DEEPFAKES IN INTERROGATIONS

Wayne A. Logan[*]

In recent years, academics, policymakers, and others have sounded the alarm over police use of artificial intelligence in areas such as predictive policing, gunshot detection, and facial recognition. One area not receiving attention is the interrogation of suspects. This Article addresses that gap, focusing on the inevitable coming use by police of AI-generated deepfakes to secure confessions, such as by creating and presenting to suspects a highly realistic still photo or video falsely indicating their presence at a crime scene, or an equally convincing audio recording of an associate or witness implicating them in a crime.

That police should resort to trickery to secure a confession is nothing new. Indeed, in Frazier v. Cupp *(1969), the Supreme Court condoned a police lie to a suspect that an associate implicated him in a crime, holding that the deceit did not violate due process because it did not render the confession secured involuntary, while positing that an innocent individual would not falsely confess. Building upon the now-recognized reality that innocents do indeed confess and research demonstrating the coercive impact of police use of the "false evidence ploy" (FEP) in securing confessions, scholars have urged a general ban on its use. Courts, while often expressing dismay over police resort to FEPs, typically conclude that they do not violate due process, but at times have held otherwise, expressing particular concern over police presentation of fabricated physical evidence to suspects (as opposed to orally relating its existence, as in* Frazier*).*

While sympathetic to a ban on police deceit in interrogations more generally, this Article singles out deepfakes for specific concern, based on their unprecedented verisimilitude, the demonstrated inability of the lay public to identify their falsity (despite confidence to the contrary), and

---

* Steven M. Goldstein Professor, Florida State University College of Law. Thanks to Brandon Garrett, Anjelica Hendricks, Richard Leo, David Logan, Michael Mannheimer, Eve Brensike Primus, Chris Slobogin, and George Thomas for their comments on prior drafts, and thanks to Kat Kelpher for her research assistance.

the common belief that police are not permitted to lie about evidence, much less fabricate it. Ultimately, the Article makes the case for reconsideration of Frazier, *based on research findings of the past fifty years, as well as the many major changes to the criminal legal system since 1969, especially the significantly increased pressure felt by defendants to plead guilty (very often on the basis of confessions, rendering them more susceptible to FEPs).*

A ban on deepfakes will also have important functional benefits. These include providing ex ante guidance to police, who lack clarity on the parameters of permissible interrogation techniques, and judges, who must decide motions to suppress based on application of the notoriously indeterminate due process voluntariness standard. More broadly, a ban will act as a partial yet important bulwark against the deleterious wave of disinformation now sowing distrust in governmental actors and institutions. If deepfakes are condoned in interrogations, it is not hard to imagine that judges, jurors, witnesses, and members of the public will be skeptical of the reliability of evidence in criminal cases, undermining a cornerstone of the nation's constitutional democracy.

The Article concludes with a discussion of how a ban can be achieved and why ameliorative tweaks to the current framework regulating confessions are not up to the challenge of checking the formidable threats posed by police use of deepfakes in interrogations.

TABLE OF CONTENTS

INTRODUCTION

Imagine that police arrest D1 and D2 based on suspicion that they robbed a store. At the police station, both arrestees waive their *Miranda* rights,[1] as individuals commonly do,[2] or are not in "custody" or being "interrogated," avoiding *Miranda*'s protections.[3] Police isolate the suspects and engage in two distinct interrogation strategies. In one room, they falsely inform D1 that D2 has implicated him in the robbery, and D1 confesses to the crime. In another room, police show D2 a video they created, employing highly realistic, artificial-intelligence-generated imagery and voice of D1—a "deepfake"—implicating D2 in the robbery, and D2 confesses. Were the confessions of D1 and D2 voluntary, as due process requires?

Under current law, the answer regarding D1's confession is "yes," and D2's confession "maybe." For decades, police have been permitted to make oral misstatements about the existence of incriminating physical and testimonial evidence (D1's confession).[4] The judicial reaction to police fabrication of actual tangible forms of evidence and presentation of it to a suspect to secure a confession (D2's confession), however, has been less positive. While most courts condone the practice,[5] several find it coercive and violative of due process.[6]

To the latter camp, there exists a "qualitative difference between the verbal artifices deemed acceptable and the presentation of the

1. *See* Miranda v. Arizona, 384 U.S. 436, 444–45 (1966).

2. *See* Ian Farrell & Nancy Leong, *How Crime Dramas Undermine* Miranda, 14 U.C. IRVINE L. REV. 211, 221–24 (2024) (noting the high rate of *Miranda* waivers and discussing possible reasons for it). Waiver is especially common among factually innocent individuals. *See* Brandon L. Garrett, *The Substance of False Confessions*, 62 STAN. L. REV. 1051, 1092 (2010). This is perhaps due to the belief that they "did not have anything to hide" and that "the power of their own innocence [will] set them free." Saul M. Kassin, *On the Psychology of Confessions: Does Innocence Put Innocents at Risk?*, 60 AM. PSYCH. 215, 218 (2005).

3. *See Miranda*, 384 U.S. at 467 (extending protections to only "in-custody interrogation"); *see also* Richard A. Leo, *Questioning the Relevance of* Miranda *in the Twenty-First Century*, 99 MICH. L. REV. 1000, 1017 (2001) (discussing the strategic value to police of circumventing *Miranda* protections by avoiding circumstances constituting custody and/or interrogation).

4. *See* Frazier v. Cupp, 394 U.S. 731, 739 (1969) (holding that due process was not violated when a confession was obtained by police who falsely told the defendant that a codefendant had confessed and implicated the defendant); *see also infra* notes 69–76 and accompanying text.

5. *See infra* notes 77–78 and accompanying text.

6. *See infra* notes 79–112 and accompanying text.

falsely contrived [evidence]."[7] This is because an individual is likely "more impressed and thereby more easily induced to confess when presented with tangible, official-looking reports as opposed to merely being told that some tests have implicated him."[8] Moreover, such deception is problematic because it threatens to erode faith in the integrity of police and is contrary to the common belief that police will not "knowingly fabricate tangible documentation or physical evidence against an individual."[9] Finally, concern exists that manufactured or falsified documents could affect legal proceedings after their use in a suspect's current prosecution.[10]

This Article builds upon these concerns in an as-yet-unexamined, even more problematic realm of falsified tangible evidence: audio, audiovisual, or still-image content generated by artificial intelligence (AI), colloquially known as deepfakes. Today, a variety of software tools such as FakeApp, FaceSwap, Voice Engine, Sora, Photoshop, Pixlr, GIMP, OpenShot, ZAO, WaveNet, and Filmora allow anyone, including police, to readily create media of highly compelling realism.[11] Whereas in the past police had to resort to splicing and doctoring to create even remotely realistic interrogation room deceits, they can now fabricate highly realistic and genuine-appearing media easily and at little cost.[12] Although there exist no reported instances of police interrogation deepfakes to date, it is safe to assume that the absence will not last and that deepfakes will be coming soon to a police station near you.[13] This Article is intended to provide a framework for courts and policymakers to employ when this comes to pass.

---

7. State v. Cayward, 552 So. 2d 971, 973 (Fla. Dist. Ct. App. 1989).

8. *Id.* at 974.

9. *Id.* ("[W]e think the manufacturing of false documents by police officials offends our traditional notions of due process of law under both the federal and state constitutions.").

10. *Id.* at 975.

11. Michael D. Murray, Generative Artifice: Regulation of Deepfake Exploitation and Deception Under the First Amendment 2 (June 21, 2024) (unpublished manuscript), https://perma.cc/8AHU-89BZ; Benj Edwards, *OpenAI Collapses Media Reality with Sora, a Photorealistic AI Video Generator*, ARS TECHNICA (Feb. 16, 2024), https://arstechnica.com/information-technology/2024/02/openai-collapses-media-reality-with-sora-a-photorealistic-ai-video-generator/; Cade Metz, *OpenAI Unveils A.I. That Instantly Generates Eye-Popping Videos*, N.Y. TIMES (Feb. 15, 2024), https://perma.cc/6JWN-EQJB.

12. *See* Rebecca A. Delfino, *Pornographic Deepfakes: The Case for Federal Criminalization of Revenge Porn's Next Tragic Act*, 88 FORDHAM L. REV. 887, 892–93 (2019); *see also* Eryn J. Newman & Norbert Schwarz, *Misinformed by Images: How Images Influence Perceptions of Truth and What Can Be Done About It*, CURRENT OP. PSYCH., Apr. 2024, at 1, 1.

13. A view shared by two Kansas Supreme Court justices who recently dissented from their colleagues' approval of a police interrogation lie about the supposed ironclad scientific reliability of a "computerized voice stress analysis" machine. *See* State v. Garrett, 555 P.3d 1116, 1119–20, 1132 (Kan. 2024) (Rosen, J., dissenting, joined by Wall, J.) (noting the "[h]yper-realistic" nature of

The discussion proceeds as follows. Part I discusses police use of what is known as the false evidence ploy (FEP), the central role it plays in securing confessions, and the judicial response to the practice.[14] Over fifty years ago, in *Frazier v. Cupp*,[15] the Supreme Court condoned a particular FEP, holding that falsely telling a murder suspect that a confederate had confessed did not render the suspect's confession involuntary for due process purposes.[16] Since then, most courts have relied on *Frazier* to categorically deny challenges to police use of FEPs. As noted, however, several courts have been more critical of FEPs, especially those involving presentation to suspects of police-fabricated physical evidence, deeming the practice coercive and violative of due process.

Part II examines pertinent social science research. It first surveys research demonstrating the significant impact that FEPs have on inducing confessions, including from innocents, especially when the FEP is presented in physical form (as opposed to being orally told of its existence). It then discusses research showing the inability of individuals to detect deepfakes, despite their confidence that they can do so, a deficit exacerbated by the ever-increasing quality of deepfakes and the commonly held yet mistaken belief that police are not permitted to lie or fabricate incriminating evidence in interrogations.

While sympathetic to a ban on police deceit in interrogations more generally, as advocated by several scholars, Part III singles out deepfakes for particular concern, arguing that their especially compelling verisimilitude "critically impair[s]" the "essentially free and unconstrained choice" that the Supreme Court requires for a voluntary confession.[17] When police present a suspect with such a seemingly irrefutable form of incriminating evidence, they destroy, in Judge Richard Posner's words, "the information required for a rational choice" when making a confession.[18] Part III also discusses the major changes occurring in the criminal legal system since *Frazier v. Cupp* was decided in 1969, underscoring the need for its reconsideration. Most significant is the massively increased power of prosecutors to pressure individuals to engage in plea bargaining, a

---

deepfakes and expressing fear that "it will not be long before law enforcement tests the limits of creating fabricated images of a detainee at the scene of the crime or artificially create[s] other evidence in order to convince a suspect" to confess).

14. The focus here is on use of false evidence in securing interrogations, not in the courtroom, where the knowing introduction of false evidence by the government is prohibited. *See* Miller v. Pate, 386 U.S. 1, 7 (1967). Thanks to Brandon Garrett for suggesting that I emphasize this distinction at the outset.

15. 394 U.S. 731 (1969).

16. *See id.* at 739.

17. *See* Schneckloth v. Bustamonte, 412 U.S. 218, 225 (1973).

18. Aleman v. Village of Hanover Park, 662 F.3d 897, 906 (7th Cir. 2011).

system where securing confessions (from factually innocent and guilty individuals alike) has paramount importance.[19] Aggravating matters, discovery rules during the plea-bargaining stage typically preclude suspects from being able to learn that police used deceit in securing confessions, severely limiting suspects' wherewithal to contest charges.[20]

Part IV elaborates on why a per se ban on deepfakes is needed. In addition to reconciling confession law with post-*Frazier* social science findings and institutional changes, a ban will have important functional benefits. These include providing ex ante guidance to police, who lack clarity on the parameters of permissible interrogation techniques, and judges, who lack clarity in their application of the notoriously indeterminate due process voluntariness standard. More broadly, a ban will act as a partial yet important bulwark against the deleterious wave of disinformation now sowing distrust in governmental actors and institutions. In a world where deepfakes are condoned in interrogations, it is not hard to imagine that judges, jurors, witnesses, and members of the public will be skeptical of the reliability of evidence in criminal cases. The Article concludes with a discussion of how a ban can be achieved and why ameliorative tweaks to the current framework regulating confessions—motions to suppress, plea colloquies, trials, and appeals—are not up to the challenge of checking the formidable threats posed by police use of deepfakes in interrogations.[21]

I. DECEPTION AND ITS DOCTRINE

A. Deception by Police

Until the late 1930s, police regularly employed the "third degree,"[22] using threatened and actual physical abuse to obtain confessions from individuals.[23] Perhaps the most infamous case involving the practice was *Brown v. Mississippi*,[24] where police

19. *See infra* notes 186–92 and accompanying text.

20. *See infra* notes 213–18 and accompanying text.

21. Testament to the need for action, many police are aware of the challenges deepfakes present in investigations, especially with regard to sources of video frequently used as evidence, such as surveillance and doorbell cameras and cellphone videos. *See* Frederick Dauer, *Law Enforcement in the Era of Deepfakes*, POLICE CHIEF (June 29, 2022), https://perma.cc/U6AP-738U.

22. *See* 11 NAT'L COMM'N ON L. OBSERVANCE & ENF'T, REPORT ON LAWLESSNESS IN LAW ENFORCEMENT 5–6 (1931).

23. For accounts of police engaging in the third degree, by multiple methods, see ERNEST JEROME HOPKINS, OUR LAWLESS POLICE: A STUDY OF THE UNLAWFUL ENFORCEMENT OF THE LAW 189–263 (1931); Jeffrey S. Adler, *"The Greatest Thrill I Get Is When I Hear a Criminal Say, 'Yes, I Did It'": Race and the Third Degree in New Orleans, 1920–1945*, 34 LAW & HIST. REV. 1, 3 (2016).

24. 297 U.S. 278 (1936).

extracted confessions from several African American men after brutalizing them with repeated whippings with a metal-buckled strap and threatening to hang them from a tree limb.[25] Deciding the case at a time before the Fifth Amendment's prohibition of compelled self-incrimination was incorporated and applied to the states,[26] the Court relied on the Fourteenth Amendment's Due Process Clause to deem the confessions involuntary and therefore invalid.[27] Thereafter, in multiple cases, the Court condemned a variety of abusive practices of a less brutal nature, finding the confessions obtained invalid on due process grounds.[28]

Miranda v. Arizona,[29] best known for holding that the Fifth Amendment protection against compelled self-incrimination requires that police provide warnings and secure a waiver from suspects before conducting a custodial interrogation, noted a shift in police interrogation tactics.[30] By the mid-1960s, the *Miranda* Court recognized, police tactics had become "psychologically rather than physically oriented,"[31] as evidenced in then-popular police training manuals, which considered these tactics "the most enlightened and effective means presently used to obtain statements through custodial interrogation."[32]

Among these manuals was *Criminal Interrogation and Confessions*, compiled in 1963 by Fred Inbau and John Reid (the Reid Manual).[33] The Reid Manual, which remains the most widely used interrogation protocol in the United States,[34] employs the Reid Method of interrogation, whereby police

---

25. *Id.* at 281–82.

26. *See* Malloy v. Hogan, 378 U.S. 1, 6 (1964) (incorporating the bar on compelled self-incrimination under the Fifth Amendment).

27. *See Brown*, 297 U.S. at 287.

28. *See, e.g.*, Beecher v. Alabama, 389 U.S. 35, 35–36 (1967) (gun held to the head of already wounded suspect); Davis v. North Carolina, 384 U.S. 737, 746–52 (1966) (interrogated incommunicado for sixteen days without food in a closed cell without windows); Haynes v. Washington, 373 U.S. 503, 504–05 (1963) (threatened with indefinite detention); Reck v. Pate, 367 U.S. 433, 436–39 (1961) (held for four days with inadequate food and medical attention); Ashcraft v. Tennessee, 322 U.S. 143, 154 (1944) (interrogated for thirty-six hours without sleep); Ward v. Texas, 316 U.S. 547, 549–52 (1942) (subjected to slapping, beating, and burning).

29. 384 U.S. 436 (1966).

30. *See id.* at 498.

31. *Id.* at 448.

32. *Id.* at 449.

33. *See* FRED E. INBAU ET AL., CRIMINAL INTERROGATION AND CONFESSIONS 343 (5th ed. 2013) [hereinafter REID MANUAL].

34. Wyatt Kozinski, *The Reid Interrogation Technique and False Confessions: A Time for Change*, 16 SEATTLE J. SOC. JUST. 301, 301–02 (2018).

> first seek to break down suspects' confidence in their denials of guilt by causing them to perceive that they are caught, that no one will believe their assertions of innocence, and that resisting the interrogators' accusations is therefore futile. Second, interrogators seek to induce or incentivize suspects to believe that, given the available options for someone in their situation, it is in their best short-term and long-term interests to stop denying the interrogators' accusations and comply with the interrogators' demands and requests.[35]

Under the approach, "outright lies concerning the existence of evidence" are permitted.[36] Such deceit, Professor Saul Kassin has observed, is key to convincing the suspect that his refusal to confess would be fruitless because police possess overwhelming incriminating evidence that will ensure a conviction.[37] According to a Reid training

---

35. Richard A. Leo, *The Decision to Confess Falsely Twenty-Five Years Later: Windows and Walls in Empirical Psychological and Legal Scholarship*, 100 DENV. L. REV. 541, 542–43 (2023) (footnote omitted).

36. REID MANUAL, *supra* note 33, at 351. The manual adds that deceit is especially useful with individuals "vulnerable to the ploy." *Id.* at 246. With respect to the role of the FEP in extracting false confessions in particular, the manual states that "[i]t is our clear position that merely introducing fictitious evidence during an interrogation would not cause an innocent person to confess." *Id.* at 352. The fourth edition of the manual, published in 2001, affirmatively stated that "trickery and deceit . . . are not only helpful but frequently indispensable in order to secure incriminating information from the guilty" and recommended creation and use of "visual props" to secure confessions. *Id.* at xii, 217 (4th ed. 2001).

37. *See* Saul M. Kassin et al., *Police-Induced Confessions: Risk Factors and Recommendations*, 34 LAW & HUM. BEHAV. 3, 16–17 (2010) ("[O]nce people see an outcome as inevitable, cognitive and motivational forces conspire to promote their acceptance, compliance with, and even approval of the outcome."); *see also* State v. Griffin, 262 A.3d 44, 92 (Conn. 2021) (Ecker, J., concurring in part and dissenting in part) (stating that "the drumbeat theme of the Reid method is that resistance is futile and confession is the only rational choice").

Related to the FEP but distinct is the "bluff" tactic, whereby police suggest to a suspect the existence of incriminating evidence not yet reviewed. Jennifer T. Perillo & Saul M. Kassin, *Inside Interrogation: The Lie, the Bluff, and False Confessions*, 35 LAW & HUM. BEHAV. 327, 328 (2010). Employing the tactic, an interrogator might suggest that security camera footage exists that recorded the offense and, when examined, will incriminate the suspect. *See id.* A variant of the technique involves police falsely stating that unreviewed evidence exists but that they are less certain about its results. *Id.* Research suggests that the latter tactic is especially conducive to innocents confessing because they believe the unreviewed evidence will eventually exonerate them. *Id.* The bluff promotes confessions because it "decreases belief in [the] probability of conviction even if [innocents] confess" while promising conclusion of the stressful interrogation session. Jean J. Cabell et al., *Evaluating Effects on Guilty and Innocent Suspects: An Effect Taxonomy of Interrogation Techniques*, 26 PSYCH. PUB. POL'Y & L. 154, 158 tbl.3, 160, 163 (2020); *see also* Bryan Barnes et al., *The Influence of False*

seminar instructor, the key to dealing with suspects is "to shut them up" and "[n]ever allow them to give you denials."[38] Other how-to interrogation guides similarly advocate use of the FEP to secure confessions.[39]

Use of the FEP is complemented by other strategies. Police, for instance, will often seek to persuade a suspect that a confession is in their self-interest[40] and counter denials of guilt with evidence supporting guilt[41] and contentions that the suspect is lying.[42] At the same time, police might intimate that the suspect will have some sort

---

Evidence Ploy Variants on Perceptions of Coercion and Deception, 24 J. FORENSIC PSYCH. & PRAC. 500, 504, 515 (2024).

38. Douglas Starr, *The Interview: Do Police Interrogation Techniques Produce False Confessions?*, NEW YORKER (Dec. 1, 2013), https://perma.cc/7H4F-QEVY.

39. *See, e.g.*, DEVERE D. WOODS, JR., O HARA'S FUNDAMENTALS OF CRIMINAL INVESTIGATION 195–97 (9th ed. 2019). FEPs can be divided into three types: (1) demeanor (e.g., an officer saying "I can see the guilt on your face"); (2) testimonial (e.g., suggesting than an eyewitness or video surveillance implicates the suspect); and (3) scientific (e.g., false claims that police have DNA evidence implicating the suspect). Krista D. Forrest et al., *False-Evidence Ploys and Interrogations: Mock Jurors' Perceptions of False-Evidence Ploy Type, Deception, Coercion, and Justification*, 30 BEHAV. SCI. & L. 342, 343–44 (2012).

40. Richard Leo, *Police Interrogations, False Confessions, and Alleged Child Abuse Cases*, 50 U. MICH. J.L. REFORM 693, 703 (2017). *But see* DAVID SIMON, HOMICIDE: A YEAR ON THE KILLING STREETS 213 (1993) ("The fraud that claims it is somehow in a suspect's interest to talk with police will forever be the catalyst in any criminal interrogation. It is a fiction propped up against the greater weight of logic itself . . . .").

41. Kassin et al., *Police-Induced Confessions*, *supra* note 37, at 15.

42. Police often tell suspects that they are lying, based on body language, facial expressions, and speech, when in fact, police perform poorly as lie-detectors. *See, e.g.*, DAN SIMON, IN DOUBT: THE PSYCHOLOGY OF THE CRIMINAL JUSTICE PROCESS 126–37 (2012); Saul M. Kassin et al., *"I'd Know a False Confession If I Saw One": A Comparative Study of College Students and Police Investigators*, 29 LAW & HUM. BEHAV. 211, 211–12 (2005); Christine A. Meissner & Saul M. Kassin, *"He's Guilty!": Investigator Bias in Judgments of Truth and Deception*, 26 LAW & HUM. BEHAV. 469, 479 (2002); *cf.* Wayne A. Logan, *Policing Emotions: What Social Psychology Can Teach Fourth Amendment Doctrine*, 72 BUFF. L. REV. 685, 715 (2024) (discussing research showing the lack of human capacity to accurately discern human emotions, such as fear or surprise, from the facial expressions of others).

In turn, police, believing a suspect's denial is groundless, naturally put stock in an FEP-induced confession (whether true or false). *See, e.g.*, Moa Lidén et al., *The Presumption of Guilt in Suspect Interrogations: Apprehension as a Trigger of Confirmation Bias and Debiasing Techniques*, 42 LAW & HUM. BEHAV. 336, 337–38 (2018); Fadia M. Narchet et al., *Modeling the Influence of Investigator Bias on the Elicitation of True and False Confessions*, 35 LAW & HUM. BEHAV. 452, 462–63 (2011); Yueran Yang et al., *The Effect of a Presumption of Guilt on Police Guilt Judgments*, PSYCH. CRIME & L., Nov. 7, 2023, at 1, 1.

of defense if he or she confesses[43] or suggest that the suspect will get to go home if incriminating information is provided.[44] They will also seek to downplay the perceived costs of confession or to signal empathy and understanding of why the interrogation subject would engage in the criminal act.[45] Rounding matters out, if *Miranda* does not apply—because of waiver or because the *Miranda* preconditions of custody and/or interrogation are not satisfied[46]—police can secure confessions without benefit of counsel being present (which police prefer),[47] and the prosecution will likely have no constitutional duty to disclose during interrogations (i.e., pre-charge) any information that might exculpate the individual.[48]

FEPs in turn dovetail, indeed facilitate, what Professor Anne Coughlin has called the strategic goal of interrogators to construct a narrative of a suspect's involvement in a crime.[49] She observes, based on her review of interrogation and trial transcripts:

> [T]he cop is not merely finding but creating, not merely reconstructing but constructing, the solution to the crime. The interrogator is master narrator or, maybe, improvisational playwright, one who is comfortable batting around potential plot lines, as well as pinning down specific bits of dialogue, with

---

43. REID MANUAL, *supra* note 33, at 345 (stating that an interrogator may say to a suspect, "if this is something that happened on the spur of the moment, that would be important to include in my report").

44. RICHARD A. LEO, POLICE INTERROGATION AND AMERICAN CRIMINAL JUSTICE 132 (2008) (providing examples).

45. Saul M. Kassin, *The Psychology of Confession Evidence*, 52 AM. PSYCH. 221, 221–23 (1997); Allison D. Redlich & Kyle Scherr, *Coercion in Interrogations*, CHAMPION, April 2023, at 39, https://www.nacdl.org/Article/MarApr2023-CoercioninInterrogations.

46. *See, e.g.*, Shedelbower v. Estelle, 885 F.2d 570, 573–75 (9th Cir. 1989) (concluding that FEP by officer, resulting in a confession by a suspect who had invoked his right to counsel, did not qualify as an "interrogation" and therefore did not violate *Miranda*).

47. *See* Christopher Slobogin, *Manipulation of Suspects and Unrecorded Questioning: After Fifty Years of* Miranda *Jurisprudence, Still Two (Or Maybe Three) Burning Issues*, 97 B.U. L. REV. 1157, 1171 (saying of "interrogation room negotiation[s]" that "the implicit or explicit message [is] that if counsel is consulted, the deal is off the table"); *see also* George E. Dix, *Promises, Confessions, and Wayne LaFave's Bright Line Rule Analysis*, 1993 U. ILL. L. REV. 207, 247 (1993) ("[O]fficers may be willing to 'deal' with a suspect only if the suspect is willing to deal immediately . . . without consulting counsel . . . .").

48. *See, e.g.*, State v. Harris, 105 P.3d 1258, 1265–66 (Kan. 2005) ("Although the State has a duty to reveal any exculpatory evidence to the defense before trial . . . our research fails to reveal any case law supporting Harris' theory that the same rule extends to police interrogation before the defendant has been charged with any crime.").

49. Anne M. Coughlin, *Interrogation Stories*, 95 VA. L. REV. 1599, 1603 (2009).

> his leading actors before getting them to sign off on the final script.[50]

And:

> [P]olice interrogators do not merely find facts that are buried out there somewhere—this time, deep in the mind and heart of a suspect—just waiting for the alert detective to come along and excavate them. Rather, by using interrogation stories, interrogators actively and inescapably shape the meaning of the facts by helping suspects to embed them in a coherent narrative that coincides with our normative judgments about which acts are blameworthy and which are not. Moreover, once the suspect endorses one of the plots the cops offer, the interrogation story, now an offender's confessional speech act, itself has the potential to become the past.[51]

That police should go to such lengths to secure a confession is understandable. Long regarded as the "queen of evidence,"[52] confessions are "like no other evidence"[53] and "can render other evidence "superfluous."[54] As the Supreme Court has recognized,

> The defendant's own confession is probably the most probative and damaging evidence that can be admitted against him . . . . [T]he admissions of a defendant come from the actor himself, the most knowledgeable and unimpeachable source of information about his past conduct. Certainly, confessions have profound impact on the jury, so much so that we may justifiably doubt its ability to put them out of mind even if told to do so.[55]

Research also shows that confessions are difficult if not impossible to retract without consequence. Among police, already

---

50. *Id.* at 1608.

51. *Id.* at 1600, 1609; *see also* Garrett, *supra* note 2, at 1067 ("Police are trained to construct a narrative of how the crime occurred, including the motives for committing the crime and a detailed explanation of how it was committed.").

52. *See* Talia Fishera & Issachar Rosen-Zvi, *The Confessional Penalty*, 30 CARDOZO L. REV. 871, 872 (2008); *see also* Kassin, *supra* note 2, at 222 (showing that mock jury studies have demonstrated that confessions are more potent than eyewitness and other forms of evidence and that juries do not fully discount confessions even when they are supposed to do so).

53. Arizona v. Fulminante, 499 U.S. 279, 296 (1991).

54. CHARLES MCCORMICK, HANDBOOK ON THE LAW OF EVIDENCE 312 (2d ed. 1983); *see also* Saul M. Kassin & Katherine Neumann, *On the Power of Confession Evidence: An Experimental Test of the Fundamental Difference Hypothesis*, 21 LAW & HUM. BEHAV. 469, 479, 481 (1997) (finding that confessions are more prejudicial to the defendant's case than eyewitness identification and character testimony).

55. *Fulminante*, 499 U.S. at 296 (alteration in original) (quoting Bruton v. United States, 391 U.S. 123, 139–40 (1968) (White, J., dissenting)).

subject to the Reid Manual's "guilt-presumptive" mindset[56] and the belief that an innocent party would not confess,[57] a confession very often drives the belief that a case is "solved," discouraging continued investigation into other possible suspects and other factual scenarios.[58] Among prosecutors, who are likewise disinclined to believe that an individual would falsely confess,[59] confessions drive "investigative tunnel vision and confirmation bias (e.g., seeking out evidence that points to a suspect's guilt while discounting evidence that points away from it)."[60]

A confession, moreover, can have downstream influence on forensic examiners, who, consciously or not, can align their results with what the government regards as a case-dispositive confession.[61] Confessions can also impact the testimony of witnesses, prompting a shift in their testimony.[62] In short, as social psychologist Saul Kassin observes, a confession often fosters "corroboration inflation,"[63] creating a situation in which "confession begets conviction."[64]

---

56. *See* Kassin et al., *Police-Induced Confessions*, *supra* note 37, at 27, and accompanying text; *see also, e.g.*, Meissner & Kassin, *"He's Guilty!," supra* note 42, at 477–78. *But see* Miller v. Fenton, 474 U.S. 104, 116 (1985) (stating that due process analysis considers whether the tactics employed to elicit a confession are "compatible with a system that presumes innocence").

57. *See supra* note 36.

58. *Cf.* Eyal Press, *Does A.I. Lead Police to Ignore Contradictory Evidence?*, NEW YORKER (Nov. 13, 2023), https://perma.cc/C4B3-K3YM (discussing this effect with respect to facial-recognition software used by police).

59. *See* Steven A. Drizin & Richard A. Leo, *The Problem of False Confessions in the Post-DNA World*, 82 N.C. L. REV. 891, 910 (2004) (noting that prosecutors routinely assume that "an innocent person would not falsely confess to a serious crime unless he is physically tortured or mentally ill").

60. Samantha Luna, *Defining Coercion: An Application in Interrogation and Plea Negotiation Contexts*, 28 PSYCH. PUB. POL'Y & L. 240, 241 (2022); *see also* SAUL KASSIN, DUPED: WHY INNOCENT PEOPLE CONFESS—AND WHY WE BELIEVE THEIR CONFESSIONS 138 (2022) (author, a leading social psychologist who has worked on multiple false confession cases, stating, "I've seen time and again prosecutors adhere to a confession as an act of faith in the face of overwhelming contradictory evidence that unequivocally excluded the confessor."). On confirmation bias more generally, see MOA LIDÉN, CONFIRMATION BIAS IN CRIMINAL CASES (2023).

61. *See* Saul M. Kassin et al., *Forensic Confirmation Bias: Problems, Perspectives, and Proposed Solutions*, 2 J. APPLIED RSCH. MEMORY & COGNITION 42, 44 (2013). *See generally* Glenda S. Cooper & Vanessa Meterko, *Cognitive Bias Research in Forensic Science: A Systematic Review*, 297 FORENSIC SCI. INT'L 35 (2019). The phenomenon is especially problematic given ongoing concern over the accuracy and reliability of forensic laboratory work. *See, e.g.*, John Morgan, *Wrongful Convictions and Claims of False or Misleading Forensic Evidence*, 68 J. FORENSIC SCIS. 908 (2023).

62. KASSIN, *supra* note 60, at 201–03.

63. *Id.* at 270.

64. *Id.* at 274.

B. The Judicial Response

Whether a confession was secured in violation of due process depends on whether it was "voluntary," not coerced by police,[65] based on the totality of the circumstances.[66] Voluntariness turns on whether

> the confession [was] the product of an essentially free and unconstrained choice by its maker[.] If it is, if he has willed to confess, it may be used against him. If it is not, if his will has been overborne and his capacity for self-determination critically impaired, the use of his confession offends due process.[67]

In short, when assessing whether police violated due process when securing a confession, courts must assess whether a suspect's "will was overborne" by police.[68]

As noted earlier, the Supreme Court in *Frazier v. Cupp* refused to hold that police trickery—in *Frazier*, a lie that another individual

---

65. *See* Colorado v. Connelly, 479 U.S. 157, 167 (1986) (stating that "coercive police activity is a necessary predicate to the finding that a confession is not 'voluntary' within the meaning of the Due Process Clause of the Fourteenth Amendment"); *see also* United States v. Bell, 367 F.3d 452, 461 (5th Cir. 2004) (holding that the defendant "must demonstrate that [the confession] resulted from coercive police conduct and that there was a link between the coercive conduct of the police and his confession"); United States v. Lugo, 170 F.3d 996, 1004 (10th Cir. 1999) ("[A] confession is only involuntary when 'the police use coercive activity to undermine the suspect's ability to exercise his free will.'" (quoting United States v. Erving L., 147 F.3d 1240, 1249 (10th Cir. 1998))).

66. Frazier v. Cupp, 394 U.S. 731, 739 (1969); *see also* United States v. Haak, 884 F.3d 400, 409 (2d Cir. 2018) (stating that the inquiry considers "(1) the characteristics of the accused, (2) the conditions of interrogation, and (3) the conduct of law enforcement officials" (quoting Green v. Scully, 850 F.2d 894, 901–02 (2d Cir. 1988))). *See generally* 2 WAYNE R. LAFAVE ET AL., CRIMINAL PROCEDURE § 6.2(c), at 686–721 (4th ed. 2015).

67. Culombe v. Connecticut, 367 U.S. 568, 602 (1961). Whether any confession can be considered truly voluntary has long been questioned. As Judge Richard Posner noted:

> The courts in such cases retreat to the proposition that a confession, to be admissible, must be the product of a free choice . . . . [But this] leads nowhere. Taken seriously it would require the exclusion of virtually all fruits of custodial interrogation, since few choices to confess can be thought truly "free" when made by a person who is incarcerated and is being questioned by armed officers without the presence of counsel or anyone else to give him moral support. The formula is not taken seriously.

United States v. Rutledge, 900 F.2d 1127, 1129 (7th Cir. 1990) (citations omitted).

68. Arizona v. Fulminante, 499 U.S. 279, 288 (1991); *see also* Schneckloth v. Bustamonte, 412 U.S. 218, 218, 225–26 (1973) (quoting *Culombe*, 367 U.S. at 602).

implicated the defendant—renders a confession involuntary.[69] According to the Court, "[t]he fact that the police misrepresented the statements [a codefendant] had made is, while relevant, insufficient in our view to make this otherwise voluntary confession inadmissible."[70] Lower state and federal courts have applied and expanded on the Court's tolerance of deceptive police practices for inducing confessions. While courts at times express displeasure with police utilizing the FEP,[71] they typically condone police lies about the existence of incriminating physical evidence,[72] including surveillance video,[73] forensic reports, or other incriminating documents,[74] as well

69. *Frazier*, 394 U.S. at 737. In *Frazier*, police complemented their lie with other strategies, including minimization by attributing the homicide to the victim's supposed "homosexual advances" toward the defendant. *Id.* at 738.

70. *Id.* at 739; *cf.* Oregon v. Mathiason, 429 U.S. 492, 495–96 (1977) (holding that an "officer's false statement about having discovered [the defendant's] fingerprints at the scene" of the crime had "nothing to do with whether [he] was in custody for purposes of the *Miranda* rule"). In so deciding, the Court distanced itself from earlier statements categorically condemning police deceit in interrogations. *See, e.g.*, Lisenba v. California, 314 U.S. 219, 237 (1941) ("If, by fraud, collusion, trickery, and subornation of perjury, on the part of those representing the State, the trial of an accused person results in his conviction, he has been denied due process of law.").

71. *See, e.g.*, United States v. Orso, 266 F.3d 1030, 1039 (9th Cir. 2001) ("[M]isrepresenting a piece of the evidence . . . [is] reprehensible" but deeming confession voluntary); State v. Register, 476 S.E.2d 153, 158 (S.C. 1996) (calling falsely telling the suspect multiple untruths regarding eyewitness and physical evidence "deplorable" but concluding that the suspect's will was not overborne); State v. Von Dohlen, 471 S.E.2d 689, 694 (S.C. 1996) (calling police creation and presentation to suspect of a fake "composite sketch" to make him think police had an eyewitness to offense "reprehensible" and "deplorable" but finding no due process violation).

72. *See, e.g.*, Lucero v. Kerby, 133 F.3d 1299, 1311 (10th Cir. 1998) (falsely informing the defendant that his fingerprints found at the victim's home); Holland v. McGinnis, 963 F.2d 1044, 1051 (7th Cir. 1992) (lying that a witness had seen the suspect at the crime scene); Sotelo v. Ind. State Prison, 850 F.2d 1244, 1248–49 (7th Cir. 1988) (falsely telling suspect that the results of a polygraph indicated that he was lying about his innocence); *Ex parte* Jackson, 836 So. 2d 979, 984–85 (Ala. 2002) (falsely telling suspect his fingerprints were found on item tying him to alleged accomplice); State v. Nightingale, 58 A.3d 1057, 1061, 1069–70 (Me. 2012) (falsely suggesting that police had compromising DNA evidence and satellite images of the suspect's car at the crime scene). *But see* State v. Baker, 465 P.3d 860, 879 (Haw. 2020) (finding impermissible coercion in a case where police made repeated "misrepresentations about the existence of incontrovertible physical evidence that directly implicate[d] the accused").

73. Goodwin v. State, 281 S.W.3d 258, 266 (Ark. 2008) (falsely telling suspect the crime was caught on surveillance videotape); Johnson v. State, 713 S.E.2d 376, 379 (Ga. 2011) (same); Commonwealth v. Neves, 50 N.E.3d 428, 436–37 (Mass. 2016); State v. Brown, 796 N.E.2d 506, 512 (Ohio 2003).

74. *See, e.g.*, People v. Mays, 95 Cal. Rptr. 3d 219, 227–29 (Ct. App. 2009) (fake graph "fabricat[ed]" from a fake polygraph); Arthur v. Commonwealth, 480

as information provided by victims and witnesses[75] and statements by suspected confederates implicating the defendant.[76] Courts also typically condone police presentation of items they fabricated, eschewing any distinction between oral representations of

S.E.2d 749, 752 (Va. Ct. App. 1997) ("'dummy' reports" showing that a fingerprint and hair found at the scene of a murder matched defendant's); Sheriff, Washoe Cnty. v. Bessey, 914 P.2d 618, 619 (Nev. 1996) (showing defendant a fabricated forensic report).

75. *See, e.g.*, *Holland*, 963 F.2d at 1051 (falsely telling suspect that witness saw his car at the crime scene); Luckhart v. State, 736 N.E.2d 227, 230–31 (Ind. 2000) (falsely telling suspect two witnesses placed her at the crime scene); Rodriquez v. State, 934 S.W.2d 881, 890–91 (Tex. Ct. App. 1996) (falsely telling suspect that the victim identified him as perpetrator); Farmah v. State, 789 S.W.2d 665, 671–72 (Tex. Ct. App. 1990) (same), *rev'd on other grounds*, 883 S.W.2d 674 (Tex. Crim. App. 1994).

FEPs in this context might include "non-eyewitness identifications" where a witness does not observe a crime but only views a video and later identifies the suspected perpetrator. Tamar Lerer & Kathy Pezdek, *The Witness Who Was Not There: Challenging the Reliability of Identifications Made from Images*, CHAMPION, May 2024, at 12, https://perma.cc/FJS2-ZF7Y. In such instances, the legal system imposes far fewer and less demanding limits and requirements than with conventional eyewitness testimony. *Id.* "If the non-eyewitness claims almost any degree of familiarity with the perpetrator, that identification will be admissible. The judge will not consider any suggestiveness in the circumstances leading to the identification or the risk of an irreparable misidentification. And the jury will not be told about these risks." *Id.*

76. *See, e.g.*, United States v. Lux, 905 F.2d 1379, 1382 (10th Cir. 1990) (falsely telling suspect that her codefendant had implicated her in killing); United States v. Velasquez, 885 F.2d 1076, 1088–89 (3d Cir. 1989) (falsely telling suspect that her alleged accomplice had been released after making statements against her and these false statements made the government's evidence look much stronger than it actually was); United States v. Castaneda-Castaneda, 729 F.2d 1360, 1362–63 (11th Cir. 1984) (falsely telling husband the codefendant wife had confessed); State v. Galli, 967 P.2d 930, 936 (Utah 1998) (falsely telling suspect all three accomplices implicated him when only one had); State v. Thaggard, 527 N.W.2d 804, 806, 810 (Minn. 1995) (falsely telling suspect that accomplice had confessed).

incriminating evidence and the presentation of physical items containing the false information,[77] permitting the latter.[78]

Several courts, however, have held that police creation and presentation of fabricated physical evidence renders a confession involuntary and therefore a violation of due process. In the leading case, *State v. Cayward*,[79] a Florida appellate court condemned police creation and use of two scientific reports falsely stating the presence of the defendant's biological material at the crime scene.[80] One report was on stationery of the Florida Department of Criminal Law Enforcement, and the other was on stationery of Life Codes, Inc., a scientific testing organization.[81] The reports stated, falsely, that the defendant's semen had been found on the victim's underwear.[82] The defendant, who voluntarily came to the police station for an interview and waived his *Miranda* rights,[83] repeatedly denied involvement in the crime until he was confronted with the fake reports, after which he confessed.[84]

The *Cayward* court acknowledged that oral representation of fabricated incriminating evidence does not render a confession involuntary per se but held that there is an "intrinsic distinction" with the creation and presentation of fabricated physical documents containing the information, which warranted a "bright-line" rule of prohibition.[85] The court recognized a "qualitative difference

77. *See, e.g.*, Lincoln v. State, 882 A.2d 944, 956–57 (Md. Ct. Spec. App. 2006) ("It is a simplistic generality that a written false assertion by the police, regardless of its substance, always will have a greater impact on a suspect's thinking than an oral assertion, and that every written false assertion by the police will have precisely the same coercive effect as all other false written assertions by the police."); *Bessey*, 914 P.2d at 621 (averring that police can lie about the existence of an incriminating report and that the outcome should be the same when the information is rendered in tangible form, stating that "[t]his is a distinction without a real difference"). *But see infra* notes 113–25 and accompanying text (discussing research demonstrating the more powerful impact of showing evidence, as opposed to merely stating that it exists).

78. *See, e.g.*, United States v. Carona, 660 F.3d 360, 365 (9th Cir. 2011) (fake subpoena attachments provided by prosecutors); United States v. Haynes, 26 F. App'x 123, 129, 134 (4th Cir. 2001) (per curiam) (fabricated ballistics report and boxes purportedly containing evidence in defendant's case but actually containing unrelated materials); State v. Von Dohlen, 471 S.E.2d 689, 694 (S.C. 1996) (purported composite sketch of defendant designed to have defendant believe that police had an eyewitness, when in reality the sketch was drawn by a police artist through a one-way window); *Arthur*, 480 S.E.2d at 751–52 (false lab reports).

79. 552 So. 2d 971 (Fla. Dist. Ct. App. 1989).

80. *Id.* at 972, 975.

81. *Id.* at 972.

82. *Id.*

83. *Id.*

84. *Id.*

85. *Id.* at 973–74.

between . . . verbal artifices . . . and the presentation of . . . falsely contrived scientific documents."[86] A "suspect [may be] . . . more easily induced to confess when presented with tangible, official-looking reports as opposed to merely being told that some tests have implicated him."[87] The court also expressed "distaste" for the practice, comparing it to "the horrors of less advanced centuries in our civilization,"[88] and had "practical concerns," including that manufactured documents, with their "potential of indefinite life and the facial appearance of authenticity," might be used in later legal proceedings or be reproduced and remain in disparate government files, causing "dangerous confusion," and possibly "be disclosed to the media."[89] Finally, the court reasoned that condoning fabrication and use of physical evidence by law enforcement would damage trust in police[90] and that "the type of deception engaged in here has no place in our criminal justice system."[91]

A New Jersey appellate court reached the same result in *State v. Patton*.[92] There, police created and used a fabricated audiotape of an eyewitness claiming to have seen the defendant perpetrate a killing.[93] The "eyewitness" (actually an officer) recounted the version of events the police thought transpired.[94] To lend authenticity to the fabricated statement on the audiotape, the recording also contained true information about the defendant, including about his prior misconduct, some targeting the victim.[95] When the defendant heard

---

86. *Id.* at 973.

87. *Id.* at 974.

88. *Id.* In full, the court stated:

> We think . . . that both the suspect's and the public's expectations concerning the built-in adversariness of police interrogations do not encompass the notion that the police will knowingly fabricate tangible documentation or physical evidence against an individual. Such an idea brings to mind the horrors of less advanced centuries in our civilization when magistrates at times schemed with sovereigns to frame political rivals. This is precisely one of the parade of horrors civics teachers have long taught their pupils that our modern judicial system was designed to correct.

Id.

89. *Id.*

90. *Id.* at 975.

91. *Id.* at 974; *see also* State v. Farley, 452 S.E.2d 50, 60 n.13 (W. Va. 1994) (commenting in dictum that the court would "definitely draw a demarcating line between police deception generally, which does not render a confession involuntary per se, and the manufacturing of false documents by the police, which 'has no place in our criminal justice system'" (quoting *Cayward*, 552 So. 2d at 974)).

92. 826 A.2d 783 (N.J. Ct. App. 2003).

93. *Id.* at 785.

94. *Id.* at 785–89.

95. *Id.* at 787–88.

the tape, he confessed.[96] The fake audiotape was introduced as evidence at trial, and the defendant was convicted.[97] The *Patton* court concluded that the confession was involuntary and reversed the conviction,[98] like *Cayward* adopting a "bright-line" prohibition,[99] going so far as to equate the use of "police-fabricated tangible evidence" with actual physical coercion.[100] Also, like *Cayward*, the court expressed concern that the ploy undermined the legitimacy of the legal system.[101]

More recently, the Kentucky Supreme Court in *Gray v. Commonwealth*[102] condemned police use of FEPs but refused to impose a bright-line prohibition when police employed several strategies, including the creation and presentation to a suspect of a fake document purporting to confirm the existence of DNA evidence incriminating the defendant (as in *Cayward*).[103] However, concerned about the "severe" risk of coercion in such a situation, the court held that a judge is to "presume [that] this tactic is unconstitutional until the Commonwealth can firmly establish that the document[] did not overwhelm the defendant's will and was not a critical factor in the defendant's decision to confess."[104]

According to the *Gray* court, when assessing whether will is overborne, the question is "whether the tactics employed by police would overwhelm the will of an ordinary defendant,"[105] based on a "Rational Choice Model," which accepts that "many criminal defendants choose to confess when faced with an abundance of evidence against them as the more economical solution."[106] While the

96. *Id.* at 789.

97. *Id.* at 784. The falsity of the audiotape was conceded by the parties and the jury "was told that the police created the audiotape to persuade defendant to confess." *Id.* at 799. The trial court was troubled by the tape's admission and provided the jury a limiting instruction, which the court of appeals found inadequate. *Id.* The primary focus of the court, however, was on the coerciveness of the audiotape. *Id.* at 800.

98. *Id.* at 802.

99. *See id.* at 804 ("We hold that law enforcement and the public would best be served by a 'bright-line' rule . . . . Such 'bright-line' rules serve to protect the constitutional rights of suspects while providing a clear procedure for police to follow that should produce consistent results.").

100. *Id.* at 805; *see also* State v. Chirokovskcic, 860 A.2d 986, 989, 991 (N.J. Super. Ct. App. Div. 2004) (affirming *Patton*'s "bright-line rule" barring fabricated tangible evidence, deeming involuntary a statement obtained when detectives used a fabricated laboratory report purporting to show that defendant's and victim's DNA had both been found on a pair of gloves).

101. *Patton*, 826 A.2d at 800.

102. 480 S.W.3d 253 (Ky. 2016).

103. *Id.* at 262.

104. *Id.* at 263.

105. *Id.*

106. *Id.*

court was especially concerned about the impact on individuals of fake DNA evidence,[107] it was also troubled by other police tactics, including the creation and use of video footage falsely placing the defendant at the crime scene.[108] "When faced with seemingly insurmountable evidence," the court stated, "it becomes reasonable for one to perceive the futility of maintaining innocence. Indeed, facing overwhelming documentary and verbal forensic evidence, we think an average defendant in Gray's situation would feel pressured to confess to the point that it usurps free will."[109]

Concluding, the court found that the fabricated evidence (a DNA report) played a causal role in the defendant's confession and rejected the government's effort to minimize the impact of the fabricated report, which it had characterized as having an "amateurish quality."[110] The court reasoned that the "operative assumption should not be an expectation that citizens should distrust everything law enforcement tells them or shows them."[111] Rather, "[o]rdinarily, when a police officer presents a lab report purporting to represent DNA evidence of criminality, one likely does not carefully examine the contents for detailed accuracy."[112]

---

107. *See id.* ("Given the evidentiary power DNA and forensic evidence enjoy[] in the minds of jurors, it is reasonable to conclude that documents containing incriminating scientific evidence would similarly cause the ordinary criminal defendant to consider maintaining his innocence a futile endeavor.").

108. *Id.* at 263–64.

109. *Id.* at 264.

110. *Id.*

111. *Id.*

112. *Id.* In other cases, courts have condemned police use of fabricated tangible evidence not shown to suspects. In *Commonwealth v. DiGiambattista*, 813 N.E.2d 516 (Mass. 2004), for instance, the Massachusetts Supreme Court condemned use of two videotapes with false and misleading external labels suggesting the suspect's connection to the crime (arson), complemented by police insinuation of his involvement. *Id.* at 523–24. After noting "that ongoing research has identified such use of false statements as a significant factor that pressures suspects into waiving their rights and making a confession," *id.* at 524, the court stated,

> [t]his is particularly true where, as here, the false statement suggests a form of incriminating evidence that would be viewed as incontrovertible. If a suspect is told that he appears on a surveillance tape, or that his fingerprints or DNA have been found, even an innocent person would perceive that he or she is in grave danger of wrongful prosecution and erroneous conviction.

Id. at 525. The court added that "police resort to false statements concerning ostensibly irrefutable evidence of guilt [is] a tactic that 'casts instant doubt' on the voluntariness of [a] subsequent confession" but also relied on what the court considered excessive "minimization" techniques employed by the police in deeming the defendant's confession coerced. *Id.* at 527–28; *see also* Robinson v. Smith, 451 F. Supp. 1278, 1291 (W.D.N.Y. 1978) (barring confession on due process grounds, stating that police "foist[ing]" of fake signed confession by a

II. SOCIAL SCIENCE FINDINGS

As discussed, courts are divided on the constitutionality of confessions resulting from police presentation of fabricated tangible evidence to interrogation subjects, with several courts deeming the practice coercive and therefore violative of due process. This Part explores the growing body of research supporting the view of these courts that the FEP interrogation strategy has uniquely coercive effect on interrogated individuals. Research, moreover, demonstrates that police use of FEPs has particular impact on factually innocent individuals, prompting them to confess to crimes that they did not commit. Finally, a growing body of research shows that deepfakes present unique challenges to both individuals and the adjudicatory process, heightening the importance of banning deepfakes in interrogations.

A. Research on the Impact of FEPs

A large body of research demonstrates that the presentation of fabricated physical evidence has greater psychological and emotional impact on individuals than oral communication of the existence of the same evidence. For humans, seeing is believing.[113] We take images at face value,[114] even if they are known to have been manipulated.[115] The persuasive impact is even greater with audiovisual presentations, again even when it is known that they can misrepresent facts.[116]

The impact of being shown fake evidence has been demonstrated in laboratory and real-world settings alike. In one of the best-known

---

purported accomplice was "soundly condemned" and that "[s]uch deception clearly has no place in our system of justice"); *cf.* State v. Grey, 907 P.2d 951, 955 (Mont. 1995) (holding confession coerced in part because police placed a video camera in the store where the crime took place to create the false impression that the crime had been recorded).

113. Robert A. Nash & Kimberley A. Wade, *Innocent but Proven Guilty: Eliciting Internalized False Confessions Using Doctored-Video Evidence*, 23 APPLIED COGNITIVE PSYCH. 624, 625 (2009) (recognizing that "seeing is believing: in both legal and everyday decision-making tasks people are more persuaded by visual than verbal evidence," citing studies in support); *see also* Richard K. Sherwin et al., *Law in the Digital Age: How Visual Communication Technologies Are Transforming the Practice, Theory, and Teaching of Law*, 12 B.U. J. SCI. & TECH. L. 227, 246 (2006); Carolyn Purnell, *Do We All Still Agree That "Seeing Is Believing"?*, PSYCH. TODAY (June 23, 2020), https://perma.cc/62D7-S8DR. As Susan Sontag long ago noted with respect to still photographs, they "make a claim to be true," and showing "[a] fake photograph (one which has been retouched or tampered with, or whose caption is false) falsifies reality." SUSAN SONTAG, *The Heroism of Vision*, *in* ON PHOTOGRAPHY 86 (2005) (1977).

114. *See generally* Yael Granot et al., *In the Eyes of the Law: Perception Versus Reality in Appraisals of Video Evidence*, 24 PSYCH. PUB. POL'Y & L. 93, 97 (2017).

115. *Id.*; Newman & Schwarz, *supra* note 12, at 2.

116. *See* Granot et al., *supra* note 114, at 97–98.

laboratory studies, researchers examined the impact video has on reconstructing personal observations.[117] In the study, researchers placed sixty college students in a room to engage in a "computerized gambling task."[118] Following completion of the task, researchers individually showed each subject a digitally altered video falsely depicting a co-subject cheating.[119] Nearly half of the subjects were willing to testify that they had personally witnessed a co-subject cheating after seeing the fake video; only one in ten was willing to testify to the same effect after the researcher merely told the subject about the cheating.[120] Multiple other studies report similar results.[121]

Based on his observation of actual interrogations, Professor Richard Leo concluded that in 30% of the cases, police confronted suspects with false evidence to try to convince them to confess[122]—involving what Leo termed "negative incentives," "tactics that suggest the suspect should confess because no other course of action is plausible."[123] Leo found that the presentation of false evidence was more effective in securing confessions than the use of truthful evidence.[124] Confessions were secured in 78% of cases when suspects were told of actual evidence but in 83% of the cases where the police lied about the existence of incriminating evidence.[125]

Research also shows that police FEPs are more likely to be successful when the suspect being interrogated is especially vulnerable to manipulation due to factors such as youth, interpersonal trust, naiveté, suggestibility, lack of intelligence, stress, anxiety, fatigue, and alcohol or drug use.[126]

---

117. Kimberley A. Wade et al., *Can Fabricated Evidence Induce False Eyewitness Testimony?*, 24 APPLIED COGNITIVE PSYCH. 899, 900 (2010).

118. *Id.* at 900–01.

119. *Id.* at 903.

120. *Id.* at 905.

121. *See, e.g.*, Richard A. Leo & Brittany Liu, *What Do Potential Jurors Know About Police Interrogation Techniques and False Confessions?*, 27 BEHAV. SCI. & L. 381, 394, 397 (2009); *cf.* Forrest et al., *supra* note 39, at 358 (reporting results of study of mock jurors finding that "testimonial" FEPs, such as falsely telling a suspect that an eyewitness exists, is viewed as more coercive than false "demeanor" and "scientific" evidence).

122. Richard A. Leo, *Inside the Interrogation Room*, 86 J. CRIM. L. & CRIMINOLOGY 266, 278 tbl.5 (1996).

123. *Id.* at 300.

124. *Id.* at 294 tbl.14.

125. *Id.*

126. *See* Kassin et al., *Police-Induced Confessions*, *supra* note 37, at 15–17, 29–30; *see also* R. Kent Greenawalt, *Silence as a Moral and Constitutional Right*, 23 WM. & MARY L. REV. 15, 41 (1981) (noting that deceptive interrogation tactics "work unevenly by undermining the inexperienced and ignorant [while] having little effect on the hardened criminal").

B. FEPs and False Confessions

Although falsely confessing to a crime is commonly believed to be something that no rational person would ever do,[127] and both police[128] and the Reid Manual[129] think innocent individuals will not falsely confess, we now know that individuals do indeed confess to crimes they did not actually commit.[130]

Research shows that false confessions have contributed to 13% (460) of the 3,651 wrongful convictions documented by the National Registry of Exonerations, including 23% (333) of the wrongful convictions for homicide.[131] Importantly, FEPs were used by police in the vast majority of false confession cases resulting in exonerations.[132] In his 2022 book *Duped: Why Innocent People Confess—and Why We Believe Their Confessions*, Professor Kassin notes eighteen cases in which he was personally involved where police

---

127. *See* Saul M. Kassin, *False Confessions: How Can Psychology So Basic Be So Counterintuitive?*, 72 AM. PSYCH. 951, 952 (2017) ("The notion that anyone of sound mind would confess to a crime he or she did not commit is not intuitive to the average person.").

128. Kassin et al., *Police-Induced Confessions, supra* note 37, at 29.

129. *See* REID MANUAL, *supra* note 33, at 352 (taking a "clear position" that "merely introducing fictitious evidence during an interrogation" cannot lead to false admissions of guilt); *id.* ("It is absurd to believe that a suspect who knows he did not commit a crime would place greater weight and credibility on alleged evidence than his own knowledge of innocence.").

130. *See* Corley v. United States, 556 U.S. 303, 320–21 (2009) (stating that there is mounting empirical evidence that the pressures inherent in custodial interrogations can induce a "frighteningly high percentage of people to confess to crimes they never committed"). Professor Kassin identities three types of false confessions: (1) "voluntary false confessions," in which "people claim responsibility for crimes they did not commit without prompting or pressure from police"; (2) "compliant false confessions," where the "suspect capitulates to escape a stressful in-custody situation, avoid physical harm or legal punishment, or gain a promised or implied reward"; and (3) "internalized false confessions," where "innocent but psychologically vulnerable suspects not only agree to confess as an act of compliance but become confused, lose their grip on reality, and come to believe that they committed the crime in question." KASSIN, *supra* note 60, at 12–13.

131. *% Exonerations by Contributing Factor*, NAT'L REGISTRY EXONERATIONS (2025), https://perma.cc/A8XM-HQVJ. On the occurrence of wrongful convictions from earlier eras, see EDWIN M. BORCHARD, CONVICTING THE INNOCENT: ERRORS OF CRIMINAL JUSTICE, at vi (1932).

132. Kassin et al., *Police-Induced Confessions*, *supra* note 37, at 12; *see also* Deborah Davis & Richard A. Leo, *The Problem of Interrogation-Induced False Confession: Sources of Failure in Prevention and Detection*, *in* HANDBOOK OF FORENSIC SOCIOLOGY AND PSYCHOLOGY 47, 59 (Stephen J. Morewitz & Mark L. Goldstein eds., 2014) (recognizing that "[t]he practice of lying about evidence has been strongly implicated as a cause of real-life false confessions and the conviction of the innocent").

use of the FEP resulted in false confessions.[133] The false confessions of the "Central Park Five," in New York City, are perhaps the best known and most infamous of other more recent FEP cases.[134] According to Professor Leo, "[f]alse evidence ploys often cause innocent suspects to perceive their situation as hopeless, and thus are not only present but also a significant factor in virtually every police-induced false confession in America."[135]

The impact of FEPs on false confessions is also borne out by multiple experimental studies.[136] In one study, nearly 100% of subjects who viewed "fake-video evidence" falsely confessed to an act that they did not commit,[137] and they did so earlier than those who were merely told that video evidence existed.[138] This impact has been replicated in other experiments.[139]

Why individuals falsely confess to crimes is in part explained by the persuasive power of FEPs in the cost-benefit analysis individuals face when deciding whether to confess. As noted, a prime goal of

133. KASSIN, *supra* note 60, at 116.

134. *See* State v. Griffin, 262 A.3d 44, 104–05 n.28 (Conn. 2021) (Ecker, J., concurring in part and dissenting in part) (noting the case as well as several others); Miriam S. Gohara, *A Lie for a Lie: False Confessions and the Case for Reconsidering the Legality of Deceptive Interrogation Techniques*, 33 FORDHAM URB. L.J. 791, 796–97 (2006) (discussing other cases).

135. Richard A. Leo, *Structural Police Deception in American Police Interrogation: A Closer Look at* Minimization *and* Maximization, *in* INTERROGATION, CONFESSION, AND TRUTH: COMPARATIVE STUDIES IN CRIMINAL PROCEDURE 183, 194 (Lutz Eidam et al. eds., 2020).

136. *See, e.g.*, Saul M. Kassin & Katherine L. Kiechel, *The Social Psychology of False Confessions: Compliance, Internalization, and Confabulation*, 7 PSYCH. SCI. 125, 127 (1996); Jennifer T. Perillo & Saul M. Kassin, *Inside Interrogation: The Lie, the Bluff, and False Confession*, 35 LAW & HUM. BEHAV. 327, 335 (2011); *see also* Danielle E. Chojnacki et al., *An Empirical Basis for the Admission of Expert Testimony on False Confessions*, 40 ARIZ. ST. L.J. 1, 18 (2008) (citing and discussing studies). Perhaps the most widely cited study, based on the "ALT Key" experiment conducted by Kassin and Kiechel (cited at the outset of this footnote), has been criticized over the years as being an unrealistic basis to support that false evidence induces false confessions. For a spirited defense of the study's findings and its importance to the field, see Alan Hirsch, *So Misunderstood: The ALT Key Experiment and False Confessions*, CHAMPION, April 2022, at 34, https://perma.cc/CZM8-K9X8.

137. Nash & Wade, *supra* note 113, at 633.

138. *Id.* at 629.

139. *See, e.g.*, Robert Horselenberg et al., *Individual Differences and False Confessions: A Conceptual Replication of Kassin and Kiechel (1996)*, 9 PSYCH. CRIME & L. 1, 1 (2003); Saul M. Kassin & Gisli H. Gudjonsson, *The Psychology of Confessions: A Review of the Literature and Issues*, 5 PSYCH. SCI. PUB. INT. 33, 33 (2004); Kassin et al., *Police-Induced Confessions*, *supra* note 37, at 17; Jessica R. Klaver et al., *Effects of Personality, Interrogation Techniques and Plausibility in an Experimental False Confession Paradigm*, 13 LEGAL & CRIMINOLOGICAL PSYCH. 71, 79 (2008).

interrogators is to instill in their subjects the sense that the amassed evidence makes confessing a rational, self-interested course of action;[140] that confessing is sensible because they are likely to be convicted anyway;[141] and that confessing can possibly secure lenience.[142] As the Utah Supreme Court put it, a suspect faced with police misrepresentation of evidence may well "conclude that, given the futility of resistance, it is most prudent to cooperate and even confess falsely in order to get leniency."[143] This pressure is likely

---

140. *See supra* notes 40–41 and accompanying text; *see also* State v. Fernandez-Torres, 337 P.3d 691, 702 (Kan. App. 2014) ("Although innocent, an individual may attribute the purported evidence against him to a horrible mistake rather than to the interrogator's deception. And the interrogator's categorical dismissal of each protest of innocence can cement that fear. The individual then considers the minimalized admission of guilt the interrogator has offered to be the best way out of an exceptionally bad predicament." (citations omitted)); Richard J. Ofshe & Richard A. Leo, *The Decision to Confess Falsely: Rational Choice and Irrational Action*, 74 DENV. L. REV. 979, 985 (1997) (observing that interrogators must make a suspect believe that confessing to a crime is "rational and appropriate"); Kyle C. Scherr et al., *Cumulative Disadvantage: A Psychological Framework for Understanding How Innocence Can Lead to Confession, Wrongful Conviction, and Beyond*, 15 PERSPS. ON PSYCH. SCI. 353, 353 (2020).

141. *See* Ofshe & Leo, *supra* note 140, at 985–86 ("Police elicit the decision to confess from the guilty by leading them to believe that the evidence against them is overwhelming, that their fate is certain (whether or not they confess), and that there are advantages that follow if they confess."); *id.* at 1013 ("The fact that the suspect's denials of the evidence ha[ve] failed to convince the investigator of . . . his innocence is intended to serve as a demonstration. The implication is that he will also be unable to convince a prosecutor, a judge, or a jury of his innocence. The investigator strives to create the impression that because his opinion is based on hard facts, all other equally reasonable and informed persons will reach the same conclusion.").

142. *See* Drizin & Leo, *supra* note 59, at 916–17. The Reid Manual itself concedes that, although lying to a suspect about inculpatory evidence in and of itself would not cause a false confession, "it becomes much more plausible that an innocent person may decide to confess" if "such false statements were . . . used to convince the suspect that regardless of his stated innocence, he would be found guilty of the crime and . . . sentenced to prison" but would be afforded leniency "if he cooperates by confessing." REID MANUAL, *supra* note 33, at 352.

143. State v. Rettenberger, 984 P.2d 1009, 1015 (Utah 1999); *see also* State v. Baker, 465 P.3d 860, 877 (Haw. 2020) ("The presentation of falsified incontrovertible evidence is designed to demonstrate to the accused that they will inevitably be found guilty of the alleged crime. The fact that the accused has failed to convince the interrogating officer of their innocence demonstrates that the accused will also be unable to convince a prosecutor, judge, or jury . . . . What ultimately makes deception about incontrovertible evidence insidious is the implied threat that it carries: independent incriminating evidence exists, so the accused should confess in order to enter a mitigating statement into the record." (citations omitted)); *Gray*, 480 S.W.3d at 264 ("When faced with seemingly insurmountable evidence, it becomes reasonable for one to perceive the futility of

especially acute with false representations of scientific evidence, as "both the guilty and the innocent have a harder time explaining away evidence that is allegedly derived from scientific technologies."[144] This trust extends to pseudoscientific technologies of little or no actual probative value.[145]

Research and experience also demonstrate the significant impact of multiple situational factors. In particular, emotional and psychological stress and fatigue, exacerbated by the duration of interrogation and isolation from the world outside, can reduce individuals' capacity to make reasoned long-term decisions in favor of short-term goals (such as ending the interrogation based on a false promise of being able to go home).[146] As one of the exonerated Central Park Five defendants said of his confession, extracted by an interrogation in which an FEP figured prominently,

> It's hard to imagine why anyone would confess to a crime they didn't commit. But when you're in that interrogation room, everything changes. During the hours of relentless questioning that we each endured, detectives lied to us repeatedly . . . . It felt like the truth didn't matter. Instead, it seemed as though they locked onto one theory and were hellbent on securing

---

maintaining innocence . . . . [W]e think an average defendant in [defendant's] situation would feel pressured to confess to the point that it usurps free will.").

144. Ofshe & Leo, *supra* note 140, at 1023; *see also Baker*, 465 P.3d at 877 (observing that resistance to interrogation is especially undercut "when the false evidence is characterized as scientific because people generally expect scientific tests to be accurate and trustworthy" (citations omitted)).

145. *See, e.g.*, WOODS, *supra* note 39, at 198 (recognizing that "[m]any people believe that forensic science can solve virtually any case. By mixing pseudoscience into their statements, investigators can often convince subjects that incriminating evidence has been found"). *See generally* Catherine E. White, Comment, *"I Did Not Hurt Him . . . . This Is a Nightmare": The Introduction of False, but Not Fabricated, Forensic Evidence in Police Interrogations*, 2015 WIS. L. REV. 941 (2015). For discussion of police use of the scientifically debunked methods of the polygraph and the "Computer Voice Stress Analyzer" as FEPs, see LEO, *supra* note 44, at 144–48.

146. *See* Deborah Davis & Richard A. Leo, *Interrogation-Related Regulatory Decline: Ego Depletion, Failures of Self-Regulation, and the Decision to Confess*, 19 PSYCH. PUB. POL'Y & L. 673, 673 (2012); Kassin et al., *Police-Induced Confessions*, *supra* note 37, at 14; Stephanie Maddon et al., *How Factors Present During the Immediate Interrogation Situation Produce Short-Sighted Confession Decisions*, 37 LAW & HUM. BEHAV. 60, 60 (2013).

On the tendency of interrogated individuals to discount the long-term consequences of making statements and over-estimate the short-term benefits of making statements that may shorten an interrogation see, for example, Yueran Yang et al., *Short-Sighted Confession Decisions: The Role of Uncertain and Delayed Consequences*, 39 LAW & HUM. BEHAV. 44 (2015).

> incriminating statements to corroborate it. A conviction rather than justice felt like the goal.[147]

Adding to the pressure and stress, police can demand an immediate now-or-never decision to confess—"Help yourself now, or we can't help you later."[148]

The impact of police deceit is such that it can convince individuals that they are guilty even though they have no memory of committing the crime.[149] Police encourage this belief by telling suspects that they may have repressed their memories of the crime, that people can commit crimes without remembering,[150] or perhaps that they are experiencing a split personality,[151] with FEPs playing a key role in suspects' confabulations.[152] As the Canadian Supreme Court noted, "police tactics [can] cause the innocent person to 'become confused, doubt his memory, be temporarily persuaded of his guilt and confess to a crime he did not commit."[153] While factually guilty parties "can

147. Y. Salaam et al., *Act Against Coerced Confessions*, N.Y. TIMES, Jan. 5, 2021, at A19.

148. KASSIN, *supra* note 60, at 133.

149. Ofshe & Leo, *supra* note 140, at 1031–32; *see also* Robert A. Nash et al., *Digitally Manipulating Memory: Effects of Doctored Videos and Imagination in Distorting Beliefs and Memories*, 37 MEMORY & COGNITION 414, 421 (2009) ("False evidence can, in effect, change the past . . . . [W]e have shown that even our memories of recent self-involving events can be modified by subtle and compelling digital trickeries . . . ."); Nash & Wade, *supra* note 113, at 625 ("[F]eeding subjects false information can transform their beliefs; it can cause them to think their memory is unreliable and it encourages them to turn to external sources [such as police] to infer whether an event genuinely happened."); Newman & Schwarz, *supra* note 12, at 3 ("Memory may be affected even when people correctly reject a photo as manipulated at the initial exposure, paralleling effects observed in the repetition of text. Having seen a claim before increases the odds of its later acceptance, even when people were aware at the initial exposure that the claim is false or comes from an untrustworthy source." (citations omitted)).

150. *See* Richard A. Leo, *False Confessions: Causes, Consequences, and Implications*, 37 J. AM. ACAD. PSYCHIATRY & L. 332, 339 (2009) (noting that when facing evidence proving a suspect's involvement in a crime, the "suspect offers up the remaining basis for his belief in his innocence: that he has no memory of committing the crime"); Ofshe & Leo, *supra* note 140, at 1107.

151. KASSIN, *supra* note 60, at 13 (recounting confession provided by 14-year-old suspect after police falsely told him that they had forensic evidence of his guilt and convinced him that he had a split personality, with his "bad" version killing his sister out of a jealous rage).

152. In one recent study, an interviewer "well trained in police interview tactics," including the presentation of "incontrovertible false evidence," induced 70% of participants to provide a detailed account of a false memory of committing theft or assault. *See* Julia Shaw & Stephen Porter, *Constructing Rich False Memories of Committing Crime*, 26 PSYCH. SCI. 291, 294, 296, 299 (2015).

153. R. v. Oickle, [2000] 2 S.C.R. 3, 28 (Can.). Indeed, it is not beyond the realm of possibility that police—using the ploy that the suspect "blacked out"—

counter . . . eyewitness evidence by claiming it is in error, and co-perpetrators' evidence by asserting it is a lie,"[154] Professors Ofshe and Leo observe, innocent suspects, who are more likely to waive their *Miranda* rights, believing they "have "[no]thing to hide,"[155] "have a harder time explaining away evidence that is allegedly derived from scientific technologies."[156] Research also demonstrates that juveniles,[157] individuals with an intellectual disability,[158] and those suffering from a mental illness[159] are especially prone to false confessions.

Finally, lay persons, including criminal suspects, commonly believe that police are not permitted to lie about the existence of incriminating evidence,[160] let alone fabricate and use it to obtain a confession.[161]

C. Research on Deepfakes

A portmanteau combining AI-enabled "deep learning" and "fake," deepfakes include fabricated audio, audiovisual, or still-image content created or altered to appear to be a genuine account of the

---

create and share a deepfake of a suspect herself confessing, prompting an additional confession. Thanks to Brandon Garret for noting the possibility.

154. Ofshe & Leo, *supra* note 140, at 1023.

155. Kassin, *supra* note 2, at 218.

156. Ofshe & Leo, *supra* note 140, at 1023 (concluding that "[b]oth the guilty and the innocent have a harder time explaining away evidence that is allegedly derived from scientific technologies" and that individuals find it difficult to contradict evidence carrying "the prestige and incomprehensibility of modern science").

157. Allison Redlich & Gail S. Goodman, *Taking Responsibility for an Act Not Committed: The Influence of Age and Suggestibility*, 27 LAW & HUM. BEHAV. 141, 151 (2003).

158. *See* Samson J. Schatz, *Interrogated with Intellectual Disabilities: The Risks of False Confession*, 70 STAN. L. REV. 643, 687 (2018) ("[A]round one quarter of those who have been proved to have falsely confessed . . . display indicia of intellectual disability."); *see also* Gisli H. Gudjonsson & Lucy Henry, *Child and Adult Witnesses with Intellectual Disability: The Importance of Suggestibility*, 8 LEGAL & CRIM. PSYCH. 241, 241 (2003).

159. *See generally* Allison D. Redlich et al., *Comparing True and False Confessions Among Persons with Serious Mental Illness*, 17 PSYCH. PUB. POL'Y & L. 394 (2011). *See also* Allison D. Redlich et al., *Self-Reported False Confessions and False Guilty Pleas Among Offenders with Mental Illness*, 34 LAW & HUM. BEHAV. 79, 81 (2010).

160. *See, e.g.*, Richard Rogers et al., *"Everyone Knows Their* Miranda *Rights": Implicit Assumptions and Countervailing Evidence*, 16 PSYCH. PUB. POL'Y & L. 300, 310 (2010) (finding that a majority of recent offenders mistakenly believed that it was illegal for police to lie about eyewitness evidence).

161. *See* Gray v. Commonwealth, 480 S.W.3d 253, 264 (Ky. 2016) ("Our operative assumption should not be an expectation that citizens should distrust everything that law enforcement tells them or shows them. The contrary should be true.").

speech, conduct, image, or likeness of an individual or event.[162] "They create a fake reality by superimposing a person's face on another's body, or by changing the content of one's speech."[163]

Rapidly evolving AI technology enables deepfakes to grow more sophisticated by the day, posing ever-greater challenges to their detection.[164] Research shows that individuals have little wherewithal to detect deepfakes,[165] even after they are educated regarding their possible earmarks;[166] that wherewithal to detect deepfakes can depend on whether the demographic group of the person depicted matches one's own;[167] and that older individuals are less capable of detecting deepfakes.[168] Research also shows that individuals are at once inclined to mistake deepfakes for authentic content and to significantly overestimate their ability to detect deepfakes,[169] making them especially credulous.[170] Also, lower image or video resolution

---

162. Douglas Harris, *Deepfakes: False Pornography Is Here and the Law Cannot Protect You*, 17 DUKE L. & TECH. REV. 99, 99–100 (2019).

163. Rebecca A. Delfino, *Deepfakes on Trial: A Call to Expand the Trial Judge's Gatekeeping Role to Protect Legal Proceedings from Technological Fakery*, 74 HASTINGS L.J. 293, 298 (2023).

164. *See* Nina I. Brown, *Deepfakes and the Weaponization of Disinformation*, 23 VA. J.L. & TECH. 1, 25–26 (2020) (discussing the belief among computer science and digital forensics experts that deepfake detection methods cannot keep pace with innovations in deepfake technology); Tiffany Hsu & Steven Lee Myers, *Another Side of the A.I. Boom: Detecting What A.I. Makes*, N.Y. TIMES (May 18, 2023), https://perma.cc/RMZ6-2PMG ("Detection tools inherently lag behind the generative technology they are trying to detect. By the time a defense system is able to recognize the work of a new chatbot or image generator . . . developers are already coming up with a new iteration that can evade that defense.").

165. *See* Nils C. Köbis et al., *Fooled Twice: People Cannot Detect Deepfakes but Think They Can*, ISCIENCE (Nov. 19, 2021), https://perma.cc/8764-HN9R.

166. *Id.* at 7.

167. Alena Birrer & Natascha Just, *What We Know and Don't Know About Deepfakes: An Investigation into the State of the Research and Regulatory Landscape*, NEW MEDIA & SOC'Y: ONLINEFIRST 6 (May 22, 2024), https://doi.org/10.1177/14614448241253138.

168. *Id.* at 6–7.

169. *See* Newman & Schwarz, *supra* note 12, at 2 ("People report high confidence that they can distinguish AI-generated images from real one but perform poorly at the task"); *see also* Sergi D. Bray et al., *Testing Human Ability to Detect "Deepfake" Images of Human Faces*, 9 J. CYBERSECURITY 1, 2 (May 11, 2023), https://doi.org/10.1093/cybsec/tyad011; Köbis et al., *supra* note 165; Elizabeth J. Miller et al., *AI Hyperrealism: Why AI Faces Are Perceived as More Real Than Human Ones*, 34 PSYCH. SCI. 1390, 1391 (2023); Sophie J. Nightingale & Kimberley A. Wade, *Identifying and Minimising the Impact of Fake Visual Media: Current and Future Directions*, MEMORY, MIND & MEDIA, Oct. 20, 2022, at 1, 1, https://perma.cc/G8TJ-MNM3.

170. *See* Nightingale & Wade, *supra* note 169, at 2 (noting that "people are not only poor at detecting manipulated photos, but also often unaware of just how poor they are"); *see also* Jeremy Kahn, *Facebook Contest Shows Just How Hard It Is to Detect Deepfakes*, FORTUNE (June 12, 2020), https://perma.cc/H2JE-5P78;

makes it more difficult to recognize whether content is authentic,[171] which can advantage a deepfake creator seeking to deceive.

The foregoing research findings assume greater importance when one considers two other research findings. One, noted above, is that suspects commonly lack awareness that police can lie about the existence of incriminating evidence, much less fabricate it.[172] Second, given that individuals are especially affected by police presentation of tangible fake evidence,[173] it is safe to presume that the same—and then some—will be the case with AI-generated deepfakes, which will become ever more sophisticated and persuasive as technology improves.[174]

III. WHY INTERVENTION IS NEEDED

The preceding discussion advanced three main points: (1) that police regularly use FEPs; (2) that physical presentation by police of fabricated evidence is especially impactful in securing confessions from guilty and innocent individuals alike; and (3) that use of AI-generated deepfakes will have significant appeal to police and will present unprecedented detection challenges, especially as technology outpaces human wherewithal to discern fakes. This Part builds upon these points, making the case that police use of deepfakes warrants particular due process concern. It also argues that critically important changes in the criminal legal system in the over half-century since *Frazier* was decided make imposing a per se prohibition on police deepfakes all the more important.

Sophie J. Nightingale & Henry Farid, *AI-Synthesized Faces Are Indistinguishable from Real Faces and More Trustworthy*, PNAS, Feb. 22, 2022, at 1, 1, https://doi.org/10.1073/pnas.2120481119.

171. Matthew Groh et al., *Deepfake Detection by Human Crowds, Machines, and Machine-Informed Crowds*, PNAS, Jan. 5, 2022, at 1, 9, https://perma.cc/HFH2-2VK4; Nils Hulzebosch et al., *Detecting CNN-Generated Facial Images in Real-World Scenarios*, CVPR (2020), https://perma.cc/X6RR-YR5R; Rashid Tahir et al., *Seeing Is Believing: Exploring Perceptual Differences in DeepFake Videos*, ACM DIGIT. LIBR. (May 7, 2021), https://dl.acm.org/doi/pdf/10.1145/3411764.3445699.

172. *See supra* notes 160–61 and accompanying text.

173. *See supra* notes 113–25, 136–39 and accompanying text.

174. Not all agree that use of the FEP is problematic. Professor Christopher Slobogin, for instance, reasons that if "the evidence is made-up, the pressure to talk is, at worst, no more intense and is probably much less, since the suspect, whether guilty or innocent, can often smell out the ruse." Christopher Slobogin, *The Legality of Trickery During Interrogation*, *in* INTERROGATION, CONFESSIONS, AND TRUTH: COMPARATIVE STUDIES IN CRIMINAL PROCEDURE, *supra* note 135, at 61, 77. Such a view, however, is at odds with research discussed *supra* showing the impact of FEPs on confessions (including false ones), the inability of individuals to discern AI-created deepfakes, and the common belief that police cannot use fabricated evidence in interrogations.

A. The Hazy Contours of Coercion

Although the Supreme Court long attached central importance to the reliability and trustworthiness of confessions in its due process analysis,[175] this is no longer the case.[176] According to the Court, today the focus is on coercion and involuntariness, with the due process analysis asking, based on an assessment of the totality of the circumstances, "Is the confession the product of an essentially free and unconstrained choice by its maker? . . . If it is not, if his will has been overborne and his capacity for self-determination critically impaired, the use of this confession offends due process."[177]

The question thus becomes whether and how the use of FEPs—and deepfakes in particular—"critically impair[]" the "capacity for self-determination" of an interrogated party.[178] As the foregoing discussion demonstrates, deepfake deception should indeed be considered coercive because, in the words of Judge Richard Posner, it "destroy[s] the information required . . . for a rational choice," one informed by factual reality.[179] Judge Posner recognized this coercion in excluding a defendant's confession after police lied about medical evidence implicating the defendant:

> [Officer] Micci induced Aleman's "confession" by lying to him about the medical reports. The lies convinced Aleman that he must have been the cause of Joshua's shaken-baby syndrome because, according to Micci, the doctors had excluded any other possibility . . . . Not being a medical expert, Aleman could not contradict what was represented to him as settled medical opinion. He had shaken Joshua, albeit gently; but if medical opinion excluded any other possible cause of the child's death, then, gentle as the shaking was, and innocently intended, it

---

175. Donald A. Dripps, *Guilt, Innocence, and Due Process of Plea Bargaining*, 57 WM. & MARY L. REV. 1343, 1368 (2016).

176. *See* Colorado v. Connelly, 479 U.S. 157, 167 (1986) ("The aim of the requirement of due process is not to exclude presumptively false evidence but to prevent fundamental unfairness in the use of evidence, whether true or false." (citation omitted)). Lack of reliability, however, such as is reflected in false confession cases involving FEP, is not irrelevant. This is because "[f]airness, voluntariness, and reliability overlap" and, while they embody "different values, they often move in the same direction." Lisa Kern Griffin, *Silence, Confessions, and the New Advocacy Imperative*, 65 DUKE L.J. 697, 741 (2016); *see also* Dickerson v. United States, 530 U.S. 428, 433 (2000) (stating that the common law "recognized that coerced confessions are inherently untrustworthy"); Jackson v. Denno, 378 U.S. 368, 386 (1964) (noting the "probable unreliability of confessions that are obtained in a manner deemed coercive").

177. Schneckloth v. Bustamonte, 412 U.S. 218, 225–26 (1973); *see also* Culombe v. Connecticut, 367 U.S. 568, 602 (1961) (stating that a court must assess whether an individual's "capacity for self-determination [was] critically impaired").

178. *See Schneckloth*, 412 U.S. at 225.

179. Aleman v. Village of Hanover Park, 662 F.3d 897, 906 (7th Cir. 2011).

> must have been the cause of death. Aleman had no rational basis, given his ignorance of medical science, to deny that he had to have been the cause.[180]

For the reasons discussed, police use of deepfakes will likewise deprive individuals of the capacity for voluntary rational choice when making the all-important decision to confess. As the Kentucky Supreme Court put it, "[w]hen faced with seemingly insurmountable evidence, it becomes reasonable for one to perceive the futility of maintaining innocence . . . . [W]e think an average defendant in [the defendant's] situation would feel pressured to confess to the point that it usurps free will."[181]

The *Frazier* Court in 1969 presumed that police deceit regarding the existence of incriminating evidence lacks significant persuasive force in securing confessions. We now know that this is not the case.[182] Moreover, as intuited by *Cayward* and other courts, we also now know that police physical presentation of false evidence (versus orally relating its existence, as in *Frazier*) has a uniquely powerful impact on individuals.[183] Such research should prompt revisiting the empirical baseline informing the *Frazier* Court's understanding of the impact of FEPs when assessing coercion.[184] The unprecedented verisimilitude of deepfake technology makes revisiting *Frazier* all the more important.[185]

---

180. *Id.*; *see also* State v. Johnson, 812 S.E.2d 739, 748 (S.C. Ct. App. 2018) (police trickery is coercive if it precludes free choice); Michael J. Zydney Mannheimer, *Fraudulently Induced Confessions*, 96 NOTRE DAME L. REV. 799, 841 (2020) ("Deceptive practices potentially skew the suspect's decision whether to confess by artificially altering her perception of the relative benefits of speaking and of remaining silent."); Patrick M. McMullen, *Questioning the Questions: The Impermissibility of Police Deception in Interrogations of Juveniles*, 99 NW. U. L. REV 971, 975 (2005) (finding that deception during interrogation often "effectively leaves the suspect with no rational choice but to confess"). *See generally* SISSELA BOK, LYING: MORAL CHOICE IN PUBLIC AND PRIVATE LIFE 19–20 (1978) (recognizing the distorting effect of deceit on the free will and decision-making of individuals).

181. Gray v. Commonwealth, 480 S.W.3d 253, 264 (Ky. 2016).

182. *See supra* notes 113–61 and accompanying text.

183. *See supra* notes 113–26, 136–39 and accompanying text.

184. *See* Dassey v. Dittmann, 877 F.3d 297, 331 (7th Cir. 2017) (en banc) (Rovner, J., dissenting, joined by Wood & Williams, JJ.) (recognizing that "reform of our [judicial understanding of coercion] . . . is long overdue. When conducting a totality of the circumstances review, most courts' evaluations of coercion still are based largely on outdated ideas about human psychology and rational decision-making. It is time to bring our understanding of coercion into the twenty-first century."); *id.* at 336 ("What has changed [since *Frazier*] is not the law, but our understanding of the facts that illuminate what constitutes coercion under the law.").

185. *See supra* notes 164–71 and accompanying text.

B. Systemic Changes

Major changes in the criminal legal system since 1969, when *Frazier* was decided, heighten the need for a per se prohibition of deepfakes.

One change of singular importance is the significantly increased pressure defendants face to plead guilty, which has heightened the importance of confessions.[186] Today, well over 90% of state and federal criminal convictions result from guilty pleas, not trials.[187] As the Supreme Court has recognized, "[i]n today's criminal justice system, . . . the negotiation of a plea bargain, rather than the unfolding of a trial, is almost always the critical point for a defendant."[188] Although guilty pleas also predominated in the late 1960s[189] the Supreme Court did not place its formal imprimatur on pleas until 1970,[190] and since then the number and rate of plea-based convictions has steadily risen,[191] along with the power of prosecutors.[192]

The dominance of pleas stems from several critically important developments. One is the growth of sentence lengths, based on the proliferation of mandatory minimum sentences and increases in allowable prison terms,[193] providing contemporary prosecutors very significant charging options and therefore also plea bargaining leverage.[194] Aggravating matters, jurisdictions have enacted a welter of newly codified offenses, many of which overlap with extant

---

186. *See* Brandon L. Garrett, *Why Plea Bargains Are Not Confessions*, 57 WM. & MARY L. REV. 1415, 1426 (2016) (noting that "[t]he act of confessing . . . has been seen as crucial to plea bargaining").

187. Missouri v. Frye, 566 U.S. 134, 143 (2012) (citing data from the U.S. Bureau of Justice Statistics).

188. *Id.* at 144.

189. *See* Albert W. Alschuler, *Plea Bargaining and Its History*, 79 COLUM. L. REV. 1, 6, 33–35 (1979).

190. Brady v. United States, 397 U.S. 742, 758 (1970).

191. *See* Allison D. Redlich et al., *The Psychology of Defendant Plea Decision Making*, 72 AM. PSYCH. 339, 340 (2017).

192. *See* Jed S. Rakoff, *Why Prosecutors Rule the Criminal Justice System—and What Can Be Done About It*, 111 NW. U. L. REV. 1429, 1430 (2017) ("The plea bargain is the ultimate source of this ever-increasing prosecutorial power.").

193. *See* Cynthia Alkon, *An Overlooked Key to Reversing Mass Incarceration: Reforming the Law to Reduce Prosecutorial Power in Plea Bargaining*, 15 U. MD. L.J. RACE, RELIGION, GENDER & CLASS, 191, 196 (2015).

194. *See* Mary Price, *Weaponizing Justice: Mandatory Minimums, the Trial Penalty, and the Purposes of Punishment*, 31 FED. SENT'G REP. 309, 309–12 (2019); *see also* Shima Baradaran Baughman & Megan S. Wright, *Prosecutors and Mass Incarceration*, 94 S. CAL. L. REV. 1123, 1128 (2021) ("[S]ince 94% of criminal convictions are resolved by plea bargain, prosecutors—not judges—determine a defendant's fate the vast majority of the time.").

offenses.[195] As a result, prosecutors can engage in "charge stacking," charging additional and/or more serious offenses than warranted by the facts,[196] and avail themselves of very substantial increases in sentences for offenses, compared to the era of *Frazier*.[197] Significant discounts for acceptance of responsibility under sentencing guidelines, "fact bargaining," and other changes have also enhanced the plea-bargaining power of prosecutors,[198] with their discretion going largely unconstrained and unsupervised.[199]

These changes have been complemented by Supreme Court decisions augmenting prosecutors' charging authority. *Bordenkircher v. Hayes*,[200] decided in 1978, almost a decade after *Frazier*, held that the prosecution can carry out a threat, made during plea negotiations, to bring far more serious charges against an accused who refused to plead guilty to a charged offense.[201] Also, prosecutors can and regularly do wield the cudgel of what is known as the "trial penalty": defense awareness of the difference between the typically discounted sentence offered in a plea and the much harsher sentence that would result based on conviction at trial,[202] which in homicide cases can mean the difference between life and death.[203] And, as a result of the

195. Paul H. Robinson et al., *The Modern Irrationalities of American Criminal Codes: An Empirical Study of Offense Grading*, 100 J. CRIM. L. & CRIMINOLOGY 709, 712 (2010) ("In the past three decades . . . legislatures have introduced a proliferation of new offenses that often overlap with prior existing laws and sometimes grade the same conduct at different levels of [crime] seriousness.").

196. Andrew Manuel Crespo, *The Hidden Law of Plea Bargaining*, 118 COLUM. L. REV. 1303, 1313–14, 1313 n.31 (2018); *see also* Kyle Graham, *Overcharging*, 11 OHIO ST. J. CRIM. L. 701, 704, 707, 710, 712 (2014).

197. *See* RACHEL ELISE BARKOW, PRISONERS OF POLITICS: BREAKING THE CYCLE OF MASS INCARCERATION 129 (2019); WILLIAM J. STUNTZ, THE COLLAPSE OF AMERICAN CRIMINAL JUSTICE 257–60, 263 (2011).

198. *See* GEORGE FISHER, PLEA BARGAINING'S TRIUMPH: A HISTORY OF PLEA BARGAINING IN AMERICA 210–30 (2003).

199. *See* Megan S. Wright et al., *Inside the Black Box of Prosecutor Discretion*, 55 U.C. DAVIS L. REV. 2133, 2186–89 (2022).

200. 434 U.S. 357 (1978).

201. *Id.* at 360–65.

202. *See* Brian D. Johnson, *Trials and Tribulations: The Trial Tax and the Process of Punishment*, 48 CRIME & JUST. 313, 313 (2019). For discussion of pronounced racial disparities in the plea-bargaining process, see Carlos Berdejo, *Criminalizing Race: Racial Disparities in Plea-Bargaining*, 59 B.C. L. REV. 1187 (2018); Elayne E. Greenberg, *Unshackling Plea Bargaining from Racial Bias*, 111 J. CRIM. L. & CRIMINOLOGY 93 (2021).

203. *See, e.g.*, Chris Ochoa & Carlita Salazar, *How the Threat of the Trial Penalty Coerces the Innocent to Plead Guilty: A First-Hand Account of an Exoneree*, 31 FED. SENT'G REP. 299, 299 (2019) (one author recounting how he was pressured to plead guilty to avoid the death penalty, and received a life sentence, only to be exonerated thirteen years later); *see also* Susan Ehrhard, *Plea Bargaining and the Death Penalty: An Exploratory Study*, 29 JUST. SYS. J. 313, 323 (2008).

Court's decision in *United States v. Goodwin*,[204] prosecutors can threaten to bring—and actually bring—more serious charges if the defendant requests a jury trial.[205]

For criminal suspects, a confession promises a degree of leverage in plea negotiations. Research shows that suspects quite understandably consider the strength of the government's evidence against them and the probability of conviction,[206] increasing pressure to accede to police deceit and cajoling to confess.[207] Use of the FEP enables interrogators to coerce confessions by artificially inflating a suspect's estimated likelihood of conviction and thereby makes a plea bargain appear rational and in their self-interest.[208]

Research also shows that defense counsel do not always protect their clients against such pressures. As noted earlier, counsel are typically absent from the interrogation room,[209] and, even if in a position to weigh in, often do not provide much resistance. Commonly facing overwhelming caseloads, poor training, and limited resources,[210] and acutely aware of the trial penalty,[211] defense counsel are known to encourage their clients (guilty and innocent alike) to confess and to plead guilty.[212]

---

204. 457 U.S. 368 (1982).

205. *See, e.g.*, United States v. Muldoon, No. 90–5057, 1991 WL 65768, at *2 (4th Cir. Apr. 30, 1991) (concluding that no "presumption of vindictiveness" arises "from plea negotiations when the prosecutor threatens to bring additional charges if the accused refuses to plead guilty to pending charges. The Due Process Clause does not bar the prosecutor from carrying out his threat.").

206. *See* Greg M. Kramer et al., *Plea Bargaining Recommendations by Criminal Defense Attorneys: Evidence Strength, Potential Sentence, and Defendant Preference*, 25 BEHAV. SCI. & L. 573, 575 (2007).

207. Luna, *supra* note 60, at 241, 250.

208. Melanie D. Wilson, *An Exclusionary Rule for Police Lies*, 47 AM. CRIM. L. REV. 1, 20–21 (2010).

209. *See supra* note 47 and accompanying text.

210. *See, e.g.*, Eve Brensike Primus, *Effective Trial Counsel After* Martinez v. Ryan*: Focusing on the Adequacy of State Procedures*, 122 YALE L.J. 2604, 2606 (2013) (noting consensus "that excessive caseloads, poor funding, and a lack of training plague indigent defense delivery systems throughout the states"); *see also* Jenia I. Turner, Ronald F. Wright & Michael Braun, *Neglected Discovery*, 73 DUKE L.J. 1173, 1174 (2024) (discussing results of empirical study demonstrating that defense counsel often fail to examine electronic discovery provided by the prosecution and offering several explanations, including "lack of technological skills and support; the overwhelming volume of digital discovery; the client's desire for fast resolution of the case; the lesser gravity of some cases; high caseloads; low compensation; and . . . lack of diligence").

211. *See supra* notes 202–03 and accompanying text.

212. Shari Seidman Diamond & Jessica M. Salerno, *Reasons for the Disappearing Jury Trial: Perspectives from Attorneys and Judges*, 81 LA. L. REV. 119, 149, 157–59 (2020); *see also* Rebecca K. Helm et al., *Limitations on the Ability to Negotiate Justice: Attorney Perspectives on Guilt, Innocence, and Legal Advice in the Current Plea System*, 24 PSYCH. CRIME & L. 915, 921 (2018)

Aggravating the power imbalance, post-*Frazier*, defendants at the pretrial stage very likely lack access to exculpatory information that the government might possess. Although *Brady v. Maryland*[213] requires that the government disclose to the defense any materially exculpatory information before trial commences,[214] the government might not have a constitutional duty to disclose to suspects exculpatory information before or during interrogations (i.e., pre-charge)[215] or plea negotiations.[216] Predictably, this knowledge deficit figures significantly in plea bargains. Research shows that individuals (innocent and guilty alike) who lack access to exculpatory information are much more likely to plead guilty.[217] As a consequence, even presuming the use of a deepfake constitutes exculpatory evidence, the government is not constitutionally obliged to acknowledge its use during interrogations or plea negotiations. As a result, as one prosecutor put it, defense attorneys are placed "in the

(discussing results of survey of defense counsel finding that almost 90% had cases where their innocent client pleaded guilty and 45% had cases where they advised an innocent client to plead guilty); Scott W. Howe, *Five Faces of the Public Defender*, 64 B.C. L. REV. 1507, 1523–33 (2023) (discussing roles of public defender as plea "mediator," plea "emissary," and "guilty-plea advocate").

213. 373 U.S. 83 (1963).

214. *Id.* at 86.

215. *See supra* note 48 and accompanying text.

216. *See* United States v. Ruiz, 536 U.S. 622, 633 (2002) (holding that "the Constitution does not require the Government to disclose material impeachment evidence prior to entering a plea agreement with a criminal defendant"). Disclosure was not constitutionally mandated because "impeachment information is special in relation to the fairness of a trial, not in respect to whether a plea is voluntary." *Id.* at 629. Whether *Ruiz* totally exempts the government from disclosure of exculpatory evidence at the plea stage is a matter of disagreement among state and lower federal courts. *See* Kelly S. Smith, Comment, *Assessing the National Landscape of Constitutional and Ethical Disclosure Requirements During Plea Bargaining: Louisiana Comes Up Short*, 98 TUL. L. REV. 537, 543–48 (2024).

217. *See* Samantha Luna & Allison D. Redlich, *Unintelligent Decision-Making? The Impact of Discovery on Defendant Plea Decisions*, 1 WRONGFUL CONVICTION L. REV. 314, 317, 330–31 (2020); *see also id.* at 332 ("[A]ccess to discovery information impacts defendants' ratings of the strength of the evidence against them and perceptions of the information itself, which in turn affects the decision to accept or reject pleas."). There is reason to believe that this lack of obligated disclosure disproportionately harms innocent defendants because they know less about the crime for which they are charged and therefore are less capable of evaluating the strength of the prosecution's purported evidence and seeking exculpatory evidence. *See* John G. Douglass, *Can Prosecutors Bluff?* Brady v. Maryland *and Plea Bargaining*, 57 CASE W. RSRV. L. REV. 581, 582 (2007) (noting that "nondisclosure disproportionately harms the innocent since, almost by definition, guilty defendants know more about the facts surrounding a crime than do those who are factually innocent").

untenable position of having to advise their clients without the tools they need to make sure their advice is solid."[218]

Another factor pressuring pleas is that prosecutors often impose time limits,[219] with some offices adopting a "best first" policy prescribing that the first plea offered be the most favorable.[220] These constraints limit the wherewithal of defense counsel to conduct factual investigations that might turn up exculpatory information,[221] which, as noted, need not always be disclosed by the government during plea negotiations.[222]

Even if the defense is somehow aware of a possible deepfake, experts capable of discerning the fakery are hard to come by, and

---

218. David A. Lord, *Breaking the Faustian Bargain: Using Ethical Norms to Level the Playing Field in Criminal Plea Bargaining*, 35 GEO. J. LEGAL ETHICS 73, 95 (2022); *see also id.* (noting that "it is in plea negotiations where divulging exculpatory evidence is most needed because it is those cases where exculpatory evidence is most likely to distort the playing field between the prosecution and defense if not disclosed").

219. Cynthia Alkon, *The U.S. Supreme Court's Failure to Fix Plea Bargaining: The Impact of* Lafler *and* Frye, 41 HASTINGS CONST. L.Q. 561, 597 (2014) (observing the commonality of plea-offer "time limits" and "'[e]xploding offers' . . . as prosecutors will regularly say, 'If your client doesn't take this deal today, I will add that prior and he will be looking at double the time.'"); Jed S. Rakoff, *Why Innocent People Plead Guilty*, N.Y. REV. BOOKS, Nov. 20, 2014 ("If, however, the defendant wants to plead guilty, the prosecutor will offer him a considerably reduced charge—but only if the plea is agreed to promptly (thus saving the prosecutor valuable resources). Otherwise, he will charge the maximum, and, while he will not close the door to any later plea bargain, it will be to a higher-level offense than the one offered at the outset of the case."); *see also* United States v. Gonzalez-Vasquez, 219 F.3d 37, 42 (1st Cir. 2000) ("It is axiomatic that a prosecutor may withdraw a plea offer before a defendant accepts it.").

220. Besiki Luka Kutateladze et al., *Opening Pandora's Box: How Does Defendant Race Influence Plea Bargaining?*, 33 JUST. Q. 398, 418 (2016).

221. A vigorous (yet time-consuming) client interview by defense counsel might illuminate details of the interrogation that can serve as the basis to learn of and challenge police use of a deepfake. For a helpful how-to by a Wisconsin public defender, see generally Deja Vishny, *Defending Unrecorded False Confession Cases*, CHAMPION, Dec. 2007, at 22, https://www.nacdl.org/Article/December2007-DefendingUnrecordedFalseConfes.

222. *See supra* notes 213–18 and accompanying text. The practical significance of the defense knowledge deficit is highlighted by the case of Joseph Buffey, who pled guilty to robbing and sexually assaulting an elderly woman. *See* Redlich et al., *supra* note 191, at 339. Faced with a plea offer that was set to expire, Buffey's lawyer urged him to accept the offer, believing that any later offer would not be as attractive. *Id.* Buffey accepted the plea, not knowing that DNA evidence in the possession of the prosecution excluded him as the perpetrator. *Id.* Despite his actual innocence, Buffey stated at his plea hearing that he "broke into an elderly lady's house and robbed her and forced her to have sex with [him]." *Id.*

when available,[223] very costly,[224] an expense likely beyond the reach of indigent individuals who comprise the vast majority of criminal defendants.[225] As a result, defendants must often evaluate an offer almost exclusively based on the perceived likelihood of conviction, a calculus in which a confession weighs heavily. Furthermore, in plea deals, prosecutors commonly secure a defense waiver of the right to appellate review,[226] which further lessens the possibility of learning of the government's use of a deepfake in securing a confession.

In short, significant changes occurring over the past half-century in the criminal legal system, especially those increasing pressure on defendants to confess and plead guilty, have made it all the more important to address police use of deepfakes in interrogations.

IV. THE PATH AHEAD

Over time, a growing chorus of commentators has condemned police use of deception in interrogations. To many, police deceit of any kind regarding evidence should be banned altogether, based inter alia on a moral objection to the practice[227] due to concern that it contravenes principles of procedural justice[228] and undermines

223. *See* Delfino, *supra* note 163, at 334 (noting that "only a handful of technology experts and digital forensic experts fully grasp" how AI technology operates).

224. The cost can range from several thousand dollars to well over $100,000, with the typical analyses being somewhere in the $5,000 to $15,000 range. Betsy Mikalacki, *How Much Does Digital Forensic Services Cost?*, VESTIGE (Feb. 2, 2017), https://perma.cc/RB9J-MNQU; *see also* Deborah G. Johnson & Nicholas Diakopoulos, *Computing Ethics: What to Do About Deepfakes*, 64 COMMC'NS ACM 33, 33–35 (2021) (discussing the role of expertise in addressing deepfakes).

225. *See* CAROLINE WOLF HARLOW, U.S. BUREAU OF JUST. STAT., NCJ 179023, DEFENSE COUNSEL IN CRIMINAL CASES 1, 5 (2000) (stating that public defenders represent approximately 80% of felony criminal defendants in large state courts due to defendant indigence).

226. In their analysis of data from 971 randomly selected federal cases, King and O'Neill found that defendants waived their rights to appellate review in nearly two-thirds of the cases resolved by plea agreement. Nancy J. King & Michael E. O'Neill, *Appeal Waivers and the Future of Sentencing Policy*, 55 DUKE L.J. 209, 209 (2005).

227. *See, e.g.*, Margaret L. Paris, *Trust, Lies, and Interrogation*, 3 VA. J. SOC. POL'Y & L. 3, 9, 44–45 (1995); Deborah Young, *Unnecessary Evil: Police Lying in Interrogations*, 28 CONN. L. REV. 425, 426 (1996); Amelia Courtney Hritz, Note, *"Voluntariness with a Vengeance": The Coerciveness of Police Lies During Interrogation*, 102 CORNELL L. REV. 487, 487 (2017).

228. *See, e.g.*, Margareth Etienne & Richard McAdams, *Police Deception in Interrogation as a Problem of Procedural Legitimacy*, 54 TEX. TECH L. REV. 21, 21–22 (2021); Julia Simon-Kerr, *Public Trust and Police Deception*, 11 NE. U. L. REV. 625, 627 (2019).

human dignity interests because it deprives a suspect of the capacity to make decisions based on full and accurate knowledge.[229]

The argument advanced here stops short of a universal ban on police deceit in interrogations,[230] arguing instead that police creation and presentation of AI-generated deepfakes should be singled out in particular for prohibition.[231] This Part first makes the affirmative case for why a per se ban on deepfakes is needed, then discusses why tweaks to the current procedural framework governing the regulation of confessions are not up to the job of addressing the unique problems presented by deepfakes.

A. Why a Per Se Ban Is Needed

Several reasons support a per se ban on police creation and use of deepfakes in interrogations.

First and foremost is research demonstrating the potent impact of police presentation of fabricated tangible evidence,[232] which several courts over time have rightly identified as especially problematic for its coercive effect in securing confessions.[233] The unprecedented verisimilitude of deepfakes and the inability of humans to detect their

---

229. *See, e.g.*, George E. Dix, *Federal Constitutional Confession Law: The 1986 and 1987 Supreme Court Terms*, 67 TEX. L. REV. 231, 344–45 (1988); Eugene R. Milhizer, *Rethinking Police Interrogation: Encouraging Reliable Confessions While Respecting Suspects' Dignity*, 41 VAL. U. L. REV. 1, 88 (2006).

230. Of note, the American Law Institute, in its recently promulgated *Principles of the Law, Policing*, urges that police "avoid the use of deceptive techniques that are likely to confuse or pressure suspects in ways that might undermine accuracy of evidence." PRINCIPLES OF THE LAW, POLICING § 11.04(d) (AM. L. INST. 2023). The commentary expresses particular concern about the false evidence ploy:

> Use[s] of false evidence, such as fictional accounts that lead suspects to believe that DNA conclusively ties them to the crime, that an eyewitness has identified them as the perpetrator, or that another suspect has called them a fellow conspirator, are especially likely to increase both the risk of false confessions and of the police receiving unreliable information from a guilty suspect; they also can be highly coercive, and can harm the legitimacy of investigations.

Id. § 11.04 cmt. e.

231. Although the focus here is on deepfake audio, audiovisual, and still images, AI is capable of generating other fabrications, including handwriting style. *See, e.g.*, Sawdah Bhaimiya, *Researchers Have Developed an AI Tool That Can Closely Imitate Your Handwriting Style*, BUS. INSIDER (Jan. 16, 2024), https://www.businessinsider.com/researchers-develop-ai-tool-that-can-imitate-your-handwriting-style-2024-1. While perhaps comparatively rare, police have fabricated signatures as part of an FEP. *See, e.g.*, Lincoln v. State, 882 A.2d 944, 959 (Md. Ct. Spec. App. 2005) (addressing police fabrication of handwritten statement (with signature) of witnesses known to suspect).

232. *See supra* notes 113–25, 136–39 and accompanying text.

233. *See supra* notes 79–112 and accompanying text.

falsity will only become more problematic over time as deepfake technology progresses,[234] making a ban all the more important.

A second, related reason is that the fabricated content could well have a shelf life and influence beyond the interrogation room. The court in *Cayward* invoked this concern when outlawing police creation and use of fabricated interrogation props,[235] a concern which the markedly greater verisimilitude of deepfakes amplifies. So too does the advent of the internet, where content now lives in worldwide ineradicable perpetuity, which was not yet publicly available when *Cayward* was decided in 1989,[236] let alone in 1969, when *Frazier* was decided.

A third reason in support of a ban is that it will provide a clear prohibitory anchor in the notoriously amorphous voluntariness standard. Voluntariness, as noted earlier, depends on whether the confession was coerced by police—whether the confessor's will was overborne.[237] The standard has been justly criticized,[238] with the Supreme Court itself recognizing that this test "yield[s] no talismanic definition of 'voluntariness' mechanically applicable to the host of situations where the question has arisen,"[239] and that "[t]he notion of 'voluntariness' is itself an amphibian."[240] With only a select few practices prohibited, such as actual or threatened physical mistreatment or deprivation of food or water, or an express promise of a more lenient sentence if a suspect confesses,[241] courts have

---

234. *See supra* notes 164–71 and accompanying text.

235. State v. Cayward, 552 So. 2d 971, 974 (Fla. Dist. Ct. App. 1989).

236. *See id.* (expressing concern over "the common availability of photo reproduction processes, conducive to the widespread dissemination of documents").

237. *See supra* notes 65–68 and accompanying text.

238. *See, e.g.*, Richard A. Leo et al., *Promoting Accuracy in the Use of Confession Evidence: An Argument for Pretrial Reliability Assessments to Prevent Wrongful Convictions*, 85 TEMP. L. REV. 759, 783–84 (2013) (citations and footnotes omitted):

> The due process voluntariness test has long been criticized by numerous legal scholars. As many have pointed out, it is vague, ambiguous, and ultimately indeterminate, if not incoherent. Moreover, because it lacks clarity, the . . . test, in application, has proven inconsistent and unpredictable. It has thus failed to provide meaningful guidance to judges, lawyers or even the police themselves . . . . Perhaps not surprisingly, trial courts rarely find confessions to be involuntary; what's more, they routinely find confessions voluntary that are the product of extreme pressure, threats, and promises. Whether involuntary or not, confessions are rarely excluded from evidence at trial.

239. Schneckloth v. Bustamonte, 412 U.S. 218, 224 (1973).

240. Culombe v. Connecticut, 367 U.S. 568, 604–05 (1961); *see also* LEO, *supra* note 44, at 277 ("The voluntariness test . . . invites inconsistent application.").

241. *See* Paul Marcus, *It's Not Just About* Miranda*: Determining the Voluntariness of Confessions in Criminal Prosecutions*, 40 VAL. U. L. REV. 601,

essentially unfettered rein in assessing voluntariness, usually finding it to be present.[242] Exacerbating matters, courts very often fail to undertake a due process test altogether, finding that a defendant's waiver of *Miranda* rights, which is very common, means that a resulting confession was voluntary for due process purposes.[243]

Fourth, an outright ban will send a clear signal to all justice system actors that deepfakes are not allowed. Police, who currently lack meaningful ex ante guidance on the parameters of permissible interrogation techniques,[244] will know that deepfakes are not permitted.[245] So will interrogated individuals, who—if aware of police authority to deceive—need not be wary of police resorting to deepfakes.[246]

Fifth and finally, a ban will guard against erosion of public faith in the reliability of judicial proceedings.[247] Although the prosecution cannot knowingly introduce fabricated evidence at trial,[248] the actual

---

619, 623–24 (2006); *cf.* State v. Patton, 826 A.2d 783, 805 (N.J. Super. Ct. App. Div. 2003) (likening use of "police-fabricated tangible evidence" to actual physical coercion).

242. *See* Marcus, *supra* note 241, at 643 (reporting on results of review of twenty years of state and lower federal court decisions, finding among courts a high tolerance for abusive practices, and stating that "[t]he due process test offers almost no guidance for lawyers and judges").

243. GEORGE C. THOMAS III & RICHARD A. LEO, CONFESSIONS OF GUILT: FROM TORTURE TO *MIRANDA* AND BEYOND 219 (2012) ("[C]ourts tend to treat a *Miranda* waiver as a near-conclusive presumption that all subsequent statements are uncoerced.").

244. *See* Mark A. Godsey, *Rethinking the Involuntary Confession Rule: Toward A Workable Test for Identifying Compelled Self-Incrimination*, 93 CALIF. L. REV. 465, 470 (2005) ("[T]he theoretical ambiguity inherent in the voluntariness standard leaves police officers with little guidance in the field . . . ."); Kate Levine, *Police Suspects,* 116 COLUM. L. REV. 1197, 1215 (2016) ("[T]he voluntariness standard puts almost no restrictions on what police may do to induce a confession.").

245. *See supra* note 99.

246. *See* Katie Wynbrandt, Comment, *From False Evidence Ploy to False Guilty Plea: An Unjustified Path to Securing Convictions*, 126 YALE L.J. 545, 559 (2016) (advocating prohibition of FEPs in general because prohibition would provide a "signaling value to suspects by giving them a measure of confidence that police are telling the truth when they make certain claims in an interrogation, enabling defendants to evaluate their likelihood of conviction more accurately during the crucial time when they must decide whether to plead guilty").

247. This is already a matter of significant concern in civil litigation. *See, e.g.*, Delfino, *supra* note 163, at 297 (urging changes to evidence law and trial procedures to better ensure the authenticity of images admitted into evidence and stating that "[a]s deepfake technology improves and it becomes harder to tell what is real, juries may start questioning the authenticity of properly admitted evidence, which may in turn have a corrosive effect on the justice system").

248. *See* Napue v. Illinois, 360 U.S. 264, 269 (1959). *But cf.* Patrick Ryan, *"Deepfake" Audio Evidence Used in UK Court to Discredit Dubai Dad*, NATIONAL

deepfake used to secure a confession in an interrogation would not be subject to the prohibition, absent a ban. In a world where deepfakes are condoned in interrogations, it is not hard to imagine that judges, jurors, witnesses, and members of the public alike, especially if made aware of their inability to reliably discern fabrications,[249] will be skeptical of all they see, undermining both the utility of the evidence in prosecutions and the public confidence in the government's credibility in criminal cases.[250] As Professor Rebecca Delfino has written,

> as the public becomes more aware that video and audio can be convincingly faked, liars will exploit this awareness to escape accountability for their actions by denouncing authentic video and audio as deepfakes. Even technology experts fear that a skeptical public will be primed to doubt the authenticity of real audio and video evidence.[251]

Moreover, allowing deepfakes by government agents will have yet another problematic consequence: the disinformation they embody will exacerbate the growing distrust in governmental institutions, undermining the nation's social fabric and democratic political order.[252]

(Feb. 8, 2020), https://perma.cc/LWB2-DH8D (describing use in child custody case of deepfake audio containing falsified violent threats by child's father against the child's mother).

249. *See supra* notes 164–71 and accompanying text.

250. This environment will be conducive to what is known as the "deepfake defense" and "liar's dividend." *See* Delfino, *supra* note 163, at 310. The former is a courtroom strategy "built around the premise that the audiovisual material introduced as evidence against the defendant is claimed to be fake . . . . [L]awyers may plant the seeds of doubt in jurors' minds to question the authenticity of all audiovisual images even where the lawyers know the evidence is real." Rebecca A. Delfino, *The Deepfake Defense—Exploring the Limits of the Law and Ethical Norms in Protecting Legal Proceedings from Lying Lawyers*, 84 OHIO ST. L.J. 1067, 1070 (2024) (footnotes omitted). The "liar's dividend" occurs "when the ability to create convincing fakes allows the image creators to undermine the veracity of real information by claiming that it, too, is a fabrication . . . . Deepfakes make it easier for the liar to deny the truth." *Id.* at 1074 (footnotes omitted).

251. Delfino, *The Deepfake Defense*, *supra* note 250, at 1074; *cf.* State v. Cayward, 552 So. 2d 971, 975 (Fla. Dist. Ct. App. 1989) (expressing concern that police-fabricated documents would "severely diminish our confidence in relying upon facially valid documents in court files").

252. *See* Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753, 1776–77 (2019) (identifying risks from deepfakes, including distortion of democratic discourse, erosion of public trust in institutions, and exacerbation of social divisions); Mika Westerlund, *The Emergence of Deepfake Technology: A Review*, 9 TECH. INNOVATION MGMT. REV. 39, 42–43 (2019) (describing how the public may begin to distrust authorities deemed reliable in the past because of deepfakes); *see also* NAT'L CTR. FOR STATE CTS., PUBLIC TRUST AND CONFIDENCE RESOURCE

B. How a Ban Can Be Achieved

A ban could be achieved by judicial or legislative action. As discussed, several courts have held that police creation and presentation of false physical evidence violates due process,[253] and courts will be even more justified in banning deepfakes. And if a case should reach the Supreme Court, the Court should at a minimum, if not overruling it outright, distinguish *Frazier v. Cupp* by holding that police creation and use of a deepfake intrinsically differs from an oral lie about the existence of evidence (condoned in *Frazier*). Such a doctrinal shift, the Court should recognize, is especially warranted given the many significant changes to the criminal legal system occurring over the past half-century, as discussed.[254]

Legislation would likely face the countervailing headwinds common to bills limiting police authority,[255] but recent laws enacted in several states prohibiting police use of deceit in interrogations of juveniles hold precedential promise.[256] Moreover, in Texas a confession obtained as the result of fraudulently altering a document is subject to exclusion,[257] and in New York a bill was recently proposed (but not passed) that would have deemed a confession involuntarily when it is obtained from a defendant "by knowingly communicating false facts about evidence to the defendant."[258] Outside the interrogation context, multiple governments have acted to ban (indeed, often criminalize) deepfakes, not only in regard to

---

GUIDE 1 (2020) ("Because the judicial branch relies heavily on public support to perform its role in our system of government, public trust and confidence is a precious commodity for the courts.").

253. *See supra* notes 79–112 and accompanying text.

254. *See supra* notes 186–226 and accompanying text; *cf.* Lafler v. Cooper, 566 U.S. 156, 170 (2012) (relying on empirical reality that today the criminal legal system is a "system of pleas, not a system of trials" when deciding to extend the right to effective assistance of counsel to plea bargains).

255. *See generally* Stephen Rushin & Zoe Robinson, *The Law Enforcement Lobby*, 106 MINN. L. REV. 1965 (2023). A statute prohibiting police creation and presentation of deepfakes presumably would not present First Amendment free speech or invasion of privacy concerns. *See, e.g.*, Shannon Reid, Comment, *The Deepfake Dilemma: Reconciling Privacy and First Amendment Claims*, 23 U. PA. J. CONST. L. 209, 224 (2020). A deepfake created and used by police in an interrogation would not likely qualify as protectible speech, so far as police are concerned, and privacy would be implicated in only a highly unusual case and likely be raised by an interrogation subject. *See id.* at 216.

256. Eve Brensike Primus, *The State[s] of Confession Law in a Post*-Miranda *World*, 115 J. CRIM. L. & CRIMINOLOGY (forthcoming 2025) (manuscript at 30–31), https://perma.cc/4JWB-24L4; *see also* Alexandra Warnock, *The Truth About Police Deception and Minors: Why North Carolina Should Ban Police Lying to Minors During Interrogations*, 12 WAKE FOREST J.L. & POL'Y 427, 428–29 (2022).

257. Wilson v. State, 311 S.W.3d 452, 455, 465 (Tex. Crim. App. 2010) (interpreting TEX. CODE CRIM. PROC. art. 38.23).

258. S. 324-A, 2021-2022 Reg. Sess. (N.Y. 2021).

"revenge porn" and other sexual content, but in a broad array of other contexts such as defamation, copyright and trademark infringement, harassment, and election influence.[259]

Police, of course, are not permitted to manufacture or fabricate physical evidence to secure a conviction.[260] They should not get a pass with respect to deepfakes in the high-stakes realm of interrogations, where not only the physical liberty, life, and personal reputation of suspects is at stake, but also the integrity and trustworthiness of a critically important government actor (police) and the criminal legal system more generally.[261]

A key issue remains, however: How to smoke out government use of a deepfake in an interrogation? Interrogations themselves are notoriously opaque,[262] defense counsel are very often absent,[263] and,

---

259. *See* Michelle M. Graham, *Deepfakes: Federal and State Regulation Aims to Curb a Growing Threat*, REUTERS (June 26, 2024), https://perma.cc/65YS-M6NK.

260. *See* MICHAEL AVERY ET AL., POLICE MISCONDUCT, LAW AND LITIGATION § 2:30 (2024–2025 ed.); *see also, e.g.*, Black v. Montgomery County, 835 F.3d 358, 371–72 (3d Cir. 2016) (plaintiff acquitted at trial may state a fabricated evidence claim under 42 U.S.C. § 1983 by showing a reasonable likelihood that she would not have been criminally charged).

China has employed deepfake videos to actually create (not induce) false confessions, which have multiple advantages over traditional torture-induced confessions. According to one recent media report:

> Using deepfake AI, the [Chinese Communist Party] can now skip the middleman; professional torturers will lose their jobs. AI-powered human impersonation can create vides of the accused person confessing, crying, begging, groveling—all the details appearing in a video so lifelike that even relatives will believe it . . . . Engage AI to study body, face, movements, and voice and then just "disappear" the person. All the regime will need is the video file to broadcast nationally or worldwide. Disappearing a person without torture pays off because the victim can never escape, smuggle out messages, or otherwise reveal the regime's methods. Deepfake gives the regime all the propaganda and terror without the downsides of a living victim.

Richard W. Stevens, *AI's Next Gift to Evil: Forced Confession Propaganda in China*, EPOCH TIMES (Mar. 26, 2024), https://perma.cc/6T2S-25TZ.

261. *See* Gray v. Commonwealth, 480 S.W.3d 253, 264 (Ky. 2016); State v. Cayward, 552 S.E.2d 971, 975 (Fla. App. 1989); State v. Patton, 826 A.2d 783, 800 (N.J., 2003); *see also* Simon-Kerr, *supra* note 228.

262. *See, e.g.*, Richard A. Leo, *Police Interrogation and Social Control*, 3 SOC. & LEGAL STUD. 93, 98 (1994) ("It is difficult to investigate American police questioning practices empirically, for we know very little about what actually happens during custodial interrogation. Since interrogation has always been—and continues to be—shrouded in secrecy, the historical record is highly incomplete.").

263. *See supra* note 47. The knowledge gap might be mitigated in jurisdictions requiring the taping of confessions (but only to the extent that the deepfake utilized is discernible). For discussion of the limits of videotaping see, for example, G. Daniel Lassiter et al., *Evaluating Videotaped Confessions: Expertise*

as noted, discovery will not likely provide a knowledge conduit.[264] At the same time, it is unlikely that police sua sponte will willingly disclose use of a deepfake to secure a confession, given what we know of the effectiveness of FEPs in securing confessions.[265] However, precedent for transparency exists in laws recently requiring that the government acknowledge use of facial recognition technology,[266] "cell site simulators" (a.k.a. "stingrays"),[267] and eyewitness identifications.[268] Additional support for requiring disclosure can come from the growing chorus of public demands for greater transparency in government use of AI more generally.[269]

C. Why the Current Procedural Framework Is Inadequate

1. Motions to Suppress

If an individual wishes to contest government use of a confession, the primary avenue is a pretrial suppression hearing in which a judge determines whether the confession was voluntary and is therefore admissible as evidence.[270] As noted above, the voluntariness standard is famously indeterminate and difficult to satisfy.[271] One option, short of an outright ban, is to attach more analytic weight to police use of deepfakes in the voluntariness assessment undertaken by courts.[272]

Provides No Defense Against the Camera-Perspective Effect, 18 PSYCH. SCI. 224, 225 (2007). *Cf.* Remi Boivin et al., *The Body-Worn Camera Perspective Bias*, 13 J. EXPERIMENTAL CRIMINOLOGY 125, 126 (2016).

264. *See supra* notes 213–18 and accompanying text.

265. *See supra* notes 113–26, 136–39 and accompanying text.

266. Kaitlin Jackson, *Challenging Facial Recognition Software in Criminal Court*, CHAMPION, July 2019, at 16.

267. Spencer McCandless, Note, *Stingray Confidential*, 85 GEO. WASH. L. REV. 993, 995, 1020 (2015).

268. *See, e.g.*, N.Y. CRIM. PROC. L. § 710.20; *id.* § 710.30 (creating right to notice of and opportunity to challenge evidence that defendant was observed at the crime scene or otherwise identified by "pictorial, photographic, electronic, filmed or video recorded reproduction").

269. *See, e.g.*, Brandon L. Garrett & Cynthia Rudin, *The Right to a Glass Box: Rethinking the Use of Artificial Intelligence in Criminal Justice*, 109 CORNELL L. REV. 561, 561–62 (2024); State v. Arteaga, 296 A.3d 542, 558 (N.J. Super. Ct. App. Div. 2023) (holding that due process requires that defendant have access to information used to develop facial recognition technology employed to identify him).

270. Jackson v. Denno, 378 U.S. 368, 376–77, 395 (1964). At the motion to suppress stage, the court is only concerned with coercion, not reliability. *Id.* at 384–85. For an argument in favor of a pretrial hearing to assess reliability (especially at issue with possibly false confessions), based on Rule of Evidence 403, not coercion, see Leo et al., *supra* note 238, at 792.

271. *See supra* notes 237–42 and accompanying text.

272. *Cf.* State v. Griffin, 262 A.3d 44, 102 (Conn. 2021) (Ecker, J., concurring in part and dissenting in part) (urging an approach under which the FEP "is given

The approach has some appeal, but simply attaching more weight to the problematic strategy risks allowing it to get lost in the voluntariness totality-of-the-circumstances quagmire.[273]

Another option is to increase the government's burden of proof that a confession is voluntary. The Supreme Court has concluded that the government need prove voluntariness by a preponderance of the evidence,[274] meaning that a presiding judge need only have a "degree of confidence marginally greater than he or she would if flipping a coin."[275] Several states have seen fit to impose a more demanding standard, such as clear and convincing or beyond a reasonable doubt.[276] However, because most courts continue to abide by *Frazier v. Cupp*'s permissible stance regarding FEPs,[277] increasing the government's burden of persuasion does not hold much promise.

Less problematic would be a regime wherein police use of a deepfake raises a presumption of involuntariness that the government must rebut. Such an approach was adopted by the Kentucky Supreme Court with respect to police use of tangible FEPs.[278] Moreover, a bill was introduced in the Connecticut legislature in 2014 which would have established a rebuttable presumption of coercion when police knowingly present a suspect with false evidence or knowingly misrepresent evidence about the

---

greater weight in assessing the coerciveness of an interrogation under the totality of the circumstances test than it is currently given").

273. *Id.* at 90.

274. Lego v. Twomey, 404 U.S. 477, 489 (1972).

275. Michael D. Pepson & John N. Sharifi, Lego v. Twomey*: The Improbable Relationship Between an Obscure Supreme Court Decision and Wrongful Convictions*, 47 AM. CRIM. L. REV. 1185, 1193 (2010). According to the authors, use of a preponderance of evidence standard "leads to the erroneous admission of coerced confessions, which, in turn, often leads to unreliable verdicts." *Id.* at 1218.

276. *See, e.g.*, State v. Spooner, 404 So. 2d 905, 906 (La. 1981); Commonwealth v. Baye, 967 N.E.2d 1120, 1129 (Mass. 2012); State v. L.H., 215 A.3d 516 (N.J. 2019).

277. *See supra* notes 71–78 and accompanying text.

278. Gray v. Commonwealth, 480 S.W.3d 253, 263 (Ky. 2016); *see also supra* notes 102–12 and accompanying text. The *Gray* court failed to specify the poof standard the government must satisfy, merely stating that it "will presume this tactic is unconstitutional until the Commonwealth can firmly establish that the documents(s) did not overwhelm the defendant's will and was not a critical factor in the defendant's decision to confess." 480 S.W.3d at 263.

Of note, recently enacted laws in in Illinois and Oregon, prompted by concern over police deceit in interrogating juveniles, also impose a presumption of involuntariness that the state must rebut. K'reisa Cox, Note, *Curtailing Coercion of Children: Reforming Custodial Interrogation of Juveniles*, 49 J. LEGIS. 393, 405 (2023). Illinois requires that the government overcome the presumption of involuntariness by a preponderance of the evidence, while Oregon requires that the government do so by clear and convincing evidence. *Id.*

case, though it failed to pass.[279] However, use of a rebuttable presumption likewise lacks the several important clarifying benefits for system actors, noted above, of an outright ban.[280]

2. Guilty Plea Colloquies

The Supreme Court has described the guilty plea as a "grave and solemn act to be accepted only with care and discernment."[281] Judges must be satisfied that there is a sufficient factual basis to support the defendant's guilt[282] and accept a plea only if the defendant entered the plea knowingly, voluntarily, and intelligently.[283] In reality, the plea process regularly falls woefully short of these lofty expectations.

A foremost reason for this lack of protection is that courts often fail to ensure that a factual basis exists for a plea.[284] Courts frequently simply regard a guilty plea as itself of proof of factual guilt without further inquiry.[285] Moreover, they do not require, and defendants do not provide, detailed factual information supporting guilt, only formulaic admissions, not under oath.[286] A plea colloquy, the Supreme Court has instructed, is a "narrow inquiry" in which the defendant may provide little input, simply acceding to a "joint statement" with the prosecution.[287] Indeed, the factual basis can be based entirely on the government's recitation of the facts.[288] A defendant need only admit "in open court that he committed the acts charged in the indictment."[289] As longtime federal trial judge Jed Rakoff put it, judges "barely question the defendant beyond the basic bare bones of his assertion of guilt, relying instead on the prosecutor's

279. An Act Concerning Custodial Interrogation, H.B. 5589, 2014 Sess. § 1 (Conn. 2014). In written testimony submitted to the Judiciary Committee, the Division of Criminal Justice successfully urged no action on the bill, suggesting that the courts should address this concern on a case-by-case basis under the current state of the law rather than adopt a per se rule. *Joint Standing Comm. Hearings, Judiciary, Pt. 8*, 2014 Sess., 3564–65 (Conn. 2014).

280. *See supra* notes 244–52 and accompanying text; *see also supra* note 99.

281. Brady v. United States, 397 U.S. 742, 748 (1970).

282. *See, e.g.*, FED. R. CRIM. P. 11(b)(3).

283. *Brady*, 397 U.S. at 756; Boykin v. Alabama, 395 U.S. 238, 242 (1969).

284. *See, e.g.*, FED. R. CRIM. P. 11(b)(3); CAL. PENAL CODE § 1192.5(c) (2023); FLA. R. CRIM. P. 3.170(k) (2017); N.J. R. CT. 7:6-2(a)(1) (2007); WYO. R. CRIM. P. 11(f) (1993).

285. *See* Stephanos Bibas, *Plea Bargaining Outside the Shadow of Trial*, 117 HARV. L. REV. 2463, 2542–43 (2004).

286. In one study, for instance, 54% of individuals related that they were not asked by the court about the information the government had against them. Allison D. Redlich & Alicia Summers, *Voluntary, Knowing, and Intelligent Pleas: Understanding the Plea Inquiry*, 18 PSYCH. PUB. POL'Y & L. 626, 637 (2012).

287. Mitchell v. United States, 526 U.S. 314, 322–23 (1999).

288. *See* United States v. Ozuna-Cabrera, 663 F.3d 496, 501 (1st Cir. 2011).

289. Brady v. United States, 397 U.S. 742, 748 (1970).

statements (untested by a cross-examination) of what the underlying facts are."[290]

Moreover, the factual basis requirement, nominally a requirement in federal court, may be absent in state court proceedings,[291] which resolve the vast majority of U.S. criminal cases.[292] To the extent courts do undertake an inquiry, they draw upon a wide array of sources. As the Arizona Supreme Court put it, "The evidence of guilt may be derived from any part of the record including presentence reports, preliminary hearing transcripts, or admissions of the defendant."[293] Findings of fact, a key element of a valid plea colloquy, are thus "not subject to any meaningful testing."[294]

The availability and use of *nolo contendere* pleas exacerbate the problem. A *nolo* plea is not an admission of guilt but rather is "a consent by the defendant that he may be punished as if he were guilty and a prayer for leniency."[295] The plea does not require that there be an adequate factual basis for the plea.[296]

As noted, the court must also confirm that a defendant enters a plea voluntarily. It can be argued, of course, that the plea process itself is inherently coercive, given the array of significant pressures driving guilty pleas.[297] Nonetheless, voluntariness remains a formal requirement, but like the factual record requirement, courts do not assess it with rigor. Indeed, the Supreme Court has held that a plea can be deemed voluntary when based on a confession that was

---

290. Rakoff, *supra* note 219.

291. *See* Earl G. Penrod, *The Guilty Plea Process in Indiana: A Proposal to Strengthen the Diminishing Factual Basis Requirement*, 34 IND. L. REV. 1127, 1138–43 (2001); Jenia Iontcheva Turner, *Judicial Participation in Plea Negotiations: A Comparative View*, 54 AM. J. COMPAR. L. 199, 212–13 (2006).

292. *Compare* Christopher Slobogin, *The Case for a Federal Criminal Court System (and Sentencing Reform)*, 108 CALIF. L. REV. 941, 942 (2020) (estimating that 89,000 federal criminal cases were commenced between 2007 and 2017), *with CSP STAT Criminal*, COURT STAT. PROJ., https://www.courtstatistics.org/court-statistics/interactive-caseload-data-displays/csp-stat-nav-cards-first-row/csp-stat-criminal (Oct. 2024) (estimating that 39,390,000 non-traffic criminal cases were commenced in state courts across the United States between 2012 and 2023).

293. State v. Salinas. 887 P.2d 985, 987 (Ariz. 1994); *see also* Joshua Marquis, *The Myth of Innocence*, 95 J. CRIM. L. & CRIMINOLOGY 501, 509 n.56 (2005) (prosecutor explaining that "[w]hen a defendant agrees in a plea bargain that the state could prove a certain set of facts . . . that becomes the truth as much as it can ever be established in the eyes of the law").

294. Christopher Slobogin, *Plea Bargaining and the Substantive and Procedural Goals of Criminal Justice*, 57 WM. & MARY L. REV. 1505, 1518 (2016).

295. North Carolina v. Alford, 400 U.S. 25, 36 n.8 (1970).

296. *Id.*

297. *See supra* notes 67, 193–222 and accompanying text. For a more dramatic account, see John H. Langbein, *Torture and Plea Bargaining*, 46 U. CHI. L. REV. 3, 12–13 (1978).

unconstitutionally coercive[298] or entered into without access to material impeachment information,[299] and research shows that voluntariness is often absent from pleas.[300] Most germane to the discussion here, one must question how a confession-based plea can be deemed voluntary when it is predicated upon government-fabricated incriminating evidence (based on an interrogation strategy known to be coercive),[301] which, having heavily stacked the evidence cards against a defendant, makes the defendant's personally destructive decision appear "rational."[302]

3. Trials

Trials serve as another gatekeeping locus. After a trial judge makes the threshold voluntariness determination at a motion to suppress hearing, a confession is admissible at trial.[303] Thereafter, the jury, with the Supreme Court's blessing,[304] serves as the ultimate arbiter of the voluntariness and reliability of confessions when it adjudicates guilt.[305]

A growing body of research, however, suggests that juries are not up to the job.[306] Research shows that jurors are prone to believe that police have a superior ability to detect lies in the interrogation room,[307] despite a large body of evidence showing the contrary.[308] They also underestimate the degree to which psychologically coercive interrogation techniques will elicit false confessions.[309] Exacerbating matters, although juries consider FEPs deceptive and coercive,[310]

298. Parker v. North Carolina, 397 U.S. 790, 796 (1970); *see also* McMann v. Richardson, 397 U.S. 759, 770 (1970) (rejecting request to withdraw plea when defense counsel "misjudged the admissibility of the defendant's confession").

299. United Sates v. Ruiz, 536 U.S. 622, 629 (2002).

300. In one study, for instance, almost one-fifth of tender-of-plea forms did not mention any aspect of voluntariness. Allison D. Redlich & Catherine L. Bonventre, *Content and Comprehensibility of Juvenile and Adult Tender-of-Plea Forms: Implications for Knowing, Intelligent, and Voluntary Guilty Pleas*, 39 LAW & HUM. BEHAV. 162, 167 (2015).

301. *See supra* notes 113–26, 132–61 and accompanying text.

302. *See supra* notes 179–81 and accompanying text.

303. *See* Crane v. Kentucky, 476 U.S. 683, 687 (1986).

304. *See* Lego v. Twomey, 404 U.S. 477, 484–86 (1972).

305. *Crane*, 476 U.S. at 688.

306. *See, e.g.*, Linda A. Henkel et al., *A Survey of People's Attitudes and Beliefs About False Confessions*, 26 BEHAV. SCI. & L. 555, 576–82 (2008); Angela M. Jones et al., *Sensitivity to Psychologically Coercive Interrogations: A Comparison of Instructions and Expert Testimony to Improve Juror Decision-Making*, 21 J. FORENSIC PSYCH. RSCH. & PRAC. 373, 375–76 (2021).

307. *See* Mark Costanzo et al., *Juror Beliefs About Police Interrogations, False Confessions, and Expert Testimony*, 7 J. EMPIRICAL LEGAL STUD. 231, 238 (2010).

308. *See supra* note 42 and accompanying text.

309. Leo & Liu, *supra* note 121, at 394–95 (summarizing studies).

310. *See* Barnes et al., *supra* note 37, at 506–07 (summarizing studies).

they nonetheless lend evidentiary weight to confessions induced by them,[311] in significant part because they believe that individuals will not falsely confess.[312] Thus, even if use of a deepfake in an interrogation was somehow revealed, jurors will not likely serve as reliable gatekeepers.

Judges do little to ameliorate the situation. Like jurors, they have difficulty discerning coercion and deciding whether and how much weight to attach to confessions.[313] Also, like jurors, they may recognize the problematic coerciveness of FEPs but lend weight to the confessions they induce.[314] Judges view confessions, Professor Kassin observes, as "such powerful evidence that they do not discount [confessions] when it is legally and logically appropriate to do so."[315]

Finally, courtroom procedures fail to meaningfully regulate confessions.[316] As Professor Jeffrey Bellin recently observed, "[c]ourts and litigants spend little energy pondering the admissibility of criminal defendants' statements under the evidence rules."[317] "The only question is did the statements come from the party's mouth, pen, keyboard, etc. If the answer is yes, . . . the statements qualify for admission."[318] Bellin adds that the only recourse is for a defendant to take the witness stand and try to explain or refute a confession, an option discouraged by "[m]odern criminal procedure and evidence rules" that impose burdens in the form of "unfavorable legal and practical consequences."[319]

---

311. *See* Iris Blandón-Gitlin et al., *Jurors Believe Interrogation Tactics Are Not Likely to Elicit False Confessions: Will Expert Witness Testimony Inform Them Otherwise?*, 17 PSYCH. CRIME & L. 239, 247 (2011); Leo & Liu, *supra* note 121, at 395; William D. Woody et al., *Effects of False Evidence Ploys and Expert Testimony on Jurors, Juries, and Judges*, COGENT PSYCH., May 2018, at 1, 15; William D. Woody et al., *Comparing the Effects of Explicit and Implicit False-Evidence Ploys on Mock Jurors' Verdicts, Sentencing Recommendations, and Perceptions of Police Interrogation*, 20 PSYCH. CRIME & L. 603, 612 (2014).

312. *See* Blandón-Gitlin et. al., *supra* note 311, at 247; Leo & Liu, *supra* note 121, at 383–84.

313. D. Brian Wallace & Saul Kassin, *Harmless Error Analysis: How Do Judges Respond to Confession Errors?*, 36 LAW & HUM. BEHAV. 151, 156 (2012); *see* Woody et al., *Comparing the Effects of Explicit and Implicit False-Evidence Ploys*, *supra* note 311, at 612.

314. Wallace & Kassin, *supra* note 313, at 156.

315. Saul M. Kassin, *Why Confessions Trump Innocence*, 67 AM. PSYCH. 431, 434 (2012).

316. For discussion of the limited utility of cross-examination in testing the government's evidence, see Lisa K. Griffin, *False Accuracy in Criminal Trials: The Limits and Costs of Cross-Examination*, 102 TEX. L. REV. 1011, 1023–41 (2024).

317. Jeffrey Bellin, *The Evidence Rules That Convict the Innocent*, 106 CORNELL L. REV. 305, 331 (2021).

318. *Id.* (quoting 30B CHARLES ALAN WRIGHT & JEFFREY BELLIN, FEDERAL PRACTICE & PROCEDURE § 6773 (2020)).

319. *Id.* at 346.

The *corpus delicti* (literally, "body of the crime") rule plays a regulatory role, but only to a modest extent.[320] The rule allows a confession to be admissible only if there exists independent evidence—other than the confession—that the charged crime occurred.[321] However, it does not consider whether the confession was coerced, require that a confession be reliable, or corroborate that a defendant committed the charged crime—requiring only that a crime occurred.[322] As a result, as one commentator recognized, "the only unreliable confessions the rule screens out are confessions to nonexistent crimes."[323]

A trial might entail two other interventions. One is the use of an instruction informing the jury of how it is to assess the voluntariness of a defendant's confession.[324] However well-intended such an instruction might be, the jury's determination will still be subject to the common pitfalls discussed earlier.[325]

Another possible intervention is the use of expert testimony. In light of the unique challenges associated with discerning deepfakes, the defense could well convince the trial court that the standard prescribed by Rule 702 is satisfied—that "the expert's scientific, technical, or other specialized knowledge will help the trier of fact to

---

320. *See* David A. Moran, *In Defense of the Corpus Delicti Rule*, 64 OHIO ST. L.J. 817, 817 & n.1 (2003).

321. *Id.*

322. *See* State v. Mauchley, 67 P.3d 477, 484 (Utah 2003) (stating that the rule does "nothing to ensure that a particular defendant was the perpetrator of a crime").

323. Eve Brensike Primus, *The Future of Confession Law: Toward Rules for the Voluntariness Test*, 114 MICH. L. REV. 1, 53 n.285 (2015). Professor Primus notes that some jurisdictions have adopted a trustworthiness rule providing that the government cannot introduce a confession unless it provides substantial independent evidence tending to establish its trustworthiness. *Id.* She adds, however, that the provisions "have been impotent in practice." *Id.*; *see also* Moran, *supra* note 320, at 852 (describing the trustworthiness rule as "so malleable that almost any independent evidence of anything can serve to 'corroborate' the confession").

324. For one such instruction, see State v. Griffin, 262 A.3d 44, 82–83 n.29 (Conn. 2021):

> The test of voluntariness is whether an examination of all the circumstances present surrounding the rendering of the statement shows that the conduct of the police was such as to overbear the defendant's will to resist and resulted in a statement that was not truly self-determined . . . . Whether the statement was coerced means considering . . . whether it was forced or compelled out of the defendant by abusive conduct, by promises, implied or direct, or by deceit or artifice by the police [that] overbore the defendant's will to resist and critically impair[ed] his capacity for self-determination and, thus, brought about a statement that was not freely self-determined.

325. *See supra* notes 306–12 and accompanying text.

understand the evidence or to determine a fact in issue."[326] Experts are employed by courts to counteract potential juror misperceptions,[327] such as with the risks associated with eyewitness identification.[328] With respect to confessions, research shows that experts are more impactful than jury instructions.[329] Psychologist William Woody and his colleagues, for instance, assessed the impact of FEPs with and without expert testimony, explaining their impact on inducing false confessions.[330] Confirming other research showing that triers of fact can recognize the deception and coercion resulting from use of FEPs, yet nonetheless lend weight to the confessions yielded,[331] the study concluded that expert testimony made them less likely to convict.[332] Courts, moreover, are mixed in their receptivity to such testimony.[333]

4. Appeals

Finally, the appellate process holds little promise of curbing coerced confessions. This is because the Supreme Court in *Arizona v. Fulminante*[334] held that the improper admission of a coerced confession is subject to harmless error analysis by appellate courts.[335] In *Fulminante*, as two commentators observed, the Court placed "great faith in the ability of a jury to properly evaluate a confession and the evidence about how it is obtained."[336] However, the very idea that improper admission of a confession can constitute harmless error

326. FED. R. EVID. 702(a); *see also* Delfino, *supra* note 163, at 334 ("Although jurors and judges may have a general awareness that deepfakes exist, understanding the processes by which digital audiovisual images, fake or real, are created is well beyond the knowledge of most judges, jurors, and lawyers.").

327. *See, e.g.*, State v. Robinson, 431 N.W.2d 165, 172–73 (Wis. 1988).

328. *See, e.g.*, United States v. Rodriguez-Felix, 450 F.3d 1117, 1124 (10th Cir. 2006); People v. Boone, 91 N.E.3d 1194, 1199–200 (N.Y. 2017); State v. Lawson, 291 P.3d 673, 696 (Or. 2012); Workman v. State, 771 S.E.2d 636, 638 (S.C. 2015).

329. *See, e.g.*, Dayna M. Gomes et al., *Examining the Judicial Decision to Substitute Credibility Instructions for Expert Testimony on Confessions*, 21 LEGAL & CRIMINOLOGICAL PSYCH. 319, 319 (2016); Angela M. Jones & Steven Penrod, *Can Expert Testimony Sensitize Jurors to Coercive Interrogations Tactics?*, 16 J. FORENSIC PSYCH. PRAC. 393, 393 (2016); Skye A. Woestehoff & Christian A. Meissner, *Juror Sensitivity to False Confession Risk Factors: Dispositional vs. Situational Attributions for a Confession*, 40 LAW & HUM. BEHAV. 564, 565 (2016).

330. *See generally* Woody et al., *Effects of False Evidence Ploys and Expert Testimony*, *supra* note 311.

331. *See supra* notes 310–12 and accompanying text.

332. Woody et. al, *Effects of False Evidence Ploys and Expert Testimony*, *supra* note 311, at 1.

333. Brandon L. Garrett, *Contaminated Confessions Revisited*, 101 VA. L. REV. 395, 425–29 (2015).

334. 499 U.S. 279 (1991).

335. *Id.* at 285.

336. Hollida Wakefield & Ralph Underwager, *Coerced or Nonvoluntary Confessions*, 16 BEHAV. SCI. & L. 423, 437 (1998).

is at odds with research discussed above demonstrating the major impact of confessions in criminal cases[337] and the inability (or unwillingness) of jurors to avoid attaching weight to coerced confessions in their guilt determinations.[338]

CONCLUSION

Over the past half-century, research has made great strides in illuminating the psychological and behavioral impact of various police practices. This Article has focused on the impact of one such practice: police creation and use of false evidence in securing confessions from their interrogation subjects. In 1969, in *Frazier v. Cupp*,[339] the Supreme Court condoned a relatively unsophisticated false evidence ploy—a lie to a suspect that another individual implicated them in a killing—and held that the lie did not violate due process because it did not coerce the resulting confession.[340] Today, over fifty years hence, a large body of research demonstrates the coercive impact the FEP has on inducing confessions among innocent and guilty individuals alike.[341] The coercion is especially pronounced when police present fabricated tangible evidence, rather than merely orally relating its existence to a suspect (as in *Frazier*).[342]

This Article has addressed the likely-coming police use of a uniquely powerful innovation: deepfakes, which research shows are very difficult to detect, and which will become even more difficult to detect as technology progresses, making them ever more likely to induce confessions. Building upon several decisions handed down

---

337. *See supra* notes 52–64 and accompanying text. As Professor Kassin has recognized:

> The very notion that a confession error can prove harmless when other evidence is sufficient to support conviction is flawed because it rests on the assumption that the alleged other evidence is independent of that confession. It is not. What wrongful convictions have shown is that the confession becomes the foundation in a house of cards. Upon it, other faulty evidence is built. One cannot later extract the confession and declare the rest of the evidence independently corroborative.

KASSIN, *supra* note 60, at 281. Previously, it is worth noting, the Court had expressed concern over this indivisibility, stating that "[i]t is now axiomatic that a defendant in a criminal case is deprived of due process of law if his conviction is founded, in whole or part, upon an involuntary confession." Jackson v. Denno, 378 U.S. 368, 376 (1964).

338. *See supra* notes 306, 309–12 and accompanying text; *see also* Arizona v. Fulminante, 499 US 279, 296 (1990) (recognizing that "confessions have profound impact on the jury, so much so that we may justifiably doubt its ability to put them out of mind even if told to do so" (quoting Bruton v. United States, 391 U.S. 123, 139–40 (1968) (White, J., dissenting))).

339. 394 U.S. 731 (1969).

340. *See id.* at 739.

341. *See supra* notes 113–59 and accompanying text.

342. *See supra* notes 113–26, 136–39 and accompanying text.

since *Frazier* deeming police creation and presentation of fabricated tangible evidence to violate due process, the Article makes the case for a per se prohibition of deepfakes in interrogations. It does so based on the large body of research conducted post-*Frazier* demonstrating the coercive impact of police FEPs,[343] the unprecedented verisimilitude of deepfakes,[344] and the major systemic changes in the criminal legal system over the past decades increasing the pressure on individuals to confess.[345]

Although the Supreme Court long ago recognized that the voluntariness requirement of due process must evolve in accord with changes in police interrogation methods,[346] it has not done so, and a retooling is long overdue.[347] For inspiration, the Court can look to its recent Fourth Amendment caselaw, such as evidenced in *Riley v. California*,[348] where the Court held that police must obtain a warrant to search a cell phone but not a non-digital receptacle, reasoning that the vastly greater storage capacity of a cell phone justified different doctrinal treatment.[349] To the majority, comparing digital and non-digital storage capacity was "like saying a ride on horseback is materially indistinguishable from a flight to the moon. Both are ways of getting from point A to point B, but little else justifies lumping them together."[350] This Article maintains that a similar analogy applies to police creation and use of AI-generated deepfakes in interrogations.

---

343. *See supra* notes 113–59 and accompanying text.

344. *See supra* notes 164–71 and accompanying text.

345. *See supra* notes 186–26 and accompanying text.

346. *See* Spano v. New York, 360 U.S. 315, 321 (1959) (stating that as "methods used to extract confessions [become] more sophisticated, our duty to enforce federal constitutional protections does not cease").

347. As Seventh Circuit Court of Appeals Judge Ilana Rovner has noted, the holding in *Frazier* "developed in a factual framework in which [it was] presumed that the trickery and deceit used by police officers would have little effect on the innocent." Dassey v. Dittmann, 877 F.3d 297, 332 (7th Cir. 2017) (en banc) (Rovner, J., dissenting); *see also* George C. Thomas III, *Regulating Police Deception During Interrogation*, 39 TEX. TECH L. REV. 1293, 1295 (2007) (recognizing that "[t]he law of confessions has not developed a doctrine that is appropriately sensitive to the possibility that deception can have a coercive effect"). On the "science lag" in criminal procedure more generally, see Jennifer E. Laurin, *Criminal Law's Science Lag: How Criminal Justice Meets Changed Scientific Understanding*, 93 TEX. L. REV. 1751, 1753–56 (2015).

348. 573 U.S. 373 (2014).

349. *Id.* at 384–86.

350. *Id.* at 393; *see also* Carpenter v. United States, 138 S. Ct. 2206, 2219 (2018) (recognizing the "world of difference between the limited types of personal information" contained in bank records and telephone numbers dialed, compared to the information embodied in cell site location information, enabled by use of personal cell phones, refusing to hold that the third-party doctrine applies to the latter).