Journal of Public Policy (2025), page 1 of 23
doi:10.1017/S0143814X25000030

RESEARCH ARTICLE

Police interrogation and fraudulent epistemic environments

Luke William Hunt

University of Alabama, Department of Philosophy, Box 870218, Tuscaloosa, AL 35487-0218, USA
Email: lwhunt@ua.edu

(Received 21 May 2024; revised 23 December 2024; accepted 2 January 2025)

Abstract

The police are required to establish probable cause before engaging in custodial interrogation. Much custodial interrogation relies on a fraudulent epistemic environment (FEE) in which the police knowingly use deception and dishonesty to gain an advantage over a suspect regarding a material issue, injuring the interests of the suspect. Probable cause, then, is a sort of evidentiary and epistemic standard that legally justifies the police's use of deceptive and dishonest custodial interrogation tactics that are on par with fraud. However, there are both deontological and consequentialist considerations that show why the police's use of an FEE is often unjustified. Accordingly, the paper argues that even if the use of an FEE is based on probable cause, there are other (non-epistemic) reasons to think evidence with probative value (such as a confession) should be excluded when derived from an FEE and there is no acute threat of harm to others.

Keywords: criminal procedure; consent; fraud; legitimacy; policing; probable cause

Introduction

Wilhelmina Hicks awoke and found her four-month-old son, Matthew, unresponsive, on September 21, 2008. Hicks and her husband, Adrian Thomas, took Matthew to the emergency room, where a blood test showed that Matthew was in septic shock. A subsequent CT scan revealed fluid collections in the brain and Matthew was put on life support. There was no evidence of skull fracture, but an emergency room doctor suggested to the police that the baby had a fracture and had been murdered (People v. Thomas 2014).

The police did not know what happened to Matthew but theorized that Thomas had thrown and killed Matthew. They took custody of a distraught Thomas (who was not yet aware that Matthew would die), read him his rights, and proceeded to interrogate Thomas for a total of 9.5 hours over two sessions. The police repeatedly lied and deceived Thomas regarding medical, legal, and other facts (falsely representing that his wife had blamed him for Matthew's injuries and falsely

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

representing that Matthew was still alive, for example) and promised Thomas that he could go home if he confessed. They also threatened to arrest Thomas's wife if Thomas didn't come clean.

Thomas eventually confessed to harming Matthew – agreeing to "take the fall for his wife," even though he didn't think she would hurt their son – and was convicted of murder. He was subsequently retried without the confession evidence and acquitted. However, by this time, Hicks had divorced him, and Thomas had spent six years in prison aways from his five other children. It is now clear that Matthew died of septic shock.

One might describe Thomas's confession as having *probative value*: the confession is evidence that makes it more likely that the state will meet its burden of proof regarding relevant facts in the investigation and prosecution a murder (see Gardiner 2019). Thomas's scenario is common in policing, and it raises a number of legal, epistemological, and practical questions. This paper focuses on a question that has received relatively little attention: Does evidence with probative value derived from deceptive and dishonest police interrogation contribute to epistemic injustice and provide grounds for the exclusion of such evidence?[1] The paper answers the question by relying on the following considerations:

First, it is generally necessary for the police to establish *probable cause* (an evidentiary requirement that generally must be met for the police to make an arrest, conduct a search, or receive a warrant – based on reasonableness and "fair probability") before engaging in *custodial interrogation*.[2] Second, custodial police interrogation relies on what I call a *fraudulent epistemic environment* (FEE). Probable cause is thus a sort of evidentiary and epistemic standard that legally justifies the police's use of deceptive and dishonest custodial interrogation tactics that are on par with fraud.[3] Third, however, even if the use of an FEE is based on probable cause, norms regarding voluntariness, harm, legitimacy,[4] and bargains – along with consequentialist concern – support the conclusion that evidence with

[1]The broader issue of deceptive police interrogation has been addressed in legal literature such as Etienne and McAdams (2021), Khasin (2009), Magid (2001), and Skolnick and Leo (1992).

[2]See, for example, 4th Amendment, U.S. Constitution; Illinois v. Gates (1983). Courts have been reluctant to describe probable cause in terms of numbers, but judicial opinions suggest that probable cause is something less than "preponderance of the evidence" (proof that a fact is more likely than not to be true); accordingly, probable cause probable doesn't require an over-50% likelihood, though it is a significant threshold and certainly more than, say, "reasonable suspicion" (the standard required for the police to "stop and frisk" a person; see Terry v. Ohio (1968)). As Gardiner (2019) rightly notes, probable cause (and "reasonable suspicion") are epistemic standards, though are "not verdictive, and so to satisfy the standards, evidence need not eliminate classes of error possibilities (e.g., the police can "permissibly detain and arrest more individuals for a crime than they know can be guilty of that particular crime."). Ibid., 301.

[3]Within the legal context, fraud can be described generally as deception or dishonesty to gain an unjust advantage that injures the rights and interests of others. Such definitions are broad (and somewhat question-begging), in part because there are so many different types of fraud. The basic idea is that fraud is an "unjust" use of deception and dishonesty inasmuch as it "injures the rights and interests of others." What kind of rights and interests? I address this question in sections 3 and 4 by showing that deceptive and dishonest interrogation tactics can undermine one's right to only be subjected to legal harm that results from one's voluntary actions. For legal analysis of fraud, see, for example, Green (2006).

[4]For work on policing and legitimacy, see Monaghan (2021), along with my earlier work on the liberal limits in policing in *The Retrieval of Liberalism in Policing* (2019) and elsewhere.

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

probative value (such as a confession) should often be *excluded* when derived from an FEE.[5]

Although my argument has some moving parts, it takes the form of a fairly straightforward *reductio ad absurdum* framed by two hypothetical syllogisms. The paper seeks to prove *p*: *Even if the use of an FEE is based on probable cause, evidence with probative value (such as a confession) should be excluded when derived from an FEE and there is no acute threat of harm to others*.

The first hypothetical syllogism is considered in section 3, where the paper broadly argues that for political communities (and the institutions of which they are made, such as the police) to be viable, they must account for human vulnerability (including to fraud). To account for human vulnerability, the police institution must provide a non-FEE. For the police institution to be viable, then, it must provide a non-FEE.

The second hypothetical syllogism is considered in section 4, where the paper argues that if a person's act is based on the police's fraud, then the person's act is not voluntary and consensual. If a person's act is not voluntary and consensual, then the person should not be subject to harm resulting from the person's act. If a person's act is based on the police's fraud, then the person should not be subject to harm resulting from the person's act.

After considering these points, the paper turns to objections in section 5 by assuming *not-p*: *Evidence with probative value (such as a confession) should not be excluded when derived from an FEE in which the police have probable cause because*: (1) political institutions (such as the police) can viably embrace an FEE given that such environments facilitate the important goal of law enforcement, and (2) a person should be subject to harm resulting from the person's act even if the person's act is based on the police's fraud given the important goal of law enforcement – as when the police use deception and dishonesty against a person to obtain evidence with probative value against the person and others.

From the assumption (*not-p*) we would have to conclude something like *q*: *The police are justified in fraudulently interrogating people as needed to enforce the law when they have probable cause*. The paper closes by considering why *q* is morally and practically unacceptable for both deontological and consequentialist reasons, concluding that *p* must be right after all: *Even if the use of an FEE is based on probable cause, evidence with probative value (such as a confession) should be excluded when derived from an FEE and there is no acute threat of harm to others*.[6]

[5]For related arguments, see Lackey's (2023) illuminating book – published as a I completed this paper – which considers how deceptive interrogation can entail a distinctive epistemic wrong (namely: state power targeting the epistemic agency of its citizens and extracting false testimony) perpetrated against suspects, defendants, witnesses, and victims.

[6]As will be discussed, my conclusions about police interrogation specifically build upon my prior work regarding the constraining principles of the rule of law generally. See Hunt 2019, chapter 5 (setting forth a "prerogative power test" (PPT) regarding the justification of rule of law departures by the police); Hunt 2023; and Hunt 2024, chapter 2 (arguing that good faith is a normative foundation of the police institution and applying the PPT to police deception and dishonesty generally). Where my prior work emphasized what I take to be normative foundations of the police institution (good faith, the rule of law, and so on), this paper instead focuses on two related but distinct issues: (1) how police interrogation specifically can create the sort of epistemic environment (namely, a fraudulent one that precludes voluntariness) that renders deception

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

The paper begins with a sketch of epistemic injustice and how police interrogation works.

Epistemic injustice: how police interrogation "Works"

The police relied on the following tactics – all of which are common – to obtain Thomas's confession: (1) misrepresentation of the nature of Thomas's situation, (2) misrepresentation of the evidence they have about Thomas, and (3) misrepresentation the severity of Thomas's alleged criminality (including legal severity). Which of these tactics, if any, constitute epistemic injustice that might justify the exclusion as evidence of Thomas's confession, in part because the tactic constitutes a legal injustice?

Epistemic injustice is a broad term and I use it broadly in this paper: any number of practices relating to knowledge that are an affront to one's social, legal, or moral status. For example, lying to a particular group of people (say, women) based on the belief that women are not entitled to the truth would qualify as epistemic injustice.[7]

In other cases, there might seem to be no injustice when we deceive people – such as criminal suspects. For instance, if one lies to a murderer at the door about the location of the murderer's intended victim, such a lie strikes most people as justified considering the important goal of protecting a person from being murdered (but see Kant 1996). The idea is that the murderer is not entitled to the truth given the competing value of the victim's life (see, e.g., Williams 2002).[8] Likewise, one might think there is no epistemic injustice when the police lie to suspects for good reasons.

What sorts of reasons? I will assume that one plausible reason is law enforcement based on Lockean accounts of epistemic justification – roughly, accounts that explain justification in terms of high probability (see, e.g., Sturgeon 2008; Bergmann 2006). We should thus consider standards of probability relevant to police interrogation. In fact, there is no legal (constitutional) standard that must be met in non-custodial police interrogation – for example, interrogation that occurs when a person is free to leave the interrogation.

On the other hand, a person has a constitutional right to receive a "Miranda warning" when the police engage in a custodial interrogation – in other words, interrogation in which the police have detained or arrested a person and the person is not free to leave. Although there is no epistemic standard that governs the police's

and dishonesty unjustified (despite the policing establishing probable cause), and (2) how such unjustified interrogation tactics might be addressed at the level of *policy* – in other words, policies that would *exclude* certain confession evidence as a *remedy* for unjustified interrogation (including through legislation prohibiting some of the interrogation tactics in the first place).

[7]As will be discussed, I do not consider all deception and dishonesty to be epistemic injustice, including because all persons are deceptive to some degree. Although hermeneutical or testimonial injustice are the paradigmatic cases that purportedly give rise to the phenomena of epistemic injustice, my point is that deception and dishonesty can fall into the category of epistemic injustice. See, for example, Fricker (2007), for a compelling account of what I take to be a standard view of epistemic injustice.

[8]For example, Williams (2002) argues that such people (a murderer at the door, say) have given up their right to the truth because they themselves have turned their backs on trust and mutual respect. I address this point in section 5, arguing that even interrogation rising to the level of fraud is justified in exigent circumstances in which there is an acute threat of serious harm to others (similar to how physical violence can be justified in exigent circumstances).

https://doi.org/10.1017/S014381 4X25000030 Published online by Cambridge University Press

custodial interrogation tactics *per se*, the police are generally required to satisfy the evidentiary standard of *probable cause* to maintain *custody* (arrest) over a person.[9] Roughly, then, we can say that it is generally necessary for the police to establish probable cause before engaging in custodial interrogation.

Although the idea of police interrogation may bring to mind brutal phonebook beatings (as depicted in a television series such as *The Shield*), interrogation is of course a more subtle process based on a variety deceptive tactics. The television show *The Wire* gets closer the truth in an episode in which the police tell a suspect (falsely) that the suspect's friend and alleged co-conspirator confessed to the crime in question. The police then put the suspect's hand in a copy machine – telling him it is a sophisticated lie detector – prompting the man to confess. Such fictional examples seem extreme, but we will see that what happens in the real world is sometimes stranger than fiction.

False narratives

Interrogation tactics such as the "Reid technique" rely upon false narratives to prompt confessions. Law and psychology researchers have suggested that such tactics "convey . . . [anti-*Miranda*] messages *implicitly*, through *pragmatic implication*, without violating the letter of the law" (Davis and Leo 2012; Hunt 2016). For example, before a custodial interrogation begins, law enforcement officers create an "illusion of voluntariness" that lessens both the importance of the situation as well as *Miranda* warnings; the prelude to the interrogation might run as follows: "We need to talk to you to get 'this thing' 'straightened out', but before we can do that we need to get this form out of the way" (Davis and Leo, 356–58).

The idea, then, is that the law enforcement officers have implied something such as the following: "Your situation is not too bad; everything will be okay after we complete this formality and talk to you—and we *are* going to talk to you." After the interrogation proper begins, officers may imply that speaking to the police is in the suspect's best interests, while failing to confess will damage the suspect's opportunities (Ibid.). The "set-up question" asks the suspect whether she thinks leniency should be shown to the person who committed the crime in question, with the idea that the suspect will potentially infer that the consequences of the crime in question are not set in stone and leniency is an option.

From here, the officers might state clearly that the suspect's guilt has been established and then imply that the whole point of the interrogation is to "determine what will happen to the suspect as a *result of his guilt*, rather than an investigation as to *whether* he is guilty" (Ibid., 359–61). Finally, officers may indicate that they can aid the suspect's legal predicament if she will acknowledge her guilt. This is just a snapshot, but you get the point.

Of course, the specific tactics and the specific false narrative on which the police rely will vary from case to case – depending on the sort of deception to which the police believe a suspect will be susceptible. As we will see, the interrogators in the *Thomas* case repeatedly reassured Thomas (falsely) that they understood his son's

[9]It should be noted that *Miranda* warnings are generally required for custodial interrogation whenever a person is not free to leave, regardless of whether the police have formally arrested the person.

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

injuries to have been *accidental*. They said they were not investigating what they thought to be a *crime* (or that Thomas could be arrested) and that once Thomas had told them what had happened, he could go home. These points illuminate the convergence of false narratives with false facts and false laws.

False facts

In the *Thomas* case, when Thomas continued to deny having hurt Matthew, even accidentally, the officers falsely represented that his wife had blamed him for Matthew's injuries. They also threatened that, if he did not take responsibility, they would "scoop" his wife from the hospital and bring her in because one of them must have injured the child.

These sorts of false factual representations are not uncommon. Consider notorious cases (such as the "Central Park Five" case) in which the police tell a suspect that they have the suspect's fingerprints at the crime scene, though no such evidence exists (see also Commonwealth v. Selby 1995). The key legal issue in such cases is whether a suspect's incriminating statements are "voluntary" based on the totality of the circumstances (see generally People v. Palmer 2001; 18 U.S. Code § 3501). This is a difficult question because a suspect will often change their story (and make incriminating statements) *after* the police lie to them about the evidence. Courts have interpreted "voluntary" to mean a product of a "rational intellect" and a "free will," which perhaps raises more questions than answers. In any case, the state has the burden of proving – beyond a reasonable doubt – that the suspect's statements were made voluntarily.

How is this established? Courts have simply reasoned that *misinformation alone* is not sufficient to show involuntariness. Beyond the use of lies and phony props (photocopies of fingerprints, say) by police officers in an interrogation room, what more is needed? Not much – as long as the suspect was read his Miranda rights and there is no evidence that the suspect was incapacitated or incompetent during the interrogation. In short, the police will generally have a legal right to lie and use fake evidence as long as the suspect has been properly Mirandized, is not incompetent, and is not, say, high or drunk.

I will proceed with the assumption that this legal standard is normatively inadequate. What we need is clarity about which kinds of factual deception render an environment fraudulent. To be sure, interrogations cannot be fully honest because no interpersonal relationship requires full and complete disclosure (this includes disclosure requirements in the sexual and medical consent context). In the interrogation context, the contours of factual deception and dishonesty are even more nebulous. Consider the "good-cop routine" in which an officer feigns sympathy and the desire to help a suspect. Or consider an atheist officer who misleads a suspect about what God would want them to do. It is not difficult to think of countless other examples along this spectrum of deception and dishonesty. My solution is straightforward: We draw the line as closely as possible to FEEs.

Feigning interest in a suspect's wellbeing – or disingenuously appealing to God's will – are not objectionable, then, because the deception is not *material* (see generally Neder v. United States 1999; United States v. Morris 1996). By this I mean that a suspect's reliance upon such deception and dishonesty is a collateral (not

https://doi.org/10.1017/S014381 4X25000030 Published online by Cambridge University Press

material) component of the exchange between the officer and the suspect. Vague statements regarding things such as God's will are, presumably, not easily ascertainable – at least compared to (false) statements by the police that a suspect has been implicated by another person, or that the police have physical evidence linking the suspect to a crime. The latter are material in the traditional, legal understanding of the word, while the former (feigned concern or disingenuous statements about God's will) are more collateral and akin to what we call "puffery" (see Shiffrin 2014, p. 188–91; Hunt 2023). Although a business may not engage in fraudulent advertising (misrepresenting specific facts), businesses are free from liability when using puffery advertising tactics that invoke (exaggerated) opinions about the quality of their product (Restatement (Second) of Contracts, § 168 1981).

The point is that the examples that have been discussed are less like puffery (telling a suspect, insincerely, "I don't think you should lose sleep over what you did") and more like *fraud* – as when the police tell a suspect (falsely): "We have your blood at the crime scene," when no such evidence exists. We can thus begin to see a principled framework about which kinds of factual deception render an environment fraudulent. Now consider the idea of misrepresenting the law and a suspect's legal situation.

False laws

In the *Thomas* case, the interrogators eventually acknowledged (in a post-interrogation hearing) that they did not even have probable cause to arrest Thomas, and yet they represented (falsely) that they were the defendant's last hope in forestalling criminal charges. Consider how this makes the *Thomas* case a uniquely apt example.

One of the interrogating officers advised Thomas that "he was doing all he could to stop his superior from having [Thomas] arrested." This – coupled with the fact that Thomas was interrogated for a total of 9.5 hours (an initial 2-hour session and a subsequent 7.5-hour session) – strongly implies that Thomas was *not* free to leave. In other words, even if Thomas had not been formally arrested, the evidence strongly suggests that it would be reasonable for Thomas to think – at various points in the interrogation – that he was not free to leave. It was only in a subsequent hearing that the police acknowledged that they did not have probable cause to arrest Thomas prior to receiving his confession.

Most importantly, regarding Thomas's subsequent motion to suppress his confession, the relevant Appellate Division upheld the denial – ruling that the state had met its burden of proving that Thomas's confession was voluntary beyond a reasonable doubt and that the police's fraudulent interrogation did not offend due process. The decision was only reversed on appeal to New York's highest court, resulting in a new trial. By this point, however, Thomas's life had been turned upside down (conviction, incarceration, divorce, and so on). The *Thomas* case thus illustrates three crucial points.

First, whether or not the police's fraudulent interrogation tactics are legally justified is often a live question, even if the police imply that they have probable cause. Second, there are compelling reasons (considered in the following sections) to think that fraudulent police interrogation tactics are often unjustified considering

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

basic principles of legal and political morality, even if the police have probable cause. Third, probable cause itself, then, is an inadequate epistemic standard for accessing whether fraudulent police interrogation tactics are fully (legally and morally) justified because heightened probability is not always compatible with important norms of political morality. Something more is needed to justify fraudulent interrogation tactics.[10]

In other cases, the police have lied about a suspect's criminal exposure to obtain a confession – in other words, telling a suspect (falsely) that their alleged actions do not constitute a crime (see, e.g., State v. Walker 1992). Courts have upheld such lies about the law. One might think it is just to uphold these misrepresentations of law based on the well-established *principle of legal ignorance*, which stands for the idea that *ignorance or mistake of the law is no excuse*. The principle is based on the idea that we are responsible for knowing our legal duties. One argument in favor of this principle is that there would be lawlessness without the principle because people would claim ignorance of the law all the time. Whatever you think about that argument, note the important distinction in the cases in question: A defendant (reasonably) relies on the police – whose job it is to enforce the law – to explain the nature of the offenses in question.

Indeed, there is a *reasonable reliance* legal doctrine, standing for the idea that a person will be excused because of a mistake of law if the person reasonably relies on an official (but erroneous) statement of the law obtained from a public official with responsibility for the interpretation, administration, or enforcement of the law (see, e.g., Model Penal Code 2.04 (American Law Institute 1962)). Although the legal doctrine may not be perfectly apt in the case above (because, say, the police are relatively low-ranking state officials), the underlying principle is clearly relevant (Hunt 2024). If you cannot trust the state to tell you what conduct is and is not a violation of a particular law, who can you trust?

Given that police are tasked with enforcing the law, it stands to reason that people are entitled to a non-fraudulent account of the truth in these matters – in other words, an account in which the police do not intentionally mislead a person being interrogated about a material fact in a way that harms the person. And given that the police are not held to this standard – because, say, the confession is interpreted as "voluntary" or a product of "free will" – it is not difficult to see why the police institution suffers from a lack of trust. This raises the idea of a FEE.

Fraudulent epistemic environments

I'll try to motivate the following claim in this section: For political communities (and the institutions of which they are made, such as the police) to be viable, they must account for human vulnerability (including vulnerability to fraud). Of course, to be viable, such institutions need not account for the vulnerability of *all* members of society; to be sure, the police institution has often (historically and today) focused

[10] I will suggest that this "something more" is a sort of public safety exception, that is, fraudulent interrogation is justified only in rare cases in which it is needed to address emergencies involving acute threats of harm. And note that one upshot of my analysis is that FEE interrogations *without* probable cause are clearly unjustified.

https://doi.org/10.1017/S014381 4X25000030 Published online by Cambridge University Press

on protecting some groups but not others. The point is that we should be cognizant of when the police *do* use their discretion to enhance (not mitigate) vulnerability.

With this in mind, I consider how accounting for human vulnerability entails that the police provide a non-FEE. For the police institution to be viable, then, it must provide a non-FEE for some people and groups – and it should give us pause when the institution fails to do so for some but not others.

Vulnerability and viable political institutions

When we think about the nature of social institutions, it stands to reason that we should keep in mind fundamental truisms – natural facts and natural necessity – regarding what it means to be a human in the world. Without, say, institutions that protect one from deceit and fraud, there would be no way of knowing whether one will be taken advantage of when one trusts and relies on others. So trusting and relying on others is necessary in political communities. And for law and legal systems to be viable, they must minimally protect at least some people from their human vulnerability – including vulnerability to things such as fraud. Why? A legal system that fails to address human vulnerability is not a legal system, but more akin to what is referred to as the state of nature (see Hart 2012, p. 193–200).[11]

For example, sanctions are necessary to ensure that one's trust and cooperation will not be taken advantage of by others – at least, given the limits of human understanding and strength of will. The idea is simply that societal rules must provide some degree of sanction when people choose to defect from the rules that protect human vulnerability. We are not justified in expecting perfect compliance with the rules, but we are justified in expecting recourse when one fails to act with a basic disposition good faith that precludes fraud.[12]

This is apparent at both the individual level and the institutional level. If you pay me five dollars in exchange for my agreeing to give you a bottle of beer after I brew it, then you must have some assurance that I won't simply take the money and run – defraud you, in other words. If there is no recourse when people defect from social rules, there would be no point in having social rules because the rules would not protect people from vulnerability. Likewise, if we defer to social institutions for recourse – to enforce rules and sanction the rule breakers – then we must have some degree of confidence that the social institution itself will act non-fraudulently. Consider a few examples.

Beach flag #1: To be sure, we can acquire knowledge through learning about the world through our personal experiences. Suppose I kayak to a small, uninhabited island. I can acquire knowledge about the risk of rip currents by getting out of my kayak and gradually wading into the ocean – using my senses to acquire information and become acquainted with the features of the ocean and the risk of swimming.

But to survive in a political society, one must also be able to rely on one's relations with others to acquire knowledge. If I step onto the beach near my hotel, I must rely on the flag warning system to know whether it is safe to swim. Indeed, if there is a

[11]I examined these points – Hart's idea of the "minimum content of the natural law – in chapter 1 of *Police Deception and Dishonesty* (2024).

[12]See Hunt (2023) and Hunt (2024), chapter 2.

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

double red flag I could be arrested or fined for entering the water – meaning that I do not have the option to rely on my direct experience of the world to determine whether swimming is safe. If the flag is green, then I must be able to rely on the state's representation that the water is relatively calm (and I won't be arrested for entering the water). This would be a *reliable epistemic environment*.

Any viable political and legal system requires an epistemic environment that entails some degree of trust and mutual forbearance to account for one's vulnerability. This includes vulnerability to fraudulent relations, such as relations in which a person or entity knowingly uses deception or dishonesty to gain an unjust advantage over another person in a material component of the relation, injuring the rights and interests of the person.

Vulnerability and fraud

Beach flag #2: Imagine the alternative. Suppose the state lacks revenue and decides to charge a fee to access the beach and the ocean. Suppose further that the state misrepresents the ocean's conditions to encourage better beach attendance and generate more revenue (i.e., flying a green flag when conditions merit a yellow or red flag). If I agree to pay the fee and swim based on this representation and I am harmed, the responsibility for my decision to swim is mitigated by the state's fraudulent representation. As we will see in the next section, the decision to pay the fee and swim would not qualify as voluntary or consensual given the FEE.

Were it otherwise, we would be talking about a political society that approximates the idea of the state of nature – a society in which there is no guarantee that people and institutions will do what they say and keep their commitments (or recourse when people do not keep their commitments). Common conceptions of the state of nature, then, are the epitome of an FEE, and the value of a reliable epistemic environment is one of the chief arguments in favor of leaving the state of nature.

It is worth noting that a FEE is different from a merely *unreliable* epistemic environment. If one lives in a society in which one receives a false report every time one asks whether it is raining outside, that would be an instance of an *unreliable epistemic environment*. But it wouldn't necessarily be fraudulent, which would require an attempt to use the dishonest weather report to gain an advantage that injuries one's rights and interests.

The broader point is that we should be worried if it right to say that political institutions (such as the police institution) exacerbate human vulnerability through the use of FEEs. If we defer to political institutions for recourse – to enforce rules and sanction the rule breakers – then we must have some degree of confidence that the institution itself will operate non-fraudulently.

It is unreasonable to expect people to defer to a police institution that fails to operate non-fraudulently – in other words, an institution in which the agents themselves enhance human vulnerability by acting in bad faith, defrauding people, and undermining the rule of law. Empirical work bears this, including evidence regarding the erosion of trust in the police institution (Vallier and Weber 2021; Bradford et al. 2017), the erosion of the police's legitimacy, and the fact that people literally want to *abolish* the police (Hunt 2022).

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

In sum, viable legal institutions must account for human vulnerability by providing a non-FEE. Of course, the U.S. legal system is a far cry from the "state of nature." The U.S. and similar legal systems are viable because they protect (many) people from a wide range of human vulnerabilities (including fraud). But in examining the extent to which a legal system is "just," one relevant consideration is the extent to which the legal system provides a non-FEE for some people and groups but not others. Perhaps that is okay. Perhaps the police's use of fraudulent interrogation tactics is justified – say, when they have probable cause. The paper turns to these and related issues in sections 4 and 5.

Voluntary relations and responsibility

With the idea of an FEE in mind, now consider the relations between two persons, *A* and *B*. If *A*'s act is based on *B*'s fraud, then *A*'s act is not voluntary and consensual. If *A*'s act is not voluntary and consensual, then *A* should not be subject to harm resulting from *A*'s act. If *A*'s act is based on *B*'s fraud, then *A* should not be subject to harm resulting from *A*'s act.

Fraud and voluntariness

In Joel Feinberg's classic work on harm, *Harm to Others*, he examines principles regarding the limits of one's consent to harm:

> [The *Volenti* principle: 'to one has consented no wrong is done'] is most plausible when it denies title to complain only to him whose consent was *fully voluntary*, and a person's consent is fully voluntary only when he . . . has not been threatened, misled, or lied to about relevant facts, nor manipulated by subtle forms of conditioning. It is worth giving emphasis here to two points: that both force and fraud can invalidate consent . . . . (Feinberg, 1987).[13]

Feinberg begins with the principle suggesting that no wrong is done *A* when *A* consents to the actions of *B*. Consider how the principle is central to, say, evidentiary questions in rape investigations. If *A* accuses *B* or rape, *A* generally has no legal claim if *A* consented to sexual relations to *B*. Of course, such cases are often more complicated than that, and Feinberg's next point is thus crucial.

He writes that the principle against liability for a wrong is strongest when *A*'s consent is "fully voluntary," which would preclude cases in which *B* somehow misled *A* or lied to *A* about relevant facts. If *B* lies to *A* about *B's* mother's favorite flavor of ice cream, that might not be enough to invalidate *A*'s consent to sexual relations. On the other hand, if *B* deceives *A* about a material fact to gain *A*'s consent to sex, then *B*'s deception is fraudulent and invalidates *A*'s consent. Why? Because *A*'s consent is contingent upon *B*'s representation of the material fact. Consider three examples:

[13]Relatedly, Feinberg (1987) ties consent and the *Volenti* maxim to legitimacy, including cases in which voluntary consent to serious risks is practically impossible to confirm, as well as cases that are necessarily harmful to the interests of third parties or the public at large. Ibid., 220.

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

- *Professional fraud*: A physician or other professional advises a person that sexual consent is needed for a medical operation or diagnosis (Falk, 1998, 52–64).
- *Sexual theft*: A person agrees to pay a sex worker money in exchange for sex but refuses to pay after the sex occurs (Ibid., 76–79).
- *Abuse of formal authority*: A person uses their formal authority – such as law enforcement authority – to leverage sexual consent (Ibid., 79–84).[14]

In each of these common scenarios, person *A*'s consent to sex is contingent on person *B*'s deceptive representation of a material fact, such as medical necessity, a good faith agreement, or official authority. Accordingly, if *A*'s act of consent is based on *B*'s fraud, then *A*'s act of consent is not voluntary.

Voluntariness and harm

Suppose the police interrogate a suspect in a drug case; the police's goal is to use their leverage over the suspect to obtain a confession. The police tell the suspect that she is facing a 25-year prison sentence, but that they will recommend a reduced prison sentence to the prosecutor and judge if the suspect confesses. The suspect confesses.

In fact, however, the drug charge the suspect was facing would have made her guilty of a crime for which the sentence range was only 3–5 years. Based on these facts, the police have straightforwardly obtained the suspect's confession by fraud, regardless of whether courts sanction the use of such trickery by the police.[15] It is true that the suspect's lawbreaking was a relevant cause of any harm to which the criminal justice system subjects her. However, it can also be true that the suspect's (involuntary) confession (to the extent that it was obtained through fraud) is a harm for which the police are responsible (to the extent that we think people are fully responsibly only for acts that were truly voluntary). In almost any other context, the suspect would not have been responsible for the harm resulting from her confession because her act was not voluntary and consensual – her act was based on a fraud.

We can thus tentatively conclude that if *A*'s act is based on *B*'s fraud, then *A* should not be subject to harm resulting from *A*'s act.[16] What type of harm is relevant

[14]See, for example, Philipps (2015) for reporting on a former Oklahoma City police officer who was convicted of multiple charges – including first-degree rape and oral sodomy – stemming from the officer's use of his authority to prey on vulnerable women.

[15]For a similar, real-world case in the context of agreements between the police and informants, see Alexander v. DeAngelo (2003).

[16]To be sure, some political philosophers would grant that it is sometimes permissible for, say, an innocent bystander to be harmed. This view might raise "lesser evil justifications" in the harming and defensive force literature. The topic is beyond the scope of the current paper, considering that I am approaching the justification of fraud from the perspective of a narrowly defined state prerogative power to address emergency situations (and even then, the use of such prerogative power must not be an affront to personhood, and so on). Regardless, I reserve judgment about the extent to which broader "lesser evil" questions regarding the state harming persons through fraud are applicable in criminal legal systems, which are constrained by foundational rights respecting voluntariness and freedom from fraud. As noted earlier, the prerogative power test I've developed in my earlier work is a framework (not a bright line rule) derived from executive prerogative power and the rule of law. The basic idea is that there are rare emergency cases in

https://doi.org/10.1017/S014381 4X25000030 Published online by Cambridge University Press

in the context of police interrogation? One example is a confession that jeopardizes one's liberty and legal interests.

Recall that custodial interrogation will typically require that the police satisfy the evidentiary burden of probable cause (given that probable cause is typically required for an arrest). Custodial interrogation also requires a Miranda warning, which identifies one type of harm explicitly: anything the suspect does say can and may be used against them in a court of law. The idea is to protect one's right against self-incrimination (Miranda v. Arizona 1966; U.S. Const., 5th and 6th Amends). Accordingly, if one's decision to open oneself to such harm is based on fraud, one should not be subject to the resulting harm caused by one's confession. The upshot is that that harmful evidence obtained by the police fraudulently should be excluded (even if it has probative value).

Objections and conclusions

A central objection to my argument is the priority of law enforcement: Political institutions (such as the police) can viably embrace an FEE given that such environments facilitate the important goal of law enforcement. Accordingly, *A* should be subject to harm resulting from *A*'s act even if *A*'s act is based on *B*'s fraud given the important goal of law enforcement. I respond to these objections from both deontological and consequentialist perspectives.

Objection: the priority of law enforcement (and a caveat)

An intuitive objection to my argument is that of course the police can viably embrace an FEE because – otherwise – it would be impossible for the police to do their job (enforce the law, centrally). It follows – so the objection goes – that *A* should be subject to harm resulting from *A*'s act even if *A*'s act is based on *B*'s fraud given the important goal of law enforcement, as when the police use an FEE against *A* to obtain evidence with probative value against *A* and others.

For example, perhaps the police are justified in misrepresenting facts and law when interrogating a suspect because they have probable cause to believe the suspect engaged in criminality and the misrepresentation will increase the probability of a confession – and a confession is important evidence facilitating the police's enforcement of the law. Relatedly, perhaps the resulting harm to the suspect (based on the suspect's confession) is justified despite the misrepresentation on which the confession is based.

There are several plausible responses to this objection. These responses will lead to the conclusion that some evidence with probative value should be excluded when it is derived from an FEE that opens one to harm and a loss of rights. However, before replying to the *priority of law enforcement objection*, let me note one caveat that will perhaps make my conclusion more palatable to those who are sympathetic to the objection.

---

which the executive is justified (as a matter of political morality) in deviating from the rule of law. For illuminating perspectives on the lesser evil justification, see, e.g., Frowe (2018) and Gordon-Solmon (2022).

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

First, let's assume that some uses of police force and violence are clearly justified, with an easy example being an active shooter in a school. Can we thus say that whenever force would be justified in a law enforcement investigation (to protect a victim from harm, make an arrest, and so on), an FEE during interrogation would be as well? Elsewhere, I have suggested that the answer to this question is "yes."[17] But I think that answer is misleading and requires more nuance, including in the context of interrogation.

In short, it is not simply that deception and dishonesty on par with fraud are justified *whenever* force is justified. Instead, my position is more specifically that fraudulent police tactics (including those involving interrogation) may be justified in part based on *emergencies* involving acute threats of harm – in the same way that some emergencies can justify the use of force.[18]

I think this is the better way to make the point because it is not clear that, say, lying is always permissible on the *basis* of the permissibility of force (Betz 1985).[19] For example, there may be possible situations in which the use of force is permissible, but lying (and even the mere use of deception) is not.[20] So it is not quite right to say that deception and dishonesty on par with fraud are justified on the *basis* that force is justified. Instead, both force and fraud can be justified in various cases involving *emergencies*. In the context of interrogation, a practical rule of thumb is that the police's use of an FEE is justified in rare cases (emergencies) in which it might be necessary to address an acute risk of harm to a person.[21]

Reply: principles and consequences

In replying to the *priority of law enforcement objection*, let's begin by returning to the beach flag example in section 3. The example was a simple illustration of how we need some (reliable) epistemic state authorities regarding our legal obligations if we

[17]See Hunt 2023 and Hunt 2024, Interlude, in which I argue for a "prerogative power test" to determine whether deception and dishonesty on par with fraud is justified.

[18]In addition to such emergency situations, the other elements of the prerogative power test (notably, the "personhood constraint"), must also be satisfied. Ibid.

[19]I thank James Mahon for alerting me to this point at the Central Division Meeting of the American Philosophical Association, New Orleans, in February 2024.

[20]Betz (1985) uses an example involving a barricaded fugitive murder, arguing that force (including deadly force) might be justified to make the arrest, but not lying to induce surrender. I confess that I am not convinced by this example because lying to the fugitive seems clearly justified, especially if it prevents the need for a deadly force encounter. Still, the general point may be correct, namely: lying is not justified *because* force is justified. In my view, lying (and force) are justified insofar as they are necessary to address an emergency involving an acute threat of harm.

[21]Note that "acute" often means "imminent," but not always. In the context of a fraudulent epistemic environment, acute is the appropriate term given the difficulty of establishing the imminence of a threat outside of real-time deadly force situations (when an imminence standard would be more appropriate). A simple (and common) example of what I have in mind: A case in which the police are interrogating a person in a kidnapping investigation in which there is an acute risk of harm to the kidnapped child; the police believe the person has knowledge of the kidnapper's identity and location (e.g., the kidnapper is a friend or relative of the person being interrogated). For the related legal doctrine ("public safety exception"), see New York v. Quarles (1984) in which the U.S. Supreme Court held that police may deviate from *Miranda rights* in "a situation where concern for public safety must be paramount to adherence to the literal language of the prophylactic rules enunciated in Miranda."

https://doi.org/10.1017/S014381 4X25000030 Published online by Cambridge University Press

are to have reciprocal, epistemic moral cooperation. Non-FEEs are necessary for viable legal and political institutions, but it does not follow that the police must be moral saints. They have difficult and (sometimes) dangerous jobs that (sometimes) require FEEs. However, considering the indispensability of non-FEEs in public life (and considering especially the police's ubiquitous presence in public life as agents of the state), it is unusual to think that the police are not only permitted to defraud you in *non-emergency* situations – but also in a way that undercuts one's ability to invoke fundamental rights.

Beach flag #3: Consider a variation of the beach flag scenario. Suppose the state puts up flags that accurately signal risk but a group of bitter local residents (who dislike tourism) repeatedly sneak onto the beach and replace the red flags with green ones. One might argue that in this scenario the police would be justified in lying to the suspects to obtain a confession. The argument is based on a dirty-hands style justification insofar as the only way to defeat those who prey on human vulnerability may be to play hardball with them.

Let me be clear that my argument is consistent with police deception and dishonesty in this beach-flag-swapping case (and most other cases). It reasonable to think that people become liable to deception and dishonesty in light of their (bad) decisions and actions. And it would be naïve to think the police can never lie to criminals and criminal suspects – who are often in the business of lying to the police – as a way to protect the public. Indeed, there is generally no legal prohibition on the public telling lies, even if lying to others is morally wrong. It would thus be implausible to hold the police to the unreasonably high standard of "no lies." Instead, my point is that we should draw the line when police lying rises to the level of fraud because fraud is very much a general legal (and moral) prohibition that is fundamental to any social institution (not to mention that there is slim empirical evidence of its law enforcement efficacy).[22] Why? Because – as argued in section 4 – consent to speak and confess (or engage in other acts) is not truly voluntary if the consent is based on fraud.[23]

This view is certainly accepted with respect to more dramatic police tactics. No one seriously argues that the police should be able to enforce the law (and obtain an (involuntary) confession) by using torture and brutality, at least given uncontroversial constraints upon the means by which the police enforce the law. These constraints include political and moral conceptions of personhood that are based on standard assumptions regarding moral agency, reciprocation, and human dignity (Hunt 2019). For instance, the state traditionally requires a "principle of guilt" to be satisfied before holding a person responsible for a crime; such principles mean that one much have engaged in a voluntary act that is connected to one's guilty mind.

These sorts of principles imply that persons are construed as moral agents who are capable of reciprocating within society, and that it would be wrong and an affront to their human dignity to punish outside of the principle. But it's more than

[22]Now, it might just be the case that an FEE is justified in *beach flag #3* – if, say, the fraud is necessary to address an emergency in which life is at stake (i.e., drowning tourists).

[23]As I have indicated, by fraud I generally mean knowing and willful behavior by the police (not "honest mistakes"), materiality, and harm to the person being interrogated.

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

that. Principles of guilt entail a commitment to legitimate governance by the rule of law, which of course precludes fraud. Convicting one based on the fraudulent whims of state officials is thus inconsistent with the principle of guilt. Likewise, principles of guilt entail bargaining and contractual norms that preclude bad faith and fraud considering that bad faith and fraud can circumvent the voluntariness of one's moral agency, including with respect to confessions.[24]

In short, uncontroversial principles of legal and political morality constrain the way the state (police) may seek the valuable goal of security and law enforcement, lending support to the position that some evidence with probative value ought to be excluded when it violates those principles.

One can also respond to the *priority of law enforcement objection* on consequentialist grounds. Although there might be some good consequences of embracing an FEE (namely, obtaining evidence with probative value, potentially lessening crime in vulnerable communities), it is plausible to think those good consequences can be outweighed by the bad ones. The question of whether to allow police to rely on FEEs for interrogation is relevant to what's at stake (see Moss 2023, 2021). Consider the problem of false confessions (and convictions).

It might seem odd for a person to confess to a crime they did not commit (assuming interrogation does not involve physical abuse). However, such confessions are all too common when a person (especially a young person) is subjected to hours-long interrogation in an FEE. This includes (false) promises (that the person can go home if they say what the police wants to hear, for example), as in the notorious Central Park 5 case.[25] Given an extreme power imbalance, it can indeed be rational for innocent people to confess and plead guilty.

Innocence Project, a nonprofit legal organization focused on exonerating people who have been wrongly convicted, has documented that false confessions (which may be based on FEEs) account for almost 30% of wrongful convictions, while researchers such as Saul Kassin (2022) have conducted extensive empirical studies regarding the connection between false confessions and police interrogation practices (see also Rakoff 2014).

And as noted, data regarding the relationship between social trust (general trust among members of a society) and legal trust (trust in legal institutions such as the police) shows that trust in the legal system (such as the police institution) is a function of social trust. One important take-away from this research is that social

[24]One might think that there would be an affront to one's personhood *any* time the police employee an FEE. However, as noted, there seems to be clear cases in which (deadly) force is justified, such that using force is not an affront to personhood in those cases (similar to justified cases of lethal self-defense, perhaps as well as the related idea of "death with dignity"). Killing one in self-defense suggests the value of personhood (protecting life), and that we may only kill others under very stringent conditions (which paradoxically respects their value as a person). One is only killed when they are responsible (moral agent) for breaking the social contract (reciprocator) by posing a serious threat to others; still, the killing must be done in a way that respects the criminal's dignity (instead of in, say, a tortuous or humiliating way). Likewise, as noted, we can say that an FEE can be justified in emergency situations, though never (even if an emergency exists) in the context of degrading one's human dignity – treating one as if they have less worth or a lower status than others (Hunt 2024).

[25]The reporting on the Central Park 5 case is voluminous. See, for example, Kassin (2002), Dwyer (2019), Harris (2019), and Quiroz (2021).

https://doi.org/10.1017/S014381 4X25000030 Published online by Cambridge University Press

and legal trust are only connected when legal officials (such as the police) are viewed as representative of most members of society. There is a deep relationship between trust, legitimacy, and justice: Legal trust is important to public cooperation with authorities such as the police, meaning that both the public *and* the police have a lot at stake (especially given calls to literally abolish the police).[26]

It is far from obvious, then, that FEEs should be normalized for the police compared to other members of society. If there is reason to think the data is reliable, it rases serious questions about whether FEEs can be justified on consequentialist grounds.

Policy implications

Although courts have sanctioned FEEs, there are at least two ways in which the use of FEEs could be curtailed in policing. First, we can hope that the *police* will police themselves by adopting internal guidelines restricting their use of FEEs. This strikes me as a naïve hope and I will not consider it further. A more promising option is steeped in the idea of democratic policing: The citizenry can implore the police to adopt internal guidelines restricting the use of FEEs, as well as implore elected representations to enact legislation prohibiting law enforcement officers from using fraudulent deception and dishonesty when interrogating people. These are viable options considering that several U.S. states have recently enacted legislation prohibiting the police from lying during interrogation when the suspect is under the age of 18.[27] This sort of legislation could be expanded such that it applies to all persons regardless of age – prohibiting common deceptive tactics such as false promises of leniency and false claims about the existence of incriminating evidence. Such prohibitions could also include the remedy discussed herein: excluding evidence derived from the FEE.

I want to be clear about the implications of such measures, including tradeoffs regarding the exclusion of probative evidence. Recall the questions that have been examined: (1) When is it permissible for police to use FEEs to obtain evidence from suspects in interrogations? (2) When is it permissible for evidence that was obtained through fraudulent interrogations and has probative value to be acted on as evidence? In short, I concluded that the answer to both questions is limited to

[26]See Hunt (2024). At the most abstract level, legitimacy is often conceived as a function of authority (acts conducted pursuant to authority are legitimate), authority a function of consent (authority obtained through consent), and consent a function of governance by the rule of law (persons explicitly or tacitly consent to governance by law). On the other hand, I've argued that fraud is inconsistent with true consent and is the antithesis of the rule of law. Less abstractly, the police's legitimacy depends in part on their being an honest source of reliable knowledge and information about the law, criminal cases, and suspects' and victims' rights. If legal institutions such as the police instead govern by fraud, there is a sense in which they fulfill their law enforcement function illegitimately, outside the rule of law. Of course, legitimacy can cut the other way, too, as when the police fail to uphold their end of the contract by failing to enforce the law. If the police's restraint (e.g., from getting their hands dirty to enforce the law) is perceived as letting guilty people go free, the legitimacy of the legal system may be undermined. Again, it is for this reason that my argument does not preclude the police playing dirty (lying), so long as it does not rise to the level of fraud (unless the fraud is necessary to address an emergency involving an acute threat of harm).

[27]See Illinois Senate Bill 2122 (2021), Oregon Senate Bill 418 (2021), and Indiana Senate Bill 415 (2023), for examples of such legislation.

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

emergency situations, namely situations in which an FEE is necessary to address an acute risk of harm. I reached this conclusion by arguing, first, that the rule of law and police legitimacy depends on their being an honest source of reliable knowledge and information about the law, criminal cases, and suspects' and victims' rights; and, second, that people should not be liable to harm based on an act they were defrauded into performing because people should only be liable to harms resulting from acts they performed voluntarily or consensually. A significant tradeoff is that there will be some cases in which good, probative evidence will be excluded. Consider two difficult examples.[28]

Gangster: Suppose a suspect is arrested by the police for a crime committed by members of the suspect's gang. The interrogators tell the suspect they have evidence that the suspect's friend is the ringleader and promise the suspect a lower charge and reduced prison sentence if he incriminates that friend. Now, the suspect knows that his friend is not the ringleader, but the suspect can also see that the police have decided that the friend is the ringleader and are not interested in any other explanation. So the suspect accepts the police's offer and gives them information incriminating his friend. However, the police renege on their end of the deal and do not seek a reduced charge and sentence.

Is the suspect responsible for the harm he suffers in this scenario (conviction and incarceration, along with his friend no longer speaking to him), and is the suspect responsible for the harm inflicted on his friend? There are essentially two frauds in this scenario: one committed by the police, and one committed by the suspect. The police intentionally deceive the suspect in a way that harms the suspect's interests (falsely promising to help the suspect if he talks); the suspect should not be responsible for any resulting harm to himself. On the other hand, the suspect intentionally deceives the police in a way that harms their interests (falsely identifying his friend as the gang ringleader); the suspect thus bears some responsibility for any harm against the police's interests (along with any harm to the suspect's friend). But the police continue to bear ultimate responsibility because the initial bargain with the suspect was a sham. The point is a pragmatic one: Identify each fraud where it occurs and assign (and mitigate) responsibility for any resulting harm accordingly.

The tricker question is this: What does moral responsibility for the confession and the harms that follow have to do with the question of whether the confession should be admitted as evidence or not? This is a difficult policy question because there are other reasons – beyond the voluntary or consensual nature of a confession – to question whether probative evidence should be admitted. One reason is *reliability*. Evidence derived from an FEE will often be unreliable because – as in the example above – the police made it the case that the rational thing for the suspect to do was to incriminate his friend, even though the friend was not the ringleader. We saw this in the *Thomas* case, in which Thomas was so manipulated that he believed the only way to save his wife and family was to confess to harming (murdering) his own child. The police's use of an FEE focused on extracting

[28]I thank Katerina Hadjimatheou for variations of these examples – along with accompanying analysis – in her excellent comments on an early draft of this paper, provided during the *Policing, Policy, and Philosophy Initiative* ("3PI") Symposium hosted by Penn State on March 1, 2024.

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

incriminating testimony, not truth. This raises a question: *Would an FEE be more justified if the police focused on truth – even in the absence of acute threats of harm?* Consider a second example.

Dead body: The police have sufficient evidence to charge a murder suspect with the murders of two people; the suspect confesses when faced with the overwhelming evidence. The police believe the suspect committed a third murder, but they do not have strong evidence to support their belief. The police thus tell the suspect that if he does not confess to the third murder, they will ensure that the suspect receives the death penalty. In fact, the death penalty no longer exists in the state, but the suspect does not know this. The suspect confesses and in doing so provides detailed descriptions of the murder of the third victim – and the location of the body – which is then corroborated through investigation.[29]

Should the police have refused to search the location for the body because the suspect provided information in an FEE? Or after the body is located, should the police refuse to disclose that it was the suspect who provided the location information? Of course not – and saying otherwise would rightly be met with an incredulous stare. It would be perverse to act *as if* there was no evidence in such situations. However, there still remains a question of whether the suspect should be *harmed* by the evidence, considering that the evidence was the product of an FEE. The answer is still *no* according to my argument. The answer *cannot* be otherwise if we assume that the suspect truly has *rights* (to remain silent, to only be responsible for consensual statements obtained without fraud, and so on).[30]

Ironically, legal doctrine supports this conclusion, even if the relevant doctrines have been somewhat eroded by exceptions. For example, in the United States, the "fruit of the poisonous tree" doctrine extends the exclusionary rule (the rule preventing the government to use evidence gathered in violation of the Constitution) to make evidence inadmissible in court if it was derived from evidence that was illegally obtained. As the metaphor implies, if the evidential "tree" is tainted, so is its "fruit."[31] The rule can even bar testimonial evidence resulting from excludable evidence such as a confession.

Of course, there are many exceptions under which evidence will not be excluded, including: (1) evidence discovered from a source independent of the unlawful activity; (2) evidence that would have been discovered inevitably; and (3) evidence discovered as a result of excludable but *voluntary* testimony from the defendant. Relatedly, illegally obtained evidence (and its "fruit") may be admissible if obtained under the "good faith exception."[32] These exceptions are consistent with the arguments in this paper insofar as a suspect would be harmed *even if* the police did not use an FEE. Moreover, when the police act in good faith (but make a mistake), they cannot be accused of relying on an FEE because an FEE requires intentionality (to engage in fraud) by definition.

[29]Ibid.

[30]See, for example, the 5th Amendment, U.S. Constitution, made applicable to the state through the 14th Amendment, U.S. Constitution.

[31]See, for example, Silverthorne Lumber Co., Inc. v. United States (1920), Nardone v. United States (1939), and Mapp v. Ohio (1961).

[32]See, for example, United States v. Leon (1984) and Arizona v. Evans (1995).

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

There is also legal doctrine that supports my conclusion about the justification of FEEs in emergency situations. Courts have recognized the "exigencies of the situation" as an exception to the requirement that police obtain search and arrest warrants based on probable cause – and an exception to reading a suspect their "Miranda Rights." In short, such actions are legally justified for reasons that include dangerous and life-threatening circumstances in which there is a compelling need for official action and not time to waste. In such cases, there is often an objective need for the police to act quickly in order to protect their lives or the lives of others.[33] Likewise, an FEE may be justified by emergencies involving acute threats of harm.[34]

The point is that there is general legal support for the policy implications of my argument. Admittedly, though, there is significant tension between my argument and legal doctrine in many cases. One example is the exception that "poisonous fruit" is admissible when it is a result of excludable but "voluntary" testimony from the defendant. My arguments call into question the extent to which confessions can be truly voluntary when they are based on an FEE.

* * *

My conclusions about police interrogation specifically build upon the constraining principles of the rule of law generally.[35] Where my prior work on this subject has emphasized what I take to be normative foundations of the police institution, this paper has focused on two related but distinct issues: (1) how police interrogation specifically can create the sort of epistemic environment (namely, a fraudulent one that precludes voluntariness) that renders deception and dishonesty unjustified (despite the police establishing probable cause), and (2) how such unjustified interrogation tactics might be addressed at the level of policy – in other words, policies that would *exclude* certain confession evidence as a *remedy* for unjustified interrogation (including through legislation prohibiting some of the interrogation tactics in the first place).

Although it is generally necessary for the police to establish probable cause before engaging in custodial interrogation, meeting that epistemic standard should not always justify the police's use of FEEs. Marvin Backes's makes a similar point in the context of reaching verdicts in legal cases:

> [I]n addition to reaching verdicts that are likely to be true, the law is also concerned with reaching verdicts that are fair and just—something that is

[33]See, for example, New York v. Quarles (1984), United States v. Santana (1976), and Birchfield v. North Dakota (2016).

[34]Suppose the police are investigating a series of violent bank robberies in which the robber holds up banks at gunpoint. The police believe they have identified the bank robber, but they do not know his location. However, the police locate and detain the suspect's brother because they have probable cause to believe the brother is concealing the robber's location. This is an emergency scenario in which an FEE would be justified against the brother, considering that there is an acute threat that the robber will engage in another violent robbery in which people could be harmed.

[35]See Hunt 2019, chapter 5 (setting forth a "prerogative power test" (PPT) regarding the justification of rule of law departures by the police), Hunt 2023, and Hunt 2024, chapter 2 (arguing that good faith is a normative foundation of the police institution and applying the PPT to police deception and dishonesty generally).

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

> ultimately irrelevant to epistemic theorizing . . . we should not expect epistemic theories to track legal verdicts; and similarly, we should abandon the idea that legal considerations are helpful in evaluating epistemic theories (Backes 2020, 2777).

Even if we are sympathetic to a criminal justice system based on epistemic justification tied to high probability, there are strong reasons – in the form of both deontological and consequentialist considerations – to think the police should be constrained in their use of FEEs to obtain confessions through interrogation. Justifying police tactics in terms of heightened probability is not always compatible with important norms of political morality.

What does this mean at the level of policy? There are cases in which evidence with probative value should be excluded when it is derived from an FEE that opens one to harm and a loss of rights, namely: those cases in which the FEE is not needed to address emergencies involving acute threats of harm. The upshot is that the police will be less likely to obtain probative evidence in many cases. But if it is right to say that people should not be held responsible for the results of nonconsensual relations with the police, then excluding some probative evidence is justified when the evidence is derived from a FEE.

Data availability statement. This study does not employ statistical methods and no replication materials are available.

Acknowledgments. For their thoughtful comments, I am grateful to the participants at the *Orange Beach Epistemology Workshop*, Orange Beach, Alabama, in May 2023; Central Division Meeting of the *American Philosophical Association* in New Orleans, in February 2024; *Policing, Policy, and Philosophy Initiative* ("3PI") Symposium hosted by Penn State on March 1, 2024; and *Law and Society Association* Annual Meeting, in Denver, in June 2024. I also thank the anonymous reviewers for their helpful suggestions.

Competing interests. No competing interests.

References

Alexander v. DeAngelo. 2003. 329 F.3D 912 (7th Cir. 2003).

American Law Institute. 1962. Model Penal Code (Proposed Official Draft, 1962).

Arizona v. Evans. 1995. 514 U.S. 1 (1995).

Backes, Marvin. 2020. "Epistemology and The Law: Why There is No Epistemic Mileage in Legal Cases." *Philosophical Studies* **177**: 2777.

Bergmann, Michael Bergmann. 2006. *Justification without Awareness: A Defense of Epistemic Externalism*. Oxford: Oxford University Press.

Betz, Joseph. 1985. "Sissela Bok on the Analogy of Deception and Violence." *The Journal of Value Inquiry* **19**: 217–224.

Birchfield v. North Dakota. 2016. 579 U.S. 438, 456 (2016).

Bradford, Ben, Jackson, Jonathan, and Hough, Mike Hough. 2017. "Trust in Justice." In *The Oxford Handbook of Social and Political Trust*, ed. Eric M. Uslaner, 633–54. New York: Oxford University Press.

Commonwealth v. Selby. 1995. 420 Mass. 656 (Mass. 1995), 651 N.E.2d 843.

Davis, Deborah, and Leo, Richard A. 2012. "Interrogation through Pragmatic Implication: Sticking to the Letter of the Law While Violating its Intent." In *The Oxford Handbook of Language and Law*, eds. Peter M. Tiersma and Lawrence M. Solan. Oxford: Oxford University Press.

Dwyer, Jim. 2019. "The True Story of How a City in Fear Brutalized the Central Park Five." *The New York Times*, March 30.

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press

Etienne, Margareth, and McAdams. 2021. "Police deception in interrogation as a problem of procedural legitimacy." Texas Tech Law Review 54: 21.

Falk, Patricia J. 1998. "Rape by Fraud and Rape by Coercion." *Brooklyn Law Review* **64**: 39–180.

Feinberg, Joel. 1987. *Harm to Others: The Moral Limits of Criminal Law*. New York: Oxford University Press.

Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.

Frowe, Helen. 2018. "Lesser-Evil Justifications for Harming: Why We're Required to Turn the Trolley." *Philosophical Quarterly* **68** (272): 460–480.

Gardiner, Georgi. 2019. "The Reasonable and the Relevant: Legal Standards of Proof." *Philosophy & Public Affairs* **47** (3): 288–318.

Gordon-Solmon, Kerah. 2022. "How (and How Not) to Defend Lesser-Evil Options." *Journal of Moral Philosophy* 20 (3-4): 211–232.

Green, Stuart P. 2006. *Lying, Cheating, and Stealing – A Moral Theory of White-Collar Crime*. Oxford: Oxford University Press.

Harris, Aisha Harris. 2019. "The Central Park Five: 'We Were Just Baby Boys'". *The New York Times*, May 30.

Hart, H.L.A. 2012. *The Concept of Law*. Oxford: Oxford University Press. 193–200.

Hunt, Luke William. 2016. "Legal Speech and Implicit Content in the Law." *Ratio Juris* **29** (1): 3–22.

Hunt, Luke William. 2019. *The Retrieval of Liberalism in Policing*. New York: Oxford University Press.

Hunt, Luke William. 2022. "The Limits of Reallocative and Algorithmic Policing." *Criminal Justice Ethics* **41** (1): 1–24.

Hunt, Luke William. 2023. "Good Faith as a Normative Foundation of Policing." *Criminal Law and Philosophy* **17** (3): 1–17.

Hunt, Luke William. 2024. *Police Deception and Dishonesty – The Logic of Lying*. New York: Oxford University Press.

Illinois Senate Bill 2122 (2021).

Illinois v. Gates. 1983. 462 U.S. 213 (1983).

Indiana Senate Bill 415 (2023).

Innocence Project. "Explore the Numbers: Innocence Project's Impact." available at https://innocenceproject.org/exonerations-data/

Kant, Immanuel. 1996. "On a Supposed Right to Lie from Philanthropy," In *Practical Philosophy*, trans. Mary Gregor, 605–16. Cambridge: Cambridge University Press.

Kassin, Saul. 2002. "False Confessions and the Jogger Case," *The New York Times*, November 1.

Kassin, Saul. 2022. *Duped: Why Innocent People Confess – and Why We Believe Their Confessions*. Lanham: Prometheus Books.

Khasin, Irina. 2009. "Honesty is the best policy: A case for the limitation of deceptive police interrogation practices in the United States." *Vanderbilt Journal of Transnational Law* 42: 1029.

Lackey, Jennifer. 2023. *Criminal Testimonial Injustice*. New York: Oxford University Press.

Magid, Laurie. 2001. "Deceptive police interrogation practices: How far is too far?" *Michigan Law Review* **99** (5): 1168–1210.

Mapp v. Ohio. 1961. 367 U.S. 643 (1961).

Miranda v. Arizona. 1966. 384 U.S. 436 (1966).

Monaghan, Jake. 2021. 'Boundary Policing', Philosophy & Public Affairs 49.1 (2021).

Moss, Sarah. 2021. "Pragmatic Encroachment and Legal Proof." *Philosophical Issues* **30** (1): 258–79.

Moss, Sarah. 2023. "Knowledge and Legal Proof." *Oxford Studies in Epistemology* **7**: 176–213.

Nardone v. United States. 1939. 308 U.S. 338 (1939).

Neder v. United States. 1999. 527 U.S. 1, 25 (1999).

New York v. Quarles. 1984. 467 U.S. 649 (1984).

Oregon Senate Bill 418 (2021).

People v. Palmer. 2001. 282 A.D.2d 256 (N.Y. App. Div. 2001).

People v. Thomas. 2014. N.Y. Slip Op. 1208 (N.Y. 2014).

Philipps, Dave. 2015. "Former Oklahoma City Police Officer Found Guilty of Rapes." *New York Times*, December 10.

Quiroz, Nigel Quiroz. 2021. "Five Facts About Police Deception and Youth You Should Know." *Innocence Project*, May 13.

https://doi.org/10.1017/S014381 4X25000030 Published online by Cambridge University Press

Rakoff, Jed S. 2014. "Why Innocent People Plead Guilty," *The New York Review of Books*, June 20.
Restatement (Second) of Contracts, § 168. 1981. § 168 cmts. B& c, § 169 (1981).
Shiffrin, Seana Valentine. 2014. *Speech Matters*. Princeton: Princeton University Press.
Silverthorne Lumber Co., Inc. v. United States. 1920. 251 U.S. 385 (1920).
Skolnick, Jerome H., and Leo, Richard. 1992. A. "The ethics of deceptive interrogation." *Criminal Justice Ethics* **11** (1): 3–12.
State v. Walker. 1992. 493 N.W.2d 329 (1992).
Sturgeon, Scott. 2008. "Reason and the grain of belief." *Noûs* **42** (1): 139–165.
Terry v. Ohio. 1968.392 U.S. 1 (1968).
United States v. Leon. 1984. 468 U.S. 897 (1984).
United States v. Morris. 1996. 80 F.3d 1151, 1161 (7th Cir. 1996).
United States v. Santana. 1976. 427 U.S. 38, 42 (1976).
Vallier, Kevin, and Weber, Michael. (ed.). 2021. *Social Trust*. New York: Routledge.
Williams, Bernard. 2002. *Truth and Truthfulness*. Princeton: Princeton University Press, 114–22.

Cite this article: Hunt LW (2025). Police interrogation and fraudulent epistemic environments. *Journal of Public Policy*. https://doi.org/10.1017/S0143814X25000030

https://doi.org/10.1017/S0143814X25000030 Published online by Cambridge University Press