 16, at 17.

Given how shadow forensics like call analysis are used without formal proffering at trial, early disclosure of expert consultation or law enforcement use would be an important step to combat its spread. But scholars have expressed doubts that evidentiary reform alone can lead to substantive change.³³¹ *Daubert* and *Frye* are premised on excluding unreliable forensics at trial only, and solutions must preempt and target the earlier points through which shadow forensics enter the criminal legal process. Even if enough precedent is created to exclude the use of call analysis, the history of forensics tells us that newer junk forensics will emerge. Guarding against shadow forensics requires solutions that preempt its earlier entry points.

B. Transparency Litigation

To account for the fact that shadow forensics enter cases at entry points not governed by traditional reliability analysis, potential solutions must seek disclosures in novel ways and at novel times—not just at the trial stage, where *Daubert* and *Frye* hearings are litigated, but at a macro scale so that the state's use of shadow forensics is rendered visible. At the trial level, Professor Maneka Sinha suggests that defense attorneys play a crucial “forcing function” by litigating the reliability of police technologies in Fourth Amendment litigation before the start of trial.³³² The defense forcing function is critically important, particularly where officers rely on automated technologies but lack any personal knowledge of how they operate and, therefore, of their reliability.³³³ Defense counsel could play a similar forcing function to challenge the reliability of forensic techniques and demand the disclosure of forensic practices used against their clients. However, they must be made aware of such practices to make these challenges.

Transparency litigation, which is “used to compel the disclosure of government records,”³³⁴ could help render visible the depth of shadow forensics.³³⁵ Proposed in the context of discovering the extent of the government's use of surveillance technology,

³³¹ See Capers, *supra* note 322, at 873–74 (characterizing the Rules of Evidence as selectively blind to the evidence it screens versus that which it lets “breeze unchecked past”); Murphy, *supra* note 50, at 760 (describing how the current evidentiary regime “effectively railroads” forensic evidence).

³³² See Sinha, *supra* note 52, at 4.

³³³ See *id.* at 1.

³³⁴ See Bloch-Wehba, *supra* note 39, at 922.

³³⁵ *Id.*

transparency litigation “has become a core technique for rendering obscure investigative practices visible and holding police accountable.”³³⁶

The litigation forms a “criminal procedure ‘shadow docket.’”³³⁷ While the litigation itself does not resolve the merits of law enforcement technologies, it lays the groundwork for oversight and, ultimately, resistance.³³⁸ It often includes open records requests and First Amendment suits.³³⁹ In New York City, the Legal Aid Society successfully sued the New York City Police Department (NYPD) to force disclosure of its surveillance contracts.³⁴⁰ Despite court orders to disclose the records, the NYPD repeatedly refused to turn over the records, first arguing that the requests for records were not specific enough, and later arguing that turning the records over would be burdensome.³⁴¹ The appellate court rejected the NYPD’s arguments and recently ordered that the records be turned over every three months.³⁴²

Transparency litigation is an important complement to Fourth Amendment litigation and to reliability analysis. Indeed, the extent of call analysis’s spread came to light because of extensive open records requests to prosecutor’s offices and police departments by journalists.³⁴³ Building on that project, this Article examined training advertisements and training materials obtained via hundreds of open records requests from police departments across multiple states.³⁴⁴ Transparency litigation not only has important benefits for democratic accountability³⁴⁵ and oversight but also has implications for evidentiary reliability. It brings to light practices and techniques that communities may be unaware are being used against

³³⁶ *Id.*

³³⁷ *Id.* (describing transparency lawsuits that resolve not the merits of police investigatory tools but “resolve important questions about democratic governance of policing.”).

³³⁸ *See id.*

³³⁹ *Id.* at 922 & n.17; Jonathan Manes, *Secrecy & Evasion in Police Surveillance Technology*, 34 BERKELEY TECH. L.J. 503, 512 (2019).

³⁴⁰ *See* Legal Aid Society v. Records Access Officer, 238 AD3d 17, 25, (N.Y. App. Div. 2025).

³⁴¹ The Appeals Court in reaffirming the order for the NYPD to turn over the documents noted that it was the NYPD that did not digitize the documents in order to “keep the contract documents secret” in the first place. *See id.* at 20–21.

³⁴² *See id.*

³⁴³ *See* Murphy, *supra* note 8.

³⁴⁴ *See* sources cited *supra* notes 228, 245–246.

³⁴⁵ *See generally* Barry Friedman & Maria Ponomarenko, *Democratic Policing*, 90 N.Y.U. L. REV. 1827 (2015) (describing efforts to constrain police discretion through democratic rules and policies).

them. It therefore allows defendants in particular cases to challenge the use of shadow forensics like call analysis. In *Baldwin*, evidence that the detective who testified about Baldwin's 911 call was trained in call analysis did not come to light until post-conviction counsel made open records requests for the detective's training record.³⁴⁶ It was only after call analysis was rendered visible that Baldwin was able to challenge the reliability of the practice in her post-conviction hearing. The next section turns to solutions that target the earliest entry point that shadow forensics enter the criminal legal system: police trainings.

C. Police Procurement Oversight

Regulating the police procurement of unreliable forensic techniques may be a promising step in battling the spread of junk forensics like call analysis. Borrowing a tactic that some cities have taken to regulate the use of police surveillance technology may provide a useful blueprint for regulating unreliable forensic techniques. Over the last decade, approximately twenty jurisdictions have enacted laws that empower local governments to oversee and regulate the surveillance technologies that law enforcement procures and employs.³⁴⁷

Cities like Seattle, Oakland, and Chicago have taken varying steps to require oversight of, and government control over, police spending on surveillance technology.³⁴⁸ In Seattle, the police repeatedly misled the city council about the extent of their planned surveillance.³⁴⁹ When the wide extent of the surveillance came to light, the backlash prompted the city council to enact an ordinance requiring council approval for surveillance equipment.³⁵⁰ In Oakland, concerted organizing efforts led the City Council to limit the areas that police were

³⁴⁶ See Transcript of Proceedings at 13–14, *State of Minnesota v. Baldwin*, 27-CR-18-23849 (Minn. Dist. Ct. Hennepin Cnty. Sep. 11, 2024).

³⁴⁷ See Vincent M. Southerland, *The Master's Tools and a Mission: Using Community Control and Oversight Laws to Resist and Abolish Police Surveillance Technologies*, 70 UCLA L. REV. 2, 2 (2023), see also Catherine Crump, *Surveillance Policy Making by Procurement*, 91 WASH. L. REV. 1595, 1660–62 (2016).

³⁴⁸ See Crump, *supra* note 347, at 1595, 1615; Chip Mitchell, *Researchers Are Skeptical of This Gunfire Detection System. Chicago No Longer Uses It*, NPR (Oct. 20, 2024) <https://www.npr.org/2024/10/20/nx-s1-5131546/researchers-are-skeptical-of-this-gunfire-detection-system-chicago-no-longer-uses-it> [<https://perma.cc/9SGD-A9KA>].

³⁴⁹ See Crump, *supra* note 347, at 1606, 1615.

³⁵⁰ See *id.* at 1613–15.

to surveil.³⁵¹ The council also voted to create a Privacy Advisory Commission to “engage the public and obtain public input, and to draft a surveillance oversight ordinance.”³⁵² In Chicago, Mayor Brandon Johnson refused to renew the contract of ShotSpotter, an automated gunfire detection tool that purports to be able to identify the location of gun shots in real time.³⁵³ The decision to not renew the contract with ShotSpotter came after repeated studies showing that the technology was unreliable, ineffective, and costly.³⁵⁴ Furthermore, the technology was disproportionately deployed in Black and brown neighborhoods.³⁵⁵ In all these cities, governmental action followed after law enforcement surveillance methods were surfaced.

Efforts to regulate police use and procurement of unreliable forensics could prove similarly helpful. Police departments spend thousands of dollars per officer on 911 call analysis trainings that have no basis in science.³⁵⁶ Examining the validity of such trainings before sending employees would not be controversial in the context of other governmental agencies.³⁵⁷ Regulation of the forensic trainings and technology that law enforcement pays for would thus accord with theories of democratic policing; in particular, the notion that law enforcement agencies should be governed and constrained by the same laws and regulations that constrain other agencies.³⁵⁸

³⁵¹ See *id.* at 1626.

³⁵² See Southerland, *supra* note 347, at 40.

³⁵³ See Mitchell, *supra* note 348.

³⁵⁴ See Sinha, *supra* note 95, at 83 (describing the errors that ShotSpotter makes, including: misclassifying loud noises as gunfire, missing gunfire because of background noise, and providing incorrect location of gunfire); see also Motion for Leave to File Brief as Amici Curiae in Support of Defendant’s Motion for a Frye Hearing at 10, Brighton Park Neighborhood Council et. al., *State of Illinois v. Williams*, 20 CR 0899601 (May 3, 2021) (detailing MacArthur Justice Center analysis that ShotSpotter led to over 40,000 unfounded deployments over twenty-one months in Chicago). Similarly, an audit of ShotSpotter in New York found that the system was “overwhelmingly inaccurate and leads officers to spend hundreds of hours each month investigating nonexistent shots.” Alyce McFadden, *Gunshot Detection System Wastes N.Y.P.D. Officers’ Time, Audit Finds*, NY TIMES (June 20, 2024) (citing N.Y.C. COMPTROLLER’S OFFICE, AUDIT REPORT ON THE NEW YORK CITY POLICE DEPARTMENT’S OVERSIGHT OF ITS AGREEMENT WITH SHOTSPOTTER INC. FOR THE GUNSHOT DETECTION AND LOCATION SYSTEM (June 20, 2024)), <https://www.nytimes.com/2024/06/20/nyregion/nypd-shotspotter-guns.html> [<https://perma.cc/6C45-QC75>].

³⁵⁵ Sinha, *supra* note 95, at 87 (finding that the technology was deployed in half of the city’s districts, with the twelve featured districts having the largest concentration of people of color).

³⁵⁶ See Markey, *supra* note 214, at 29.

³⁵⁷ See Friedman & Ponomarenko, *supra* note 345, at 1848–49.

³⁵⁸ See *id.*

The local efforts to regulate surveillance have also had drawbacks. Those limitations can help guide more fulsome regulation of shadow forensics. For one, Oakland's Privacy Advisory Commission, like other similar entities in different jurisdictions, has only advisory, and not substantive, power. Law enforcement is free to ignore the requests and decisions of the commission.³⁵⁹ For another, there are vast power and resource imbalances between law enforcement and the city council commissions.³⁶⁰ Even where they could issue mandates, the commissions may lack the resources to adequately research and assess the technologies.³⁶¹ Finally, there are general concerns about the representativeness of interests on such commissions. The communities most affected by police surveillance are often "locked out of the democratic process."³⁶²

Drawing on the lessons in the surveillance context, any forensic oversight commission must be imbued with actual power. That includes not just substantive budgetary power, but subpoena power so it can surface the forensic practices that the police try to procure. Furthermore, any commission should be supported with the expertise and resources to evaluate forensic techniques. To address issues of representation, oversight commissions must recruit and value the input of community members most affected by law enforcement.³⁶³

Although oversight commissions are no panacea for stopping shadow forensics, they may be able to function as a ratchet for future action that shrinks the police use of unreliable forensic techniques. Even where an oversight commission lacks full power over police spending, disclosure and examination are still critically important. Having the ability to examine how police money is spent and evaluating the reliability of a practice can give credence to other measures that may keep shadow forensics out of the criminal legal system. In *Baldwin*, the defense challenged the prosecutor's use of call analysis under *Napue v. Illinois* and its Minnesota analogue.³⁶⁴ The court in *Napue* held

³⁵⁹ See Southerland, *supra* note 347, at 29 (describing by analogy, the inefficacy of many Civilian Complaint Review Boards to rein in police misconduct).

³⁶⁰ See *id.* at 60.

³⁶¹ See *id.* at 64 (detailing the immense amount of work that oversight groups must do to review police technologies—often with no additional resources).

³⁶² See Bell, *supra* note 272, at 2143.

³⁶³ See Southerland, *supra* note 347, at 66.

³⁶⁴ *Baldwin v. State*, No. A25-0555, 2026 WL 264721, at *8–12 (Minn. Ct. App. Feb. 2, 2026) (rejecting Baldwin's claim as to the 911 analysis but reversing and remanding on other grounds); *Napue v. Illinois*, 360 U.S. 264, 269–72 (1959).

that using false evidence violates Due Process. The court fashioned the following test for relief: (1) whether false evidence was used to convict the defendant; (2) whether the prosecutor knew the testimony was false; and (3) if there was a “reasonable likelihood” that the false testimony affected the jury’s decision.³⁶⁵ An oversight commission, even absent substantive budgetary control, could scrutinize call analysis. Its conclusions about unreliability could then render dispositive the first two *Napue* factors. Had Minneapolis had an oversight commission that evaluated the use of call analysis using existing empirical studies, Baldwin could have used its conclusions as *de facto* proof that (1) the call analysis testimony was false, and that (2) the prosecutor knew, or should have known, of its falsity. Baldwin did not have that benefit, and her call analysis claim was rejected.³⁶⁶

CONCLUSION

The use of unreliable junk forensics has long plagued the criminal legal system. Despite changes in evidentiary rules, updates in case law, and reports by the scientific community, junk forensics continue to be used to convict people. 911 call analysis is one of the latest forensic techniques that has been used despite its lack of reliability or validity. But it also illustrates something bigger. Using call analysis as a case study, this Article introduces the framework of Shadow Forensics—forensic practices that elide judicial gatekeeping through multiple extra-legal entry points—to understand how unreliable forensics come to be used against defendants and how difficult it can be to regulate them. The stakes are high: unreliable forensics play a central role in wrongful convictions. A multi-prong effort that encompasses disclosure rules, transparency litigation, and police procurement oversight will be necessary to surface these practices and help stem the tide of shadow forensics.

Minnesota Courts apply a similar but more lenient test espoused by the Seventh Circuit in *Larrison v. United States*, 24 F.2d 82, 87–88 (7th Cir. 1928) (requiring that (1) the court is satisfied that testimony was false; (2) “without it the jury *might* have reached a different conclusion”; and (3) the moving party was “taken by surprise” by the evidence).

³⁶⁵ See *Napue*, 360 U.S. at 269–72.

³⁶⁶ *Baldwin*, 2026 WL 264721, at *10–12.