You Can’t Handle the Truth! Trial Juries and Credibility

Renée McDonald Hutchins*

I. INTRODUCTION

Every now and again, we get a look, usually no more than a glimpse, at how the justice system really works. What we see—before the sanitizing curtain is drawn abruptly down—is a process full of human fallibility and error, sometimes noble, more often unfair, rarely evil but frequently unequal.1

* Associate Professor of Law, University of Maryland Carey School of Law. For comments on earlier drafts, I thank my colleagues in the Criminal Law Research Collective and the Clinical Theory Workshop. I also thank for their time and incredible talents my research assistants: Saidah Grimes, Claire McLamore, Stephanie Noronha, Derek Simmonsen, Akeel St. Jean, Christina Taylor, and Amond Uwadineke.


I. INTRODUCTION

Every now and again, we get a look, usually no more than a glimpse, at how the justice system really works. What we see—before the sanitizing curtain is drawn abruptly down—is a process full of human fallibility and error, sometimes noble, more often unfair, rarely evil but frequently unequal.1

* Associate Professor of Law, University of Maryland Carey School of Law. For comments on earlier drafts, I thank my colleagues in the Criminal Law Research Collective and the Clinical Theory Workshop. I also thank for their time and incredible talents my research assistants: Saidah Grimes, Claire McLamore, Stephanie Noronha, Derek Simmonsen, Akeel St. Jean, Christina Taylor, and Amond Uwadineke.


I. INTRODUCTION

Every now and again, we get a look, usually no more than a glimpse, at how the justice system really works. What we see—before the sanitizing curtain is drawn abruptly down—is a process full of human fallibility and error, sometimes noble, more often unfair, rarely evil but frequently unequal.1

* Associate Professor of Law, University of Maryland Carey School of Law. For comments on earlier drafts, I thank my colleagues in the Criminal Law Research Collective and the Clinical Theory Workshop. I also thank for their time and incredible talents my research assistants: Saidah Grimes, Claire McLamore, Stephanie Noronha, Derek Simmonsen, Akeel St. Jean, Christina Taylor, and Amond Uwadineke.

The central question, vital to our adjudicative model, is: How well can we expect a jury to determine credibility through the ordinary adversary processes of live testimony and vigorous impeachment? The answer, from all I have been able to see is: not very well.²

It was the late 1990s, and I was working as a federal prosecutor at “Main Justice.” My special assignment to the local U.S. Attorney’s Office gave me an opportunity to sit second chair in a murder prosecution.³ Lead counsel was a highly experienced prosecutor. She had handled hundreds of major felony cases and scores of murder trials. The accused in our case was suspected of tracking a long-time rival one night and shooting him twice in the head. By the time of trial, the accused had been shot in retaliation. He was confined to a wheelchair. At trial, he chose to testify in his own defense. There was a tragic air to him—at just nineteen years of age, the retaliatory shooting had rendered him a paraplegic. Unable to mount the two short steps to the witness box, the court allowed him to testify from his place at counsel’s table. His direct examination went smoothly. Attractive and well-spoken, the defendant testified to his lack of involvement in the crime. “I was home in bed,” he assured the jury.

When it came time, my colleague worried that a rigorous cross-examination would appear heartless and offend the jurors. Going too easy on him, though, risked the possibility that the defendant’s claimed lack of involvement would survive untested. She decided to press him. The early portion of the cross-examination was unremarkable. Walking him through the details of the evening, the lead prosecutor was able to catch the defendant in only the most minor of discrepancies—what he had for dinner, what he was wearing, the name of the show he watched before going to bed. A turning point, however, came when she began questioning the defendant about his relationship with the victim.

After demonstrating the long and violent history between the two men, counsel moved to the night in question. “Isn’t it true,” she asked, “that [the victim] encroached on your territory and you wanted him dead?” “No,” he shot back. And then, as if on cue, his leg twitched, banging loudly against the edge of counsel’s table.

³ Names have been withheld and some details of the case have been modified to protect the privacy of those involved.
“Isn’t it true that you held a gun to the back of [the victim’s] head and pulled the trigger twice?” she continued. “No!” and again his leg slammed up into the table. “Isn’t it true that after the shooting you ran to your cousin’s house and told him to hide the gun in a shoebox?” The calm, composed young man from direct examination was gone. “No,” he sneered. But, once again, his leg jerked against the table. He moved his hand to his knee to keep it from knocking. I glanced at the jury. Had they seen what I’d seen? At the close of the evidence, it took less than an hour for the jury to return with a guilty verdict. Had the defendant’s involuntary muscle spasms during tough questioning been a sign of false testimony? That’s certainly the way I interpreted them. Conversations with some jurors after the verdict suggested that they thought so too. But were we right?

Everyone lies. Some people more than others. There is little question that many witnesses who take the stand and raise their hand to swear the oath thereafter present the jury with a false accounting of the facts. There is also no question that many others do not. Of those witnesses who do lie, they do so for any number of reasons—some compelling, some wholly self-serving. Whatever concessions we make with regard to the failings of witnesses, we allow very few similar concessions for the failings of jurors. We trust that, for the most part, jurors will be able to spot those who lie. We encourage jurors to watch witnesses carefully. We then trust that jurors have done their job well when, after watching a victim resolutely declare, “I’m 100% certain that’s the woman who did it,” they swiftly impose a verdict of guilt. We quiet ourselves with the alluring lullaby that blind faith in jurors is not a matter of expedience, but rather a demonstration of fairness. Jurors, we tell the naysayers, simply do it better.

From messages delivered in best-selling fiction to instructions delivered by trial judges, our popular discourse is saturated with the notion that all firsthand observers have a particular advantage when making

---

4 Uviller, supra note 2, at 814 (“They will try to tell the truth, the whole truth, and nothing but the truth when their interest in promoting a falsehood is no stronger than their scruple. But raise the stakes (say a beloved child needs an alibi) and increase the certainty of impunity (it’s her word against mine), and even people who pride themselves in their honesty will either take their chances with perdition or convince themselves that their lies are true.”).

5 Much of this Article is framed as a critique on juror ability to assess credibility. Similar concerns attach in bench trials, however, where trial judges sit in the role of fact-finder.

6 See, e.g., Uviller, supra note 2, at 824–25 (noting that in one study most judges polled believed juries reached the factually correct verdict in an overwhelming majority of cases).
reliability assessments. The system then insulates those assessments in a discourse of reliability that makes post-trial attacks difficult.\(^7\)

In a troubling rebuff, however, to the popular narrative, studies show that the cues jurors look to when assessing credibility are actually the wrong ones. This Article seeks to lay bare the fiction that most firsthand observers are well-suited to make credibility assessments. There are many legitimate reasons why jurors, as first level fact-finders, are entitled to principal responsibility for credibility assessments. And there are abundant rationales for affording deference to their findings. Reliability, however, is not one of them.\(^8\) Studies increasingly show that without additional guidance, jurors are fairly poor evaluators of witness deception. The time has come to take those findings seriously. If a primary function of criminal trials is to make accurate determinations of guilt,\(^9\) the integrity of the system requires honesty about what motivates our deference to trial level credibility assessments. Only then will it be possible to tweak the system to help ensure fairer outcomes.

In Part II, this Article examines the historic development of the jury as lie detector both in the courts and in our popular conscience. Part II concludes that the current dependence on jurors as reliable

---

\(^7\) This Article focuses entirely upon modifications that might be made at the trial level to improve the reliability of outcomes. I reserve for a later date a discussion of potential modifications to the appellate system—like universal adoption of weight of the evidence review—that might better correct faulty trial assessments after they have occurred.

\(^8\) A variety of theories have been posited to explain why we are not particularly good at lie detection. These explanations include theories of evolution ("[O]ur ancestral environment did not prepare us to be astute lie catchers"), parental deception ("[Parental] privacy may often require that they mislead their children about just what they are doing."), the drudgery of a constantly suspicious life ("It is only the paranoid who foregoes such peace of mind, and those whose lives are actually at some risk if they are not constantly alert to betrayal."), and the willing collusion of the target ("Most of us operate on the unwritten principle of postponing having to confront anything that is very unpleasant, and we may do so by collusively overlooking a liar’s mistakes."). For a fuller discussion, see Paul Ekman, Telling Lies: Clues to Deceit in the Marketplace, Politics and Marriage (Norton & Company, New York 2009) at 341–43.

\(^9\) In the late 1960s, Professor Herbert Packer articulated two theoretical models for the criminal justice system. Under the first—the Crime Control Model—speed and certitude were most important. Under the second—the Due Process Model—fairness was the coin of the realm. Herbert Packe, The Limits of the Criminal Sanction 154–73 (1968). Under both models, however, there can be little question that accuracy should be the dominant goal. See David A. Sonenshein & Robin Nilon, Eyewitness Errors and Wrongful Convictions: Let’s Give Science a Chance, 89 OR. L. REV. 263, 302–03 (2010) ("Implicit in both models is the reliable judicial separation of the guilty from the truly innocent.").
detectors of truth is a relatively new construct. In Part III, this Article discusses the growing body of empirical evidence finding that juries are not particularly well-suited to make credibility assessments. Part III also explores the substantial problems that inaccurate witness credibility assessments create. Finally, in Part IV, this Article suggests that, to avoid erroneous jury determinations of witness credibility, modification of the system should be considered.

To be clear, this Article’s primary mission is to identify the problem and suggest some possible ways we might begin to address it. This piece does not propose radical overhaul of the existing criminal trial system. Nor does it recommend abandoning the role of the jury as first-level fact-finder. Indeed, I do not suggest that a fully satisfactory resolution of the problem is even possible given the developmental life of lie detection science. Rather, this Article is intended to begin the conversation—to identify the problem and some of the reasons for concern so that, as the science develops, we can more fully deploy an appropriate response. Put more simply, the point of this piece is to suggest that we should make whatever improvements are possible—even if those improvements are made simply by acknowledging the known unknowns. And there is reason to believe we can do even more. We can tell jurors what we do know about the behaviors that have predictive utility for lie detection. We can also give jurors more meaningful opportunities to observe witnesses, albeit within the artificial arrangement that is the witness box. Finally, we can directly confront juror skepticism to avoid distortion of credibility assessments.

II. ASSESSING CREDIBILITY: THEN AND NOW

Commentators have argued that accurate assessments of witness credibility depend upon the fact-finder’s ability to observe firsthand the witness as he or she is speaking. Justice Brennan once noted “there is almost nothing more convincing than a live human being who takes the stand, points a finger at the defendant, and says ‘That’s the one!’” But the notion that jurors are (and should be) the ultimate judges of truth telling is a fairly modern one, born out of the dual needs for finality and legitimacy. It was not always the case that jurors were viewed as peculiarly well-situated to make credibility calls. Assessment of criminal culpability was once made by ordeals that

---

claimed results of divine judgment. It was only over the course of several centuries that a system that anchored its legitimacy first in God and then in the oath gradually evolved into one that is anchored in the jury’s ability to recognize candor.

A. A Slow Conversion from Trial by Fire to Trial by Jury

Prior to reliance on jury trials, trial by ordeal was the favored method of sorting guilt and innocence.12 There were two primary ordeals used to assess guilt.13 The first was the ordeal by cold water.14 A second ordeal—trial by fire—involves the infliction of burns on the hands of the accused with a hot iron.15 The popularity of these trials by ordeal rested on the notion that one’s culpability was being judged by God.16

In 1215, however, trials by ordeal came to an abrupt end when the Catholic Church prohibited priests from presiding over future ordeals.17 The exclusion of priests stripped the ordeals of their ostensible divine authority and, consequently, eroded their popular appeal.18 Concomitant with the demise of the ordeal, European countries began adopting “rational” Roman-canon methods of proof that placed great faith in sworn eyewitness testimony and confessions

---

14 With trial by cold water, the accused was stripped and lowered into a 12-foot deep pool of water with his hands bound. If the accused sank, he was deemed innocent—the theory being that the water, which was pure, had accepted him. If the accused floated, he had been “rejected” by the water, and was adjudged guilty. Id. at 582–83.
15 With trial by fire, burns were inflicted on the palm with a hot iron. The wound was then bandaged. Several days later, the bandages were removed. If the burns were healing well, the accused was declared innocent. If the injuries were infected however, it was a sign from God and the accused was deemed guilty. Id. at 588.
16 Rebecca V. Colman, Reason and Unreason in Early Medieval Law, 4 J. INTERDISC. HIST. 571, 582 (1974). As scholars have rightly noted, however, despite the common construction of the ordeal as producing a divine mandate, the outcome was a product of human judgment. A human intermediary was the judge of whether a burn healed cleanly; of whether a body sunk or floated. “The institutional brilliance of the ordeal was that it so neatly merged the appearance of divine judgment with the reality of a great measure of human control.” Fisher, supra note 12, at 601.
17 W.G. Aitchison Robertson, Trial by Ordeal, 38 JURID. REV. 70, 78 (1926) (noting that the Lateran Council forbade the clergy from participating in trials by ordeal).
By 1220, these two developments produced the first documented criminal jury trial in England. But general views about jury competency at that time were only a distant cousin to our current conceptions. Public acceptance of such trials was far from immediate, for the trials lacked the ostensible divine sanction of the ordeal. To address this concern, tribunals sought the consent of the accused. Consent, however, was a poor substitute for divine legitimacy where such consent was frequently obtained by either harsh imprisonment (the *prisone forte et dure*), or tortuous punishment (the *peine forte et dure*). A plausible claim that the new judgments were anchored in some superior form of decision making was needed if jury trials were to garner the broad social acceptance previously afforded trials by ordeal.

Reliance upon the law initially provided the superior form of decision making. As Charles Donahue said, it was critical that the “ultimate decision of the case not be within human discretion but be dictated by the rules of law. People might accept a judgment of the law rather than a judgment of God; it was less likely that they would accept a judgment of man rather than one of God.”

Reliance upon the law initially provided the superior form of decision making. As Charles Donahue said, it was critical that the “ultimate decision of the case not be within human discretion but be dictated by the rules of law. People might accept a judgment of the law rather than a judgment of God; it was less likely that they would accept a judgment of man rather than one of God.”

---


20 In his 1997 piece in the Yale Law Journal, George Fisher postulated that the popularity of the ordeal was the “apparent intervention of God.” Fisher, *supra* note 12, at 587 (“The old system of trial by ordeal bespoke a social humility, an unwillingness to take life or limb without divine sanction.”). The shift from divine judgment to human judgment, scholars have hypothesized, is what drove early skepticism of jury trials. As Fisher readily acknowledges, however, there is scholarly disagreement whether widespread opposition to human judgments really existed. See, e.g., Fraher, *supra* note 19, at 56–62 (arguing that in practice, there was a great deal of discretion in the system, the existence of which provides proof that there was confidence in human judgments). As Fisher notes, however, where the system publicly proclaimed strict application of legal rules, evidence of discretion in practice may be proof only of institutional (not public) acceptance of human judgment. Fisher, *supra* note 12, at 590 n.39.

21 Fisher, *supra* note 12, at 588 (explaining that defendants were coerced into consenting by *prisone et dure*, which was later replaced by *peine forte et dure*)

guarantor of the system in the form of the oath. While there is some dispute about the precise timing of this second evolution, by the middle of the fifteenth century the evidence of sworn witnesses had become a foundational part of the system.

Significantly, accepting a witness’s sworn testimony for jury assessment did not yet signal confidence in the jury’s abilities. It was the oath, not the jury, that was seen as the central guarantor of truth telling. Significant steps were taken, therefore, to maintain the perceived legitimacy of the oath. For example, to avoid embarrassing challenges to the notion that the oath functioned as an iron-clad guarantee of truthfulness, rules that permitted only the prosecution to call witnesses helped avoid credibility contests. Rules on witness competence and racial exclusion laws also limited those who were qualified to give sworn testimony. By limiting testimony, the rules necessarily limited the opportunities for credibility battles. Other rules, like the Rule in Bethel’s Case and the evidentiary premise on

---

23 Fisher, supra note 12, at 591, 594. Whether law completely supplanted God or rather was only a means by which divine sanction was retained, there is certainly agreement that the current formula for jury trials was not a deliberate institutional strategy. (“[G]radually and over time, the institutions of justice gravitated toward a formula for the jury trial that permitted the system to reassert its old claim to divine legitimacy.”).

24 It is generally acknowledged that for the first hundred years or so, the divine legitimacy of jury trials was far from a settled question. Compare Fisher, supra note 12, at 595, with Langbein, supra note 19, at 78. It also appears that there is no clear historical explanation for the decision to settle upon the oath as the source of divine legitimacy. See generally, Fisher, supra note 12, at 591 n.43.


26 Fisher, supra note 12, at 583 (arguing that allowing only the prosecution to call witnesses made the oath more credible).

27 Under these rules, criminal defendants, spouses, convicted felons, anyone with a monetary stake in a case, atheists, and disfavored racial minorities were all barred from testifying. John H. Wigmore, A Treatise on the Anglo-American System of Evidence in Trials at Common Law 679-80 (3d ed. 1940); see generally James Oldham, Truth-Telling in the Eighteenth-Century English Courtroom, 12 LAW & HIST. REV. 95, 107–17 (1994). By limiting those who were qualified to give sworn testimony, the rules necessarily limited the opportunities for credibility battles.

28 Bethel, a candidate for Parliament, had been charged with assault and battery for hitting one of the King’s servants outside of a polling place. At trial, the prosecution presented the sworn testimony of six witnesses. Bethel, who was entitled to call sworn witnesses on his own behalf because the charge was a misdemeanor called five witnesses. Each of the prosecution’s witnesses testified they saw the blow. Each of the defense witnesses testified they had not seen a blow. In closing argument, the prosecutor urged the jury to conclude that the defense witnesses were “mistaken (or inattentive) before concluding that [they] had lied.” Fisher, supra
“positive” evidence\textsuperscript{29} also limited the jury’s ability to resolve pure credibility contests. But these rules (and others like them) eventually eroded.

The erosion, however, was gradual. Only little by little did rules develop that undermined the limitations on witness qualification. Allowing an accused’s sworn testimony was one such erosion. The increasing acceptance of alibi evidence was another.\textsuperscript{30} Diminution in the perceived value of the oath also occurred as non-Christian witnesses were increasingly tolerated.\textsuperscript{31}

Despite increased opportunities for jurors to be presented with conflicting oaths, however, the system was still unwilling to submit completely to the jury the task of credibility assessment. Competency rules and racial exclusion laws survived—shielding jurors from the full brunt of the task of divining truth. Other evidentiary guidelines also persisted. For example, well into the nineteenth century, juries faced with irreconcilable conflicts between sworn witnesses were
guided by the number of witnesses on each side. Jurors' ability to engage credibility contests was also constrained by the rule of falsus in uno, falsus in omnibus, which was endorsed by the Supreme Court as late as 1822. The rule, mandatory at its inception, instructed jurors to reject the entire testimony of a witness who was proven to have lied about any portion therein.

But the trend toward fully allowing juries to resolve credibility contests was growing. The collapse of competency rules began in earnest in the mid-1800s. And by 1881, most states and the federal government had abolished their ban on civil party testimony. The

---


33 3 Wigmore, supra note 27, at 1010; see also The Santissima Trinidad, 20 U.S. (7 Wheat.) 283, 339 (1822) (endorsing articulation of a mandatory construction of the rule); Pass v. State, 182 S.E.2d 779, 788 (Ga. 1971).

34 See, e.g., An Act for the Regulation of Civil Actions (codified at Conn. Rev. Stat. tit. I, ch.X, § 141 (1849)) (reflecting Connecticut's statutory acceptance of party testimony in civil cases). In 1846, two years prior to the passage of the Connecticut law, interested witnesses were allowed to testify in Michigan. Joel Bodansky, The Abolition of the Party Witness Disqualification: A Historical Survey, 70 Ky. L. J. 91, 93 (1981). Three years after the ban was lifted in Connecticut, civil parties in England gained the right to present evidence. See An Act to Amend the Law of Evidence, 1851, 14 & 15 Vict., c. 99 § 2 (Eng.). Within a decade, back on this side of the Atlantic, several states quickly passed laws permitting parties to testify. The first fourteen states to permit parties to testify in civil cases were: Connecticut (1848); Vermont (1852); Ohio (1853); Maine (1856); Rhode Island (1857); Mississippi (1857); New York (1857); Massachusetts (1857); New Hampshire (1857); Kansas (1858); New Jersey (1859); Michigan (1861); and Indiana (1861). See Fisher, supra note 12, at 669. As a quick review of the list reflects, Mississippi was the only southern state among this group to abolish civil competency rules early on.

35 The federal government abolished its ban on civil party testimony in 1864 as a result of efforts that also outlawed the exclusion of black witnesses. See Cong. Globe, 38th Cong., 1st Sess. 3260–61 (1864). By 1881, the additional states to permit parties to testify in civil cases were: Oregon (1862); California (1863); Maryland (1864); Florida (1866); Virginia (1866); North Carolina (1866); South Carolina (1866); Georgia (1866); Missouri (1866); Alabama (1867); Illinois (1867); Louisiana (1867); West Virginia (1868); Tennessee (1868); Pennsylvania (1869); Texas (1871); Kentucky (1872); Arkansas (1874); and Delaware (1881). See Fisher, supra note 12, at 669. As a quick review of this list suggests, many more southern states were in this latter group. By way of explanation for the lag in the South, scholars have noted that among states with racial exclusion laws, the overwhelming majority abolished their racial exclusion laws at the same time or shortly before abolishing their civil competency laws. Indeed just three states abolished their racial exclusion laws before
demise of other limitations on witness competence soon followed.

Though somewhat beyond the scope of this Article, it does bear brief mention that the emerging American willingness to accept sworn testimony from parties (civil and criminal alike) was not driven exclusively by growing acceptance of the jury’s aptitude for divining truth. Rather, the expansion of sworn testimony was, at least in part, a response to the brewing national debate regarding the legal status of American blacks.36 In the lead up to and immediately following the Civil War, racial exclusion laws in particular—and competency laws by association—became contentious issues.37 Roughly abolishing their civil competency laws. Noting this trend, at least one scholar has opined that the demise of civil competency rules in the South was more closely related to the victory of the Union in 1865 than to the region’s wholesale adoption of the transatlantic trend toward liberalized civil competency rules. As Fisher hypothesized, the South had little reason to resist the fall of competency rules after its defeat in the Civil War (and the commiserate demise of racial exclusion laws). Fisher, supra note 12, 671–74. Indeed, as Fisher suggests, with the fall of racial exclusion laws, there may have been a sense of increased urgency in permitting parties in civil cases to testify. This is because the absence of racial exclusion created the potential for results that could not be reconciled with the South’s racial caste system. For example, in the absence of racial exclusion, a black witness would be allowed to testify against a white party who was forced to sit silent under a civil competency bar.

36 As scholars have noted, the early demise of competency rules in the North was in no small part motivated by northern political attacks on the South’s exclusion of non-whites as witnesses. In an effort to anchor their claims in something more than bald claims of racism, northern opponents in the U.S. Congress cast their opposition to the southern rules in non-racial terms, and suggested that all competency rules were a hindrance to the full development of the factual record. During the Civil War and early Reconstruction periods, the most divisive human rights issues were slavery, racial exclusion laws, and the right of black Americans to vote. Fisher, supra note 12, at 675. Though little could be done at a federal level on the first and third items until the end of the war, the senatorial debates of the period did result in legislation that liberalized civil competency rules and rejected racial exclusion laws in federal courts. It was not until the additional leverage afforded by a northern victory, however, that the balance of the agenda was addressed in earnest. At the close of the Civil War, the abolition of slavery was generally accepted as a necessary pre-condition to a seceded state’s readmission into the Union. Robert J. Kaczorowski, The Supreme Court and Congress’s Power to Enforce Constitutional Rights: An Overlooked Moral Anomaly, 73 Fordham L. Rev. 153, 205 n.241 (2004); see also Eric Biber, The Price of Admission: Causes, Effects, and Patterns of Conditions Imposed on States Entering the Union, 46 Am. J. Legal Hist. 119, 143–44, 150 (2004). The abolition of racial exclusion laws was also seen as a desirable pre-requisite. Stephen A. Siegel, The Federal Government’s Power to Enact Color-Conscious Laws: An Originalist Inquiry, 92 Nw. U. L. Rev. 477, 548 (1998). In contrast, the remaining issue of the day—the right of black Americans to vote—did not garner similar support. The Progress of Reconstruction, N.Y. Times (Oct. 3, 1865), available at http://www.nytimes.com/1865/10/03/news/the-progress-of-reconstruction.html.

37 For a detailed accounting of the North-South battle over racial exclusion laws,
concurrent to (and in some instances simultaneous with) expansions in competency rules came the fall of racial exclusion laws in the States. Each new development pushed the jury further and further toward its now central role of assessing witness credibility.

In the late 1800’s, American courts began, for the first time, to clearly express the view that the jury was the appropriate institution to assess credibility. By the early 1900s, the Georgia Supreme Court found “[c]redibility is also essentially involved. . . . [t]he jury [is to] consider not only what a witness swears, but also what credit is to be given him as a witness.” Other courts soon followed.

It was in this manner that historic developments, over the course of centuries, eventually removed the substantial barriers that once

---

38 The simultaneous fall of racial exclusion and competency laws in many southern States appears to have been motivated by a view that it would be unseemly to have blacks testify against whites who were barred from speaking as a result of competency rules. See, e.g., Our Sumter Correspondence, CHARLESTON DAILY COURIER, 4 (Sept. 5, 1866); see also An Act to Protect Freedmen in Their Rights of Person and Property in This State, No. 86, 1866 Ala. Acts 98 (adopted Dec. 9, 1865) (recognizing the right of parties to testify when cross-racial testimony was expected).

39 Scholar George Fisher has identified Humphries v. State, 100 Ga. 260 (1896) as the first case to clearly and completely assign the task of “lie detection” to the jury. Fisher, supra note 12, at 638 n.276.

40 Warrick v. Georgia, 125 Ga. 133, 141-42 (1906).

41 State v. Taylor, 261 Mo. 210, 229 (1914) (“Here the weight of the evidence and the credibility of the witnesses were wholly for the triers of fact and not for us as soon as the verdict of the jury met the approval of the trial court.”).
shielded the jury from squarely evaluating credibility. The continued discomfort with jurors’ abilities to meet the task before them, however, lingered well into the modern day. In the late nineteenth century, contemporary jury instructions continued to guide juries in the evaluation of ostensibly questionable sworn testimony. More recently, legal norms limited a jury’s ability to convict when presented with testimony deemed generally unreliable. For example, though no longer the law in any state, for many years, uncorroborated accomplice testimony was insufficient to sustain a conviction. Presumably in a similar vein, a rape victim’s testimony required corroboration. Over time, however, even the remaining rules met a death of attrition as only a minority of jurisdictions, if any, still embraced them. As with the demise of earlier rules, this erosion marked the near complete rise of the jury as a lie detector. Commenting on the transition, scholar George Fisher noted:

The jury . . . promised a remarkably reliable source of systemic legitimacy. Its usually private and inarticulate decisionmaking protected it from the sort of embarrassing public failures that so regularly threatened the oath. Although two oaths all too easily could conflict, the jury’s verdict stood alone and, at least within the system’s formal bounds, was almost immune from contradiction. Moreover, whether by tradition or conscious design, the jury’s verdict has been largely impenetrable. . . . The jury’s secrecy is an aid to legitimacy, for the privacy of the jury box shrouds the shortcomings of its methods.

So it was that the jury replaced the divine sanction of the ordeal as a guarantor of the system’s legitimacy. Once enthroned, the authority of the system required that we embrace the jury as the

---

42 See, e.g., People v. Morrow, 60 Cal. 142, 147 (1882) (approving an instruction regarding the defendant’s testimony that called the jury’s attention to the potential consequences to the defendant of an adverse trial result); State v. Maguire, 69 Mo. 197, 201–02 (1878) (approving an instruction informing the jury it could consider “the fact that [the defendant] is a witness testifying in his own behalf”); St. Louis v. State, 8 Neb. 405, 418–19 (1879) (same).


45 Fisher, supra note 12, at 705.
rightful heir. Commenting on early twentieth century juries, Edson Sunderland observed that they are a “procedural opiate” that “soothes us in the assurance that we have attained the unattainable. Because it emits no light, the black box of the jury room has become the system’s black hole, drawing into itself all of the questions of fact for which the system needs an unquestionable answer.”

B. An Enthusiastic Embrace of the New Paradigm

After anointing this new guarantor of system legitimacy, the apparatus of the revamped criminal justice system then quickly crafted a narrative to encourage jurors to relish their new role. As early as the sixteenth century, legal commentators touted the advantage of hearing testimony from the witness’s own mouth. In the modern era, popular culture routinely endorses the notion of human lie detectors. Well-respected commentators also regularly applaud the ability of jurors to make credibility assessments. And virtually every state in the nation promotes a pattern jury instruction that encourages jurors to use their “natural” powers of observation to assess credibility. We are awash in a modern narrative that is fascinated with the ability to detect lies, a narrative that encourages the belief that human beings can make reliable credibility assessments based largely on the demeanor of the witness.

Using demeanor evidence is just one tool the jury has at its disposal. There are other ways to judge credibility. Capacity, bias, consistency, corroboration, plausibility, and (in the trial setting) certain character evidence are also factors that can and should be considered. But it is the jury’s use of demeanor evidence that is the most flawed. It is also the factor that is most often the focus of

---

46 Edson R. Sunderland, Verdicts, General and Special, 29 Yale L.J. 253 (1920); see also Uviller, supra note 2, at 827 (“Just because the system works in the sense that juries do deliver the unanimous verdicts demanded of them with a fair degree of regularity, and in the sense that trial judges more often than not would have voted as their juries did, does not mean that the system works in the sense that verdicts correspond to historical fact with an acceptably high degree of certainty.”).

47 The Supreme Court has noted that in the early days of jury trials, little other than the indictment was submitted in writing. “All the rest is done openlie in the presence of the Judges, the Justices, the enquest, the prisoner, and so many as will or can come so neare as to heare it, and all depositions and witnesses given aloude, that all men may heare from the mouth of the depositories and witnesses what is saide.” Press-Enterprise Co. v. Super Ct. of Cal., 464 U.S. 501, 507 (1984) (citation omitted).

48 Some studies indicate that trial judges, when sitting as fact-finder, rely upon prior inconsistencies and contradictions within testimony most often when assessing credibility. Demeanor is the factor relied upon most thereafter. Uviller, supra note 2, at 825.
popular culture. As bestselling novels and popular television shows reflect, we are deeply intrigued by the notion that demeanor is a reliable indicator of deception. Sometimes, the constructed narrative attributes the ability to the superhuman. But, more often, popular fiction imbues mere mortals with a super-intuitive ability to sense prevarication.

In the Season One opener to the hit television show *Scandal*, the lead character, Olivia Pope, is asked to manage a public relations disaster for a fictionalized U.S. president. The president, who has been threatened by a spurned lover with public revelation of their affair, asks Pope’s crisis-management team to make the allegations go away. The president’s people assure Pope that he has denied the affair. Before taking the case, however, Pope tells the president’s aide, “You want me to shut her down? Then I need to look him in the eye and know he’s not lying.”

In another episode, Pope and her team have just met with a potential client and heard him vehemently deny using a prostitute’s services. Pope is asked to assess the man’s veracity. “What does your gut say,” her colleague inquires. Pope announces confidently, “He’s not a hooker guy or a liar. If he says he’s never heard of Stacy, I believe him.” Indeed, “I know it in my gut,” is one of the series’s most oft-used exhortations—signifying the confidence that Pope and her colleagues have in their assessments of truthfulness. *Scandal* is not alone in its adoration of lie-detecting mortals. A similar ability to detect truth was the talent of lead characters in shows ranging from *Lie to Me* to *The Closer* to *The Mentalist* to *Curb Your Enthusiasm.*

---

49 In *Breaking Dawn*, the popular conclusion to the young adult *Twilight* saga, the author describes two vampires with the unique ability to accurately discern when they are being lied to. See STEPHANIE MEYER, *BREAKING DAWN* 609 (Little Brown & Co. 2008) (“Little Maggie, with her bouncy red curls, was not physically imposing like the other two, but she had a gift for knowing when she was being lied to, and her verdicts were never contested.”); id. at 721 (“‘It is not what I see, but what I feel,’ [Charles said] in a high, nervous voice. He glanced at Garrett. ‘Garrett said they have ways of knowing lies. I, too, know when I am hearing the truth, and when I am not.’”).

50 *Scandal: Sweet Baby* (ABC broadcast Apr. 5, 2012).

51 Id.

52 Id.

53 Id.

54 *Scandal: Dirty Little Secrets* (ABC broadcast Apr. 12, 2012).

55 Id.

56 Id.

57 Id.
It is not just the superhuman leads of popular teen-lit and the uniquely-intuitive of nighttime television dramas that get in on the game. Popular culture regularly communicates that the ability to “lie detect” is within each of our capacities.\(^56\) Every one of us, the theory goes, is capable of divining truth if we just look hard enough. For example, in the December 2010 issue of the magazine, *Real Simple*, readers were advised that simply by paying better attention, they could reliably calculate the veracity of a speaker.\(^59\) Under the caption “Reading Faces,” readers were told that nose scratching and intense gazes are both indicators of deception, but the conveyed wisdom did not end there.\(^60\) If a speaker is recalling something she has seen, readers were advised, she “will angle her eyes skyward, as if to picture it.”\(^61\) Similar advice was given with regard to observed body language—if the speaker holds his hands in his lap, puts them in his pockets or holds them behind his back these are “movements of deceit—[he’s] hiding something.”\(^62\) During the Jodi Arrias trial in 2013, media outlets regularly consulted experts to scrutinize witnesses’ body language for indications of deceit. Commenting after Arrias’s time on the stand, one expert told readers, “[Jodi] gives every sign of being a sociopath who can lie easily and be detached while describing her feelings, and then snap and strike out with extreme rage as she clearly did in murdering Travis.”\(^63\)

Admittedly, the pervasive modern message that we are all able to detect liars is neither particularly original nor confined to popular media. More than two centuries ago, literary giant Victor Hugo advised readers that a calm countenance was evidence of honesty.

---

\(^{56}\) Not surprisingly, their role as lie detector is not the exclusive attribute for which modern juries are championed. Discussing the power of jury nullification, scholar Paul Butler has argued, “[O]rdinary citizens, not government officials, should have the final says as to whether a person should be punished.” Paul Butler, *Jury Nullification: Why Juries Can Say No*, N.Y. TIMES, Dec. 20, 2011, at A39; see also P. Butler, *Racially Based Jury Nullification: Black Power in the Criminal Justice System*, 105 YALE L.J. 677 (1995). Both conceptions of the jury are premised upon confidence in that institution’s decision-making ability.


\(^{60}\) Id. As noted in Part III, infra, the claims in the article require, at minimum, greater qualification to make them supportable by the existing scientific data.

\(^{61}\) Id.

\(^{62}\) Id.

Writing of his tragic lead, Jean Valjean, Hugo wrote, “Maurius looked at this man. He was grim and tranquil. No lie could possibly emerge from such calm. What is ice-cold is sincere. You could feel the truth in this coldness of the grave.”

More than a century later the message that demeanor is an important clue to truthfulness persisted. Writing for the court in the early 1950s, Judge Learned Hand explained as follows:

[A] jury . . . may, and indeed they should, take into consideration the whole nexus of sense impressions which they get from a witness . . . [S]uch evidence may satisfy the tribunal, not only that the witness’ testimony is not true, but that the truth is the opposite of his story; for the denial of one, who has a motive to deny, may be uttered with such hesitation, discomfort, arrogance or defiance, as to give assurance that he is fabricating, and that, if he is, there is no alternative but to assume the truth of what he denies.

Judge Hand’s sentiment is frequently echoed by modern jurists. In the early 1980s, the Supreme Court declared that jurors as fact-finders were entitled to “formulate[] certain common sense conclusions about human behavior.” Similarly, the Chief Administrative Judge for the Federal Trade Commission has voiced support for the belief that “[d]emeanor evidence requires fact-finders to use their ‘natural and acquired shrewdness’ and experience to assess demeanor and credibility.” Similarly, pattern jury instructions in virtually every state authorize jurors’ use of demeanor evidence to detect prevarication.

---

64 VICTOR HUGO, LES MISERABLES 1139 (Julie Rose Translation 2009).
65 Dyer v. MacDougall, 201 F.2d 265, 269 (2d Cir. 1952).
67 Timony, supra note 10, at 920.
Yet, those same instructions offer little to no guidance as to how jurors should undertake this task. Many jurisdictions simply tell jurors that a witness’ words and demeanor are relevant to credibility.\(^{69}\) Those that go further provide only a little more. For example, Idaho and Arizona additionally instruct jurors to use “everyday life” skills when evaluating truthfulness.\(^{70}\) California, Maryland, New York, and Pennsylvania, to name just a few, all recommend the guides of “common sense” or “life experience.”\(^{71}\) Connecticut is more direct, recommending that jurors simply “size up the witness.”\(^{72}\) And in a model of circularity, Georgia suggests that jurors “believe the witnesses whom you think are most believable.”\(^{73}\)

Notwithstanding the above, however, the reality is that without additional training, human beings in general, and jurors in particular, are not very good at using demeanor evidence to detect lying.\(^{74}\) Our confidence in their ability—our ability—to sort truth from fiction is largely misplaced. As at least one other commentator has noted, “[t]he trial itself may be only a roll of the dice, comparable to the ordeal from which it sprang, according the guilty defendant an opportunity for the intervention of irrational

---


\(^{70}\) Ariz. RAJI (Criminal) 3d Preliminary Crim. 10 Credibility of Witnesses; Idaho Crim. Jury Instr., ICJI 104 Trial Procedure and Evidence.


\(^{72}\) 5 Conn. Prac., Crim. Jury Instr. §3.4 Accepting Truth of Testimony (4th ed.).

\(^{73}\) 2 Ga. Jury Instr., Crim. §0.01.00 Preliminary Jury Instructions.

\(^{74}\) Many criminal cases, of course, do not require jurors to assess guilt based only upon the demeanor of witnesses. They also have additional tools at their disposal, like corroboration, plausibility, and certain character evidence. However, as has been noted, the reliability of these alternate tools is almost entirely dependent upon the strength of the investigation. Uviller, supra note 2, at 825.
III. HOW JURIES GET IT WRONG AND WHY WE SHOULD CARE

Though we ask jurors to do it every day (and place great confidence in their ability to get the job done right), in reality, there is scant evidence that jurors are accurately sorting those who lie from those who tell the truth. Indeed, studies increasingly indicate that without additional training, the average juror does little better than chance at reliably detecting truth-telling. Of particular relevance to a criminal justice system—in which prosecution witnesses are often the only witnesses a jury will hear—at least one study has documented that while individuals do somewhat better than chance at identifying witnesses who tell the truth, they fare far worse than chance when identifying liars. If the cost of potentially erroneous assessments were not so high, there would be less cause for concern. Unfortunately, the cost of factually inaccurate verdicts is staggering both individually and system-wide.

As the now-routine exoneration of convicted inmates reveals, we have not yet eradicated the risk of mistrusting some who tell the truth nor the danger of believing some who lie. More reliable outcomes are important not only to those embroiled in the system, but to all who care about its legitimacy.

A. What Counts as a Lie

As a preliminary matter, it is important to note that a distinction exists between those who lie and those who do not tell the truth. Consequently, a brief discussion of what is considered a “lie” for purposes of this paper is in order. Within the term “lie,” I mean to include, first, the deliberate transfer of false information. Witnesses,
for any number of reasons, may unintentionally provide a jury with false information. Eyewitnesses who earnestly, though mistakenly, believe the light was green must be distinguished from those who know the light was red but deliberately testify otherwise. Liars, for purposes of this paper, include only the latter.

For purposes of this paper, the term “lie” also includes the deliberate withholding of truthful information with the intention to mislead. Due to the universal admonition that witnesses tell the “truth, the whole truth, and nothing but the truth,” both concealment and affirmative falsification are considered lying for purposes of this paper.80

The distinction with regard to the speaker’s intentions is significant. Many of the clues to deception discussed in the next section are present only if the speaker is deliberately misleading the audience. Consequently, this Article offers remedies that will help the criminal justice system address only deliberate deception. A separate set of remedies will be needed to address unintentionally (but nonetheless damaging) false witness testimony.81

B. Seeing Truth and Spotting Lies

In the courtroom, jurors rely upon three categories of data to assess credibility.82 Using the first—motivational—jurors consider

---

79 The structure of criminal trials has the potential for turning witnesses who start out deliberately fabricating into witnesses who earnestly believe their testimony by the time of trial. Such witnesses will be very difficult to identify as liars. Researchers have found that the repeated recitation of a narrative can lead the speaker to believe in its truth even if they once questioned its veracity or knew it to be false. See generally Ekman, supra note 8, at 327–28; Daniel Wright, Elizabeth Loftus & Melanie Hall, Now You See It; Now You Don’t: Inhibiting Recall and Recognition of Scenes, 15 APPLIED COGNITIVE PSYCHOLOGY 471, 480–81 (2001) (discussing the malleability of memory).

80 Not everyone applies so broad a definition to lying. See, e.g., Sissela Bok, SECRETS xv (Pantheon1982) (reserving the term “lying”—and its accompanying moral reprobation—for actual falsification); cf. Timothy Noah, Bill Clinton and the Meaning of “Is,” SLATE (Sept. 13, 1998), www.slate.com/articles/news_and_politics/chatterbox/1998/09/bill_clinton_and_the_meaning_of_is.html (discussing Bill Clinton’s grand jury claim that the question of whether he lied depended on what the definition of “is” is).

81 For example, many commentators have suggested modifications to identification procedures that will reduce the risk of honest, but mistaken, misidentifications.

82 Uviller, supra note 2, at 781. Three categories of credibility assessment data available to jurors: content-oriented, motivational, and behavioral. See, e.g., United States v. Abel, 469 U.S. 45 (1984) (permitting evidence of bias). No such similar guidance, however, is offered the jury for behavioral clues provided by witnesses.
whether a witness has any reason (or motive) to testify inaccurately. For example, a witness’s relationship with a party or a witness’s financial interest in the outcome of a case is a motivational cue that may be considered. The second—content-oriented—allows jurors to question the plausibility and internal consistency of the witness’s story. Content-oriented evaluations may also consider how a witness’s story matches up against other evidence. Finally, the jury has behavioral observations to help assess the witness’s credibility. It is this category of data that is most irksome, for we are increasingly gaining appreciation for our imperfect ability to process behavioral clues to deception.

Despite the popular media messages to the contrary, discussed above, most human beings are not particularly good at identifying speakers who lie. Some studies have found that test subjects can accurately identify a speaker who is lying at a rate not much better than chance. This is in no small part because many popular myths about truth-telling are inaccurate. There is no one indicator of deceit. For example, fidgeting and shifty eyes (though commonly heralded as signs of insincerity) actually have little to do with

Indeed, as discussed in greater detail below, jurors are encouraged to make what they will of a witness’s demeanor.

Jurors’ plausibility assessments can become less accurate the greater the cultural distance between the witness and the jury. For a fuller discussion, see infra note 121.

One study found that even trained law enforcement professionals, psychotherapists, trial attorneys, and judges did no better than chance when asked to assess credibility. See Ekman, supra note 8, at 285, 335. The exception to these findings were Secret Service agents, and discrete groups of federal officers, police officers, and clinical psychologists—all of whom had volunteered to take a workshop on lie-detection. Id. at 332, 335. Generally, age, job experience, and sex were not related to one’s ability to detect deception. Id. at 333. Overall confidence in ability also had little relationship to accuracy. Id. In contrast, the ability to spot micro-expressions is related to one’s accuracy of detection. Id.

Id. at 287 (“Decisions about who was lying and who was truthful had to be made based on seeing each person only once, with no other information about that person. Under those circumstances very few people were accurate.”).

Significantly, the structure of criminal trials hampers jurors’ ability to use behavioral cues to divine truth. Both the extended delay between offense and trial, and the opportunity to rehearse or repeat narratives prior to trial interfere with an observer’s ability to identify who is speaking truthfully about events. Id. at 291–92, 308. In addition, credibility assessments are enhanced when an observer can compare a witness’s demeanor in two separate situations. Id. at 287 (“When lie catchers can compare the person’s behavior in two situations, they are more accurate, although even then most do only slightly better than chance.”); see also Maureen O’Sullivan, Paul Ekman & Wallace V. Friesen, The Effect of Comparisons on Detecting Deceit, 12 J. OF NONVERBAL BEHAV. 201, 203–15 (1988). Such recurring views of a witness is rarely the norm for criminal trials.
credibility. Similarly, while jurors are more likely to embrace a witness’s version of events based upon the witness’s confidence, self-assurance has little to do with the accuracy of a witness’s assertion. In short, lay juries embark upon the task with a false sense of confidence in their ability.

Unlike untrained jurors, trained observers look for multiple signs of the emotions that typically accompany deceit. Researchers have found that lying most commonly produces three dominant emotions in the liar: fear (of being caught), delight (at accomplishing the deception), and guilt (at deceiving others). For

87 A study concerning eyewitness reliability conducted in 1992, found that the more explicitly confident a witness was in her assertions the more likely the jury was to embrace them. Thus, where a witness professed a 100% certainty, jurors convicted 54% of the time. But, when the witness professed only an 80% certainty, conviction rates fell to 39%. Steven D. Penrod & Brian L. Cutler, Eyewitness, Experts, and Jurors: Improving the Quality of Jury Decisionmaking in Eyewitness Cases, ORGANIZATIONAL & WORK PSYCHOL. 1, 2, 5 (1992). The findings of the 1992 study mimicked the findings of an earlier study of the impact of witness confidence on conviction rates. See Robert K. Bothwell, Kenneth A. Deffenbacher, & John C. Brigham, Correlation of Eyewitness Accuracy and Confidence: Optimality Hypothesis Revised, 72 J. OF APPLIED PSYCHOL. 691 (1987). When jurors were presented with expert testimony explaining that witness certainty has a poor correlation with witness accuracy, conviction rates dropped. Penrod. & Cutler, supra.

88 Elizabeth F. Loftus, Geoffrey R. Loftus & Jane Messo, Some Facts About Weapon Focus, 11 L. & HUMAN BEHAV. 55, 59 (1987) (finding in one study that two groups of witnesses were equally confident about a particular factual assertion, though the first group correctly identified the perpetrator at a rate nearly three times that of the second group).

90 Much like popular overconfidence in lie detection, lay people also overestimate the ability of witnesses to recount an accurate factual narrative. Loftus et al., supra note 88, at 56. Psychologists studying the issue of “weapon focus” found that lay persons assume that a witness confronted with a gun would “get a good look at both the gun and the face.” Id. In fact, there is near unanimous agreement among experts that when confronted with a gun, victims “focus on a gun, which will interfere with his or her ability to remember the criminal’s face.” Id.

91 Also known as “duping delight,” the joy that is derived from deceiving others is greatest when (1) the target audience has a reputation for being tough to trick, (2) the lie is a challenging one, or (3) others know about the lie and can appreciate its performance. See id. at 76–79; see also Lynsey F. Gozna, Aldert Vrij & Ray Bull, The Impact of Individual Differences on Perceptions of Lying in Everyday Life and in a High Stakes Situation, 31 PERSONALITY & INDIVIDUAL DIFFERENCES 1203, 1211 (2001) (investigating perceptions of lying and finding that manipulative people thought lying was easy and
any lie told, each of the three emotions is expected to exist in varying degrees depending upon (1) the liar, (2) the target audience, and (3) the relationship between the two. With appreciation for this context, trained observers, thus, look for signs of fear, delight, and guilt in an effort to evaluate honesty.

Lie-detection, however, remains an evolving field. This is because the signposts of the field are not exclusive. As one researcher commented, “[a]ny behavior that is a useful clue to deceit will for some few people be a usual part of their behavior.” There are also a variety of reasons unrelated to truth-telling why a speaker might display, for example, indications of shame. As has been rightly noted, “[i]t is next to impossible to distinguish the innocent boy’s fear of being disbelieved from the guilty boy’s [fear of being caught]. The signs of fear would be the same.” Consequently, evidence of any one of the “lying” emotions is not necessarily conclusive proof of dishonesty. It is also true that innocent people may lie for reasons experienced joy at a successful deception).

Also known as “deception guilt,” the remorse associated with lying was found to be greatest when one or more of the following things are true: (1) the person being deceived is not a willing participant in the deception; (2) the lie does not benefit the target of the deception and may even cause him to suffer relative to the liar; (3) the lie is not authorized and is performed in a context in which honesty is expected; (4) the liar and the target respect common social norms; (5) the liar knows the target; (6) the target is not deserving of the lie, either in that they are peculiarly naive or wicked; or (7) the liar has actively encouraged the target’s trust. EKMAN, supra note 8, at 64–76.

For example, if the liar does not respect the target or if the two do not share a common value set, the potential for deception guilt is substantially reduced. Similarly, if the liar has had ample time to rehearse the lie or has a history of successful deceptions, the fear of being caught (also known as detection apprehension) is greatly lessened. See generally Robert Kraut & Donald Poe, Behavioral Roots of Person Perception: The Deception Judgments of Customs Inspectors and Laymen, 39 J. OF PERSONALITY & SOCIAL PSYCHOL. 784, 785 (1980); David Matsumoto et al., Evaluating Truthfulness and Detecting Deception, 80 FBI LAW ENFORCEMENT BULLETIN, U.S. DEPARTMENT OF JUSTICE 1, 5 (2011), available at http://www.fbi.gov/stats-services/publications/law-enforcement-bulletin/june_2011/school_violence (observing that “liars produced significantly more nonverbal behaviors inconsistent with the context or content of their words than truth tellers”).

See, e.g., Wellborn III, supra note 76, at 1091 (suggesting that transcripts are superior to live testimony for purposes of lie detection “because they eliminate distracting, misleading, and unreliable nonverbal data and enhance the most reliable data, verbal content”). Wellborn’s suggestion that verbal content is the most reliable indicator of veracity is at odds with studies finding that, with a handful of exceptions like tirades and slippage, the actual words uttered generally tend to be a fairly unreliable indicator of veracity. See infra note 102.

EKMAN, supra note 8, at 91.

Id. at 51.
unrelated to guilt. In such cases, it would be difficult for an untrained observer to identify the source of the speaker’s tension. Moreover, much as the presence of these emotions need not denote dishonesty, their absence similarly may not be a sign of truthfulfulness—psychopaths, among others, tend to be quite successful liars because they are generally highly confident and thus less plagued by detection apprehension. Notwithstanding these limitations, however, those studying deceit have been able to draw some broad conclusions about important clues to credibility.

The first source of such clues is the autonomic nervous system (“ANS”). Indeed, to date, the ANS, rightly or wrongly, has provided the primary “scientific” basis for “lie detection.” When a speaker experiences heightened emotions, the ANS often reflects such shifts by triggering changes in breathing patterns, increased sweating and blinking, pupil dilation, and flushed skin. Polygraph machines measure some of these changes and thereby allow a specialist interpreting the results to make predictions about the speaker’s veracity.

97 Uviller, supra note 2, at 813 (“Innocent defendants in criminal cases will of course tell the truth insofar as it promotes acquittal. But even they will lie or omit facts where the full truth would heighten suspicion.”).

98 Ekman, supra note 8, at 57; see also Uviller, supra note 2, at 814 (“[S]ociopathy or previous success with falsehood may bolster the belief . . . that . . . lies will escape detection and persuade.”).

99 The results of polygraph tests are not admissible in most states. However, if both the prosecution and the defense agree prior to administration of the test, twenty-two states will allow admission of the results. In two states, New Mexico and Massachusetts, polygraph results are admissible even over the objection of a party. The Supreme Court has yet to issue any definitive guidance on the admissibility of polygraph results. See United States v. Scheffer, 523 U.S. 303 (1998). The legal profession’s reluctance to fully embrace polygraph examinations is not unfounded. The polygraph is far from a perfect tool. For example, studies have found that the accuracy of polygraph results declines as the lying base rate (the percentage of people tested who actually lie) declines. In other words, in a perverse result the more honest the test pool the less reliable the polygraph results. See Ekman, supra note 8, at 346. Moreover, as compared to trained observers, the polygraph is less accurate. Some experts opine that with the exception perhaps of the Guilty Knowledge test, polygraph results, while better than chance, remain less reliable than a trained specialist’s observation of behavioral cues. Cf. Ekman, supra note 8, at 324.

100 A common critique of polygraph machines (and other assessments of ANS reactions) is that they measure only spikes in emotion, but offer little in the way of identifying which emotions are heightened. This lack of specificity raises questions of reliability. A highly self-righteous person might feel anger at the mere thought of being accused. Such anger would trigger the same changes in the autonomic nervous system that one would expect to see with a guilty suspect’s feelings of guilt or fear. See Cooper Ellenberg, Lie Detection: A Changing of the Guard in the Quest for Truth in Court, 33 L. & PSYCHOL. REV. 139, 141 (2009) (quoting Daniel D. Langleben,
also be observed by the naked eye. It is a juror’s observation of such changes that helps form an understanding of a witness’s truthfulness in court.

What can actually be divined from the physical changes produced by the ANS, however, is questionable. At present, both scientists and casual observers are unable to pinpoint which heightened emotion is responsible for a displayed ANS response. If a speaker’s face is flushed, without additional clues, it may be difficult to tell if the redness is caused by anger or embarrassment. Accordingly, it can be difficult to tell what, if anything, to make of such a reaction. It is perhaps not surprising that even a talented

Detection of Deception with fMRI: Are We There Yet?, 13 LEGAL AND CRIMINOLOGICAL PSYCHOL. 1, 2 (2008) (“The polygraph relies on physiological measurements from the peripheral nervous system . . . ‘which places a ceiling on the accuracy potential.’”); see also Gershon Ben-Shakhar, Maya Bar-Hillel & Mordechai Kremnitzer, Trial by Polygraph: Reconsidering the Use of the Guilty Knowledge Technique in Court, 26 L. & HUMAN BEHAVIOR 527, 528 (2002) (“A particularly problematic feature of the CQT from the legal perspective is its lack of discriminant validity . . . the test outcomes can reflect various constructs, other than deception, such as surprise, fear, and stress.”).

Our understanding of the autonomic nervous system is still rudimentary. The dominant view among scientists—that any change in emotion will produce an untargeted response by the speaker’s ANS—is facing challenges from a developing view that changes caused by the ANS are targeted to changes in specific emotions. The science in this area, however, is still very much in development. Polygraph exams monitor the emotional responses of the test subject as the subject is asked a series of questions. Both the Control Question test and the Guilty Knowledge test are commonly used by polygraph examiners. With the Control Question test, an “emotion producing” question is asked of the subject (one that the subject is expected to lie about) to obtain a baseline for assessing future responses. The Guilty Knowledge test consists of true/false/I don’t know statements. During the test, a statement containing true information is embedded in a series of false statements about the crime (e.g., “The victim was found in the kitchen,” “The victim was found in the living room,” “The victim was found in the bedroom.”). As the examiner works through each statement, he is looking for spikes in the subject’s emotions when the truthful factual assertion is encountered. Such a response is significant because, beyond random chance, a person without knowledge of the offense would not be expected to have different emotional reactions to any of the statements. Both forms of questioning have drawbacks. With the Control Question test, if the control question is not carefully crafted (or if the subject is, for reasons unrelated to guilt, more emotionally aroused by the offense specific questions than the control question), it will not produce a reliable baseline for judging responses. With regard to the Guilty Knowledge test, the true statement in the test must be one that could only be known by the guilty offender. If the subject has been exposed to “guilty knowledge” through means other than the commission of the offense, the test will not be effective. DAVID LYKEN, A TREMOR IN THE BLOOD: USES AND ABUSES OF THE LIE DETECTOR (McGraw-Hill 1981). If a subject has been tainted, the Guilty Knowledge test will produce a false positive—where an innocent person is deemed to be lying about his lack of involvement. Alternatively, it is possible an examiner will ask about details that simply went unnoticed by the offender. In such cases, the Guilty
TRIAL JURIES AND CREDIBILITY

lawyer will often be unable to accurately reveal for the jury the source of the witness’s discomfort without the time and resources to develop true familiarity with the witness, the facts, and the interplay of narratives of other parties and players.

In addition to ANS changes, jurors also have words, facial expressions, voice patterns and changes, and body movements to provide them with some quantity of information about a witness’s truthfulness. Of these, words and facial expressions are, with particular exceptions noted below, the most unreliable indicators of truthfulness. With regard to the former, this is because with practice, most liars are able to demonstrate extraordinary control over their word choice. To the extent that lying is disclosed through the speaker’s words, it is most often a function of carelessness — manifesting itself in speech errors like slips of the tongue, word and sentence repetition, or speech hesitations.

And, though it is harder to control your facial expression than your word choice, a lifetime of casual feedback renders most of us (liars included) somewhat accustomed to managing our countenance when called upon to do so. There are, however, exceptions to this general rule. To the extent facial expressions offer meaningful clues to deceit, it is often in the form of micro-expressions, abridged expressions, or reliable muscle leakage. A micro-expression is a

Knowledge test would produce a false negative—adjudging a guilty person to be truthful in her protestation of innocence. On balance, studies have found that polygraph exams do catch liars at a rate that is better than chance. However, they do make mistakes. See Scientific Validity of Polygraph Testing: A Research Review and Evaluation—A Technical Memorandum, Office of Technology Assessment (Nov. 1983), available at https://antipolygraph.org/documents/ota-polygraph-report.pdf.

There are three primary ways in which a liar might be betrayed by her words: poorly constructed lies, unintended slips of the tongue, and emotional tirades. Ekman, supra note 8, at 87–90; see also Aldert Vrij & Samantha Mann, Police Use of Nonverbal Behavior as Indicators of Deception, in APPLICATIONS OF NONVERBAL COMM. 63, 67 (Ronald E. Riggio & Robert S. Feldman eds., 2005) (“The results concerning speech errors (word and/or sentence repetition, sentence change, sentence incompletions, slips of the tongue, and so on) and speech hesitations... show a conflicting pattern. In most studies an increase in such errors (particularly word and phrase repetitions) and hesitations have been found during deception.”). A handful of studies suggest that a liar’s words might also provide deception clues if her responses are particularly convoluted or long-winded. However, the validity of the connection between circumlocution and dishonesty is disputed. Id. at 90–91; see also id. at 288 (noting that lie catchers who relied only upon the words that were spoken were far less accurate in their assessments than lie catchers who relied upon cues provided by the voice, body, and face).

As with any clue to deceit, the observer must be careful not to draw unwarranted conclusions. To the extent that facial expressions provide evidence, it is evidence of concealed emotions not conclusive evidence of deceit. For a fuller
fleeting, but complete, expression of true emotions. For example, a woman who has murdered her husband and wants to evade detection may display a momentary smile when told by investigators that “the killer skillfully covered his tracks.” Those studying micro-expressions found that they tend to be reliable evidence of true emotions—and thus useful clues for lie detection. The second detection clue offered by the speaker’s face is the abridged expression. Abridged expressions are partial expressions of true emotion that begin to erupt before being corrected by the speaker. For example, our husband killer, above, may begin to smile (but not complete the expression) before returning her countenance to one of concern, worry, and sadness. Finally, the speaker may reveal deceit through involuntary facial movements caused by so-called “reliable” muscles that cannot typically be suppressed or mimicked. For example, muscle movements that occur in the forehead and eyebrows with real feelings of sadness and fear are both difficult to duplicate and
discussion of the use of facial expressions in detecting truth, see Ekman, supra note 8, at 83–84 & Ch. 5.

Current research suggests that “micro-expressions” may be a significant source of highly accurate leakage revealing deception. Stephen Porter & Leanne ten Brinke, Reading Between the Lies: Identifying Concealed and Falsified Emotions in Universal Facial Expressions, 19 PSYCHOL. SCI. 508, 508 (2008) (examining the “presence of inconsistent emotional expressions and ‘microexpressions’ (between 1/25 and 1/5 of a second) in genuine and deceptive facial expressions”). Micro-expressions are spontaneous facial expressions lasting less than a quarter of a second. They provide a full but fleeting picture of the speaker’s true emotions. With limited practice, lay observers can become trained to recognize micro-expressions, and thereby gain a window into the true emotions of the speaker. David Matsumoto & Hyi Sung Hwang, Evidence for Training the Ability to Read Microexpressions of Emotion, 35 MOTIVATION & EMOTION 181, 181 (2011) (finding that the ability to read micro-expressions can be trained and retained). In some studies, accurate identification of micro-expressions led to deception detection rates that were accurate more than 70% of the time. Ekman, supra note 8, at 129–30, 350–53.

Researchers have established that facial expressions are a combination of both voluntary and involuntary muscle movements. Ekman, supra note 8, at 84. There are a handful of muscle movements that are difficult to produce voluntarily (and difficult to hide). Known as so-called “reliable” facial muscles, these muscles cannot be controlled by most people. Id. at 134–37. And while the eyebrow and forehead movements that indicate anger and surprise can be easily mimicked and hidden, the narrowing of the lips that occurs with authentic anger is more difficult to replicate (though research shows it can be easily covered). Id. at 135. Within these difficult to mimic/cover expressions, a trained observer might look for clues to deception. Even “reliable” facial muscle movements, however, do not provide failsafe deception cues. Natural liars, method actors, and psychopaths typically have the ability to reproduce or suppress even reliable facial muscle movements. Id. at 137. Similarly, people with particularly expressive faces may exercise greater control than others over the muscles typically thought to be “reliable.” Id. at 138–39.
difficult to mask. Thus, our husband killer’s feigned grief may be easily detectable to a trained observer.

It is the job of skilled lawyers to elicit such expressions in the hopes that jurors will catch them. While micro-expressions, abridged expressions, and reliable facial muscles are all useful clues, however, they are far from failsafe or universally applicable. Micro-expressions and abridged expressions can be difficult to detect, both because they last for such a brief period (fractions of seconds) and because an observer must know what to look for. Thus, there is a risk that such expressions will be overlooked or missed, even with multiple jurors watching a witness closely for reactions that can be shared and pondered during deliberations. While some suggest, therefore, that the probability of arriving at a correct answer increases as you increase the number of deliberators, if those deliberators are all equally uninformed about which behaviors have predictive utility, there is not reason to believe that twenty-four eyeballs will be any better than two. In addition, these expressions cannot be universally analyzed because only some speakers display them. Similarly, with regard to reliable facial muscle movements, researchers have found that some people—like actors—are able to convincingly manipulate expressions controlled by reliable facial muscles.

The information provided by a speaker’s face is not limited to the three sources discussed above. In addition, an observer may be able to tell something about the speaker’s veracity—at least with regard to positive emotions—by paying attention to the symmetry of the speaker’s expression. Researchers have found that false expressions of happiness are often accompanied by asymmetrical smiles. False smiles also tend to involve only the lower half of the

---

106 Pierre Gosselin, Madeleine Warren & Michele Diotte, Motivation to Hide Emotion and Children’s Understanding of the Distinction Between Real and Apparent Emotions, 163 J. GENETIC PSYCHOL. 479, 492 (2002) (citing Ekman for the proposition that it is difficult to neutralize the face completely when an intense negative emotion is felt); Porter & Brinke, supra note 104, at 512 (“[N]egative emotional expressions include muscle actions that are under less volitional control than those involved in expressions of happiness.”).


108 Pierre Gosselin, Mélanie Perron & Reem Maasarani, Children’s Ability to Distinguish Between Enjoyment and Non-Enjoyment Smiles, 19 INFANT & CHILD. DEV. 297, 298 (2010) (citing P. Ekman, J.C. Hager & W.V. Friesen, The Symmetry of Emotional and Deliberate Facial Actions, 18 PSYCHOPHYSIOLOGY 101 (1981)) (noting that non-enjoyment smiles have been found to differ from enjoyment smiles “by asymmetrical
face, and do not include any movement of the area around the eyes.\textsuperscript{109}

The timing of expressions also bears some relationship to veracity. In the absence of a continued stimulus, most expressions do not last long. Consequently, expressions that are held for extended periods generally reveal some level of deceptiveness. Expressions that are not in sync with the flow of conversation or that appear at odds with associated physical movements also may expose pretense.\textsuperscript{110}

Another source of information about truthfulness that is closely related to facial expression is the direction of one’s gaze. Researchers have discovered that an individual’s gaze changes with shifts in emotion. As a consequence, people tend to shift their gazes downward when sad, away when disgusted, and downward or away when guilty or ashamed.\textsuperscript{111} As with many other facial expressions, however, the direction of gaze is a highly unreliable source of deception clues—for it is a well-known subject of scrutiny that is easily controlled and thus not often present with actual deception.\textsuperscript{112}

If a false smile is broad enough, the movement of the mouth muscles may push up the eyes to form crow’s feet. This is a different muscle movement than the movement that is seen with genuine smiles, which tends to involve the eyebrows as well.\textsuperscript{110}

Id. at 353 (“Any one of the seven emotions that have a universal expression—anger, fear, disgust, contempt, sadness, surprise, or happiness—can be important in detecting lies when, but only when, it contradicts what is being said, or the line the person has taken. If the facial expression fits with the words or the general line, it has no significance for uncovering lies.”).

Reginald B. Adams & Robert E. Kleck, \textit{Effects of Direct and Averted Gaze on the Perception of Facially Communicated Emotion}, 5 \textit{Emotion} 3 (2005) (finding that averted eye gaze enhances the perception of avoidance-oriented emotions, such as fear and sadness); Dacher Keltner, Randall C. Young & Brenda N. Buswell, \textit{Appeasement in Human Emotion, Social Practice and Personality}, 25 \textit{Aggressive Behavior} 359, 362 (1997) (“The embarrassment display unfolds in the following pattern of nonverbal behavior: gaze aversion, a smile control, smile, a second smile control, and then head movements down and face touching.”).

often prepare to look the audience directly in the eye.

As demonstrated, while words and facial expressions provide some information about truthfulness, they are not always trustworthy indicators of truth-telling. Nevertheless, they persist as the signals most doggedly relied upon by untrained observers. In reality, observers would do better in their search for truthfulness by examining a speaker’s voice and the body.

Much like the face, a speaker’s voice is wired to parts of the brain that involve emotions. Unlike the face, however, most individuals lack the ability to mask some of the changes in the voice that are produced by spikes in emotion. For example, studies have shown that in most people, the pitch of their voice rises as they become more angry or fearful. In addition, they tend to talk faster and more loudly.

As related to lie detection, this means we should expect someone with a high degree of detection apprehension (i.e., fear of getting caught) to speak more quickly and loudly, and in a higher pitched voice than normal. When utilizing these types of indicators, though, caution is always warranted. In keeping with the opening observation, a higher pitch is not of itself a sign of deceit, it is simply a sign of fear or anger; what the speaker is afraid of or mad about is a separate question entirely.

Like the voice, the body can be a vital source of information about the emotional state of the speaker. Though far easier to control than one’s voice, researchers have found that liars often fail to regulate the information about truthfulness that their bodies leak,

---

113 Ekman, supra note 8, at 81.
114 “Voice,” as used here, is intended broadly to include all parts of speech, not just the speaker’s words.
115 Extended or frequent pauses may be one deception clue that is provided by the liar’s “voice.” Deception may also be indicated by speech errors, including the use of non-words (“I, ahh, really think you’re great.”), the repetition of words (“I, I, I . . . I really think you’re great.”), and the use of partial words (“I rea-really think you’re great.”). A lack of preparation and high detection apprehension are thought to be the primary causes of such voice clues. One psychologist found that information concerning deception can be found more often in the speaker’s voice than in the face or body. Blumenthal, supra note 112, at 1193.
simply "because it is ignored."  Consequently, trained observers may be able to gain substantial information from a speaker's body movement. Contrary to popular notions, however, reliable information about the speaker's veracity does not often take the form of body positioning. Rather, the revelation of true emotions by body movement occurs primarily in two ways—through the unintentional (and at times fleeting) display of emblems, and through changes in the frequency of illustrators.

Emblems are culturally-specific, learned gestures that substitute for definite verbal messages. For example, flexing your index finger quickly and repeatedly in the direction of another person (while holding your hand palm up with the other fingers bent) is an emblem interpreted fairly universally in North America as a command to "come here." Though most often executed in a deliberate fashion emblems can be leaked signs of deceit when they are performed only partially or are performed using an uncharacteristic posture. These so-called "emblematic slips" are viewed as highly reliable indicators of true emotions or honest messages. Unfortunately, much like micro-expressions, only some liars display emblematic slips.

117 EKMAN, supra note 8, at 85; see also Steven I. Friedland, On Common Sense and the Evaluation of Witness Credibility, 40 CASE W. RES. L. REV. 165, 185 (1990) ("[L]ying witnesses move their hands less, speak with higher pitched voices, and, even though they may control their facial expression, may reveal their deceit through foot and leg movements.").

118 Fidgeting and some postures (e.g., crossed arms) are often targeted in the popular conscience as symbolic evidence of lying. However, while experts have found that fidgeting does increase with nervousness, individual poses (or changes in posture) are largely unrelated to deceit. See, e.g., Kraut & Poe, supra note 93, at 784–98. The increased manipulation of body parts (twisting strands of hair, biting at cuticles, tugging on an ear lobe) while speaking has also been found to be largely unrelated to truthfulness. Indeed, some familiarity with the speaker is required before even the most limited conclusions might be drawn as a result of increases or decreases in the speaker's manipulators and body positions.

119 See A. Milton Jenkins & Randall D. Johnson, What the Information Analyst Should Know About Body Language, 1 MIS Quarterly 33, 38 (1977) ("[T]he use of illustrators decreases when the speaker is trying to deceive the listener. Generally speaking, the hands and the body are better indicators of deceptive behavior than is the face when the observer has a period of honest behavior to use as a comparison base."); see also Vrij & Mann, supra note 102, at 68 ("[L]iars tend to make fewer illustrators (hand and arm movements designed to modify and/or supplement what is being said verbally) and fewer hand and finger movements (non-functional movements of hands and fingers without moving the arms) than truth tellers. The decrease in these movements might just be the result of lie complexity.").

120 Very little is currently known about the percentage of liars who display emblematic leakage. However, the few studies that have been done suggest that a
Another window into a speaker’s true emotions is a change in the speaker’s use of illustrators. While emblems take the place of speech, illustrators accompany speech to provide emphasis or pause. Illustrators may take the form of hand gestures, arm or other bodily movements, or even expressive faces. When a speaker is carefully focused on word selection, studies have found that the speaker’s use of illustrators decreases. Similarly, the use of illustrators decreases when the speaker is fearful, saddened, or disengaged. Consequently, trained observers looking for signs of dishonesty watch carefully for changes in the speaker’s typical use of illustrators.

Using illustrators as the barometer for truthfulness can prove challenging, however. Illustrators are highly idiosyncratic and tend to vary widely from person to person and from culture to culture. Thus, gleaning information about truth-telling by observing the speaker’s use of illustrators is necessarily a comparative pursuit—it is the change in illustrators that is most telling. Therefore, some familiarity with the speaker is needed for any meaningful conclusions to be drawn. For this reason, it is next to impossible to utilize illustrators to determine when a complete stranger is lying. Lie detection is also difficult when speaker and audience do not share a common culture. Moreover, because illustrators accompany speech but do not take its place, a decrease in illustrators provides less reliable evidence of deceit than leaked emblems, for lying is just one reason a speaker may feel cautious, frightened, sorrowful, or disconnected.

relative minority of liars do so. EKMAN, supra note 8, at 102–03.

121 C.F. Bond, A. Omar, A. Mahmoud & R.N. Bonser, Lie Detection Across Cultures, 14 J. NONVERBAL BEHAV. 189, 189 (1990) (showing similar patterns of lie detection within two cultures, but finding no lie detection across cultures); see also EKMAN, supra note 8, at 261–62 (“Differences in national and cultural background can also obscure the interpretation of vocal, facial, and bodily clues to deceit. . . . If the lie catcher does not know about these differences and does not explicitly take account of them, he is vulnerable to misinterpreting all of these behaviors and making disbelieving-the-truth or believing-a-lie mistakes.”). But see Fayez A. Al-Simadi, Detection of Deceptive Behavior: A Cross-Cultural Test, 28 SOC. BEHAV. & PERSONALITY 455, 460 (2000) (finding that lies can be detected across cultures if judges have access to both auditory and visual cues).

122 Though jurors are often passive audiences that must watch and listen for clues to deceit, judges at times actively engage in trials in ways that may elicit additional clues. For example, judges may undertake the questioning of witnesses. Putting aside the question of a judge’s need to appear impartial, it is unclear at this point whether a “dubious” or “compliant” questioner is more likely to uncover a lying witness. Receiving a witness with overt skepticism may increase the clues likely produced by a liar’s detection apprehension. However, evident doubt may also put the witness “on notice,” and thereby reduce the frequency of the witness’ careless
Beyond exact clues to deception, we also know that no matter what cues are being interpreted, assessment is greatly enhanced by the evaluator’s familiarity with the speaker. The greater that familiarity, the more easily the audience can assess whether a particular speech pattern or tone of voice is evidence of an uncharacteristic emotional state or simply a usual trait.\textsuperscript{125}

Assessment is also enhanced if the audience is actively and critically engaged in the task. By this, I do not mean the audience should simply watch the speaker more closely, but rather, the audience should consider context to evaluate the likelihood of lie detection. The audience needs to assess whether it is likely the liar will make mistakes, and if so, what the mistakes are likely to be.\textsuperscript{124} Has the speaker had an opportunity to rehearse? Does the speaker have practice with lying? Is the audience a willing participant in the falsehood? An affirmative response to any of these questions indicates circumstances that make it easier for a liar to accomplish deceit. If the speaker is likely engaging in an easy lie, the audience should not expect to find clues of deception and should instead look elsewhere to determine truth. Jurors should know this. Indeed, “errors are less likely if such judgments are made more explicitly. If one is aware of the source of one’s impressions, if one knows the rules that one follows in interpreting specific behaviors, corrections mistakes. Such an approach also may increase the risk of discrediting a truthful witness by increasing that witness’ fear of being disbelieved.\textsuperscript{123} Familiarity also assists the evaluator in making plausibility determinations. As Uviller has noted, in addition to looking for behavioral clues, jurors listening to a witness’ narrative engage in a form of “plausibility matching” that tests the witness’s narrated reactions against the juror’s beliefs about a likely response to similar stimuli. Uviller, supra note 2, at 783–784. “The trouble of course is that frequently the cultural context and customs of the actors in the events recounted by the witnesses are totally alien to the jurors seeking a plausibility match. Neither the jurors nor anyone they are likely to know have had any experiences comparable to those now described from the witness stand by an adolescent drug dealer or a professional underworld hoodlum.” Id.\textsuperscript{124}

The most difficult lies are those that require the speaker to mask emotions felt at the time of the lie. The stronger and more varied the emotions felt, the more difficult it should be to conceal them. Thus, “terror is harder to conceal than worry, just as rage is harder to conceal than annoyance.” Ekman, supra note 8, at 31; see also id. at 240–41. Behavioral specialist Paul Ekman has created a checklist of questions that can help to identify the likelihood of detecting deception. Id. at 369 (Table 4). This checklist includes questions like, “Does the lie involve emotions felt at the moment?” and “Does the liar have a good memory?” Id.
As the above reflects, the task of truth-telling is complex; and while there is much we know, there also remains great uncertainty. As even the experts agree, trained observers cannot point to a single look or body movement and say with certainty that the speaker is a liar. This may cause skeptics to suggest that we simply do nothing—maintain the status quo until more certainty in the field is accomplished. But continued delay is not a wise course.

With stakes so high, concern about the ambiguity of the endeavor does not merit persistent abdication of the task to uninformed juries. As noted, some studies have found that lay observers do not do much better than chance in their attempts to identify liars. Yet, juries make decisions about honesty every day in ways that have significant consequences for defendants and victims alike. Indeed, even if it could be said that jurors are prepared by life experience to make credibility assessments, the artificial conditions under which they do so in the courtroom bear little semblance to situations individual jurors face in the real world. As one researcher has noted, “[f]aced with that ambiguity most people resolve it by becoming quite convinced they can tell from demeanor which one is telling the truth. It is usually the person with whom they were most sympathetic to begin with.” The stakes are too high to allow decisions in our criminal justice system to be driven by the party or witness with the most relatable profile.

C. The Jury’s Ability to Sort Truth and Lies Matters

Certainly, some have argued that we should not expend energy precisely calculating the accuracy of outcomes in our criminal justice system. For these naysayers, “myth is the mortar of the justice system anyway, and . . . as long as the assumptions drive a mechanism that turns out a result generally accepted by the people, the system works.” But, contrary to such claims, both as a normative matter and as a practical one, accuracy in the criminal justice system matters. “The price of false confidence in the acuity of common sense and ordinary experience is groundless faith in the jury system and

125 Id. at 254.
126 Id. at 80–81.
127 Uviller, supra note 2, at 780, 831.
128 Ekman, supra note 8, at 520.
129 Uviller, supra note 2, at 777.
dullness in its improvement.”¹³⁰ Some studies have found that a lack of confidence in the criminal justice system results in communities that are less safe as cooperation erodes and hostility toward law enforcement rises.¹³¹ Indeed, even vicarious exposure to law enforcement injustices erodes individual confidence in the system.¹³² The number of citizens touched by the criminal justice system, the substantial evidence suggesting that mistakes are being made, the relationship between system accuracy and system legitimacy, and the failure of current safeguards to prevent errors all mean that the jury’s ability to spot liars matters. Our current criminal justice system provides thousands of false-positives each year that give us reason to ponder system upgrades.

The United States currently has the highest incarceration rate in the world. The sheer number of fellow citizens we lock behind bars is staggering. Well over a million people are confined in prisons each year on felony charges.¹³³ For example, in 2010, the number of incarcerated persons was 1,612,395.¹³⁴ When the number of convicted people who are subject to any form of correctional supervision (including local jails and community supervision) is also considered, the total explodes to more than 7 million.¹³⁵ While the overwhelming majority of this total is incarcerated following negotiated pleas, a small but significant fraction is incarcerated following a jury trial.¹³⁶

¹³⁰ Id. at 784.
¹³³ Census data reflect there were a total of 1,205,273 felony convictions in federal and state courts in 2006. U.S. CENSUS BUREAU, STATISTICAL ABSTRACT OF THE UNITED STATES: 2011, at 214 tbl. 343 (2011), available at http://www.census.gov/compendia/statatab/2011/tables/11s0343.pdf. More than 1.1 million of these convictions occurred in state court. Id. Of the state court convictions, just 6% were not the result of a negotiated plea. Id.
¹³⁶ A survey completed by the National Center for State Courts found just over 97,000 criminal jury trials were held in state courts nationwide in 2006. See GREGORY E. MESE, PAULA HANNAFORD-AGOR & NICOLE L. WATERS, THE STATE-OF-THE-STATES SURVEY ON JURY IMPROVEMENTS: A COMpendium REPORT 5 (2007), available at
A recent survey by the National Center for State Courts found in one year there were approximately 148,558 jury trials held in state courts around the country. An additional 5,940 jury trials were held in federal courts. Of state court jury verdicts, 65.4%—or just over 97,000—were handed down in criminal cases. The majority of these were convictions. As an absolute matter, the accuracy of the outcomes is certainly relevant to the 97,000 state court defendants who went to trial. But it is a false response to suggest that trial accuracy (apparent and actual) is relevant only to those who invoke the right. The perceived accuracy of trial outcomes affects the entirety of the system. It has a significant impact on the willingness of the innocent and guilty alike to negotiate with the prosecution. So too, the perceptible arbitrariness of the system’s results bears a direct relationship to its continued authority.

A jury’s ability to accurately spot liars also matters because, of the cases in the system that go to trial, most are dependent upon witness testimony alone. Popular TV crime dramas imply that scientific evidence is a central feature of all criminal prosecutions. The rise of DNA exonerations, too, has reinforced a belief that a prosecution case built on biological evidence is the norm. But biological samples and DNA evidence are available in only 5–10% of all felony cases. Even in the unlikely event that some other form of physical evidence is available in half of the remaining cases, that would still mean a significant percentage of felony prosecutions rely upon witness testimony alone to establish guilt.

Moving beyond analysis of the system’s volume, we come to the

http://www.ncscjurystudies.org/~/media/Microsites/Files/CJS/SOS/OSCompendiumFinal.ashx. Data collected for that same year (2006) by the U.S. Census Bureau found approximately 67,000 felony convictions following trials in state courts. Id. By way of comparison, U.S. Census data from 2006 reflects that of the 1,132,290 felony convictions in state courts that year, 94% (or approximately 1,064,352) were the result of negotiated pleas. See U.S. CENSUS BUREAU, supra note 133, at 214 tbl. 343.


Id.

137 See supra note 136 (finding that approximately 67,000 people are convicted of felonies in state court trials each year).


139 See supra note 136 (finding that approximately 67,000 people are convicted of felonies in state court trials each year).

140 Sarah A. Mourer, Gateway to Justice: Constitutional Claims to Actual Innocence, 64 U. MIAMI L. REV. 1279, 1283 (2010).
question of error rate. A recent database on wrongful convictions suggests that there may be substantial problems with the credibility assessments jurors are currently making. In May 2012, University of Michigan Law School Professor Samuel Gross, working in collaboration with the Center on Wrongful Convictions at Northwestern University’s School of Law, launched the National Registry of Exonerations. The registry, which tracks known exonerations in the United States since 1989, currently has more than 900 entries and continues to grow. A review of the database reveals that witness perjury is the greatest source of wrongful murder convictions. Similarly, in rape cases, eyewitness error was found to be the single greatest source of wrongful conviction. The consequences of the jury’s erroneous credibility assessments are substantial for the defendants whose cases are included on the registry. Half of the exonerees in the database spent more than a decade in prison. An overwhelming majority (75%) spent at least five years behind bars. Some died before their cases were overturned.

While admittedly not conclusive proof, the rise in DNA exonerations strongly suggests a need for concern about systemic error rates. At the most basic level, DNA exonerations make plain to the general public a fact that has been obvious to criminal justice practitioners for decades. Since 1989, the Innocence Project in New York has used DNA evidence to secure the exoneration of more than 300 inmates. Not surprisingly, the general public now accepts the broader reality that that figure suggests: innocent people are sometimes convicted of serious crimes. Indeed, since 1999, a consistent and overwhelming majority of Harris poll participants


144 Id.

145 Id.

146 Id.

147 Darryl Brown, The Decline of Defense Counsel and the Rise of Accuracy in Criminal Adjudications, 93 CAL. L. REV. 1585, 1608–09 (2005) (“When adjudication cannot prevent errors, it conceals them. It is this obscuring of the periodic disconnect between resolution and truth that DNA analysis and other accuracy-enhancing developments have undermined.”).

(95%) have expressed such a belief. When asked to predict the system’s error rate, the average prediction among those polled was 12 mistakes in every 100 convictions, or 12%. That estimation, though high, was not terribly far off. Extrapolating from the information provided by DNA exonerations, some studies have found that the error rate in criminal cases nationally could be as high as 8%. 

Assuming a rate that is even a fraction of the highest estimate would mean the wrongful incarceration of thousands of citizens. As previously noted, in 2010, the Bureau of Justice Statistics estimated that approximately 1.7 million Americans were incarcerated in state and federal prisons. Thus, even using Professor Risinger’s oft-cited conservative minimum of a 3.3% error rate would mean the wrongful conviction of more than three thousand individuals annually.

The accuracy of juries’ credibility assessments also needs to be improved because there are so few safeguards in the system. For example, prosecutors are under no obligation to ensure the truthfulness of the testimony they present. Without question, the law prohibits prosecutors from knowingly presenting false testimony, but prosecutors have no duty to conduct an investigation that significantly reduces the possibility of false testimony being presented. As scholar Angela J. Davis has noted, “they may engage in willful blindness, presenting a witness who helps their case without


150 Daniel Givelber, Meaningless Acquittals, Meaningful Convictions: Do We Reliably Acquit the Innocent?, 49 Rutgers L. Rev. 1317, 1336–46 (1997). But see Joshua Marquis, The Innocent and the Shammed, N.Y. TIMES, Jan. 26, 2006, at A23 (predicting an error rate of 0.027%). Marquis’ error rate was accepted by Justice Scalia in Kansas v. Marsh, 548 U.S. 163, 197–98 (2006) (Scalia, J., concurring), but disputed by D. Michael Risinger, Innocents Convicted: An Empirically Justified Factual Wrongful Conviction Rate, 97 J. CRIM. L. & CRIMINOLOGY 761 (2006) (suggesting that previous studies have over- and under-estimated the error rate, and predicting a “corrected” estimation of 3.3%). Anecdotal evidence provided by the new online Registry of Exonerations suggests that for every one DNA exonerations, two non-DNA exonerations exist. See New Exoneration List Shows Some Patterns, supra note 142 (calculating that approximately one-third of the exonerations included on the Registry involved DNA evidence).

151 Assuming the “corrected” error rate of 3.3% proposed by Risinger, supra note 150, is accurately applied to the approximately 98,000 criminal jury trials that are held in state courts each year there are a total of 3,240 erroneous convictions annually from criminal jury trials alone.

testing the truthfulness of his testimony.\footnote{Angela J. Davis, \textit{Arbitary Justice: The Power of the American Prosecutor} 54 (Oxford University Press 2007).}

Appellate standards, too, are not doing an adequate job of ferreting out error.\footnote{See John J. Parker, \textit{Improving Appellate Methods}, 25 N.Y.U. L. REV. 1, 4–5 (1950).} If a jury’s decision is based on a credibility call, the nature of appellate (and collateral) review is such that even highly improbable prosecution narratives receive little functional scrutiny. Doctrines like clear error and harmless error abound. Under the widespread view that an ability to observe the witness makes one better able to assess credibility, appellate courts routinely refuse to reconsider the jury’s choice between competing witness narratives. For example, as the Supreme Court has instructed in the context of federal habeas, a reviewing court is “not require[d] . . . to ‘ask itself whether it believes that the evidence at the trial established guilt beyond a reasonable doubt.’”\footnote{Jackson v. Virginia, 443 U.S. 307, 318–19 (1979) (citing Woodby v. INS, 385 U.S. 276, 282 (1966)).} Indeed, acknowledging the remarkable hurdle presented by the clear error standard, at least one commentator has suggested that clear error is found only when “the judge believes physically impossible things, or disbelieves testimony supported by unrefuted documents.”\footnote{Timony, \textit{supra} note 10, at 922 (citing Reed-Union Corp. v. Turtle Wax, Inc., 77 F.3d 909, 911 (7th Cir. 1996)); \textit{see also} Fisher, \textit{supra} note 12, at 579 (arguing that the jury verdict “present[s] to the public an ‘answer’—a single verdict of guilty or not guilty—that resolves all questions of credibility in a way that is largely immune from challenge or review”). \textit{See, e.g.}, United States v. Barron-Cabrera, 119 F.3d 1454, 1456 (10th Cir. 1997) (accepting an officer’s somewhat implausible claim that he was able to make a number of extremely detailed observations in nine-elevenths of a second while travelling at a relative speed of approximately 100 m.p.h.).}

Cost-effective methods for improving the accuracy of trial outcomes should also be a priority because the safeguards in the system that do exist are, in many cases, failing.\footnote{Nothing in this Article is intended to suggest that the adequate funding of defender organizations should not be a top priority. The suggested modifications that I propose only acknowledge the reality that increased funding for indigent defense is a long-standing request that, for a variety of reasons, has been often ignored by state governments. Certainly, one can fairly argue that the criminal justice house is on fire, and the changes I propose will amount to little more than pouring a cup of water on the flames. There is merit to that critique. But the best is the enemy of the good. Increased funding is unlikely in the near term. Consequently, feasible steps should be taken to improve the system, even if only at the margins.} Testing witness credibility in the crucible of cross-examination is certainly an endeavor fraught with challenges. But in our adversarial system, a
capable and equipped defense attorney is intended as a meaningful counterbalance. Through independent investigation and rigorous cross-examination, counsel helps the jury “fact-check” the prosecution’s narrative.\footnote{158} Unfortunately, though, defense counsel is often not prepared to reveal flaws in the prosecution’s case, and thus does not serve as a reliable counterbalance.\footnote{159} Despite the Supreme Court’s fifty-year-old mandate in \textit{Gideon v. Wainwright},\footnote{160} chronically underfunded and overburdened defender organizations nationwide have led to a crisis in indigent defense that is widely acknowledged.\footnote{161} In too many cases, juries are left with little more than their untutored instincts to assess who is telling the truth. Certainly in an adversarial system improving the acuity of the fact-finder cannot compensate for unevenly matched opponents.\footnote{162} But acknowledging that change will not completely ameliorate the problem is not a sufficient reason to forgo it altogether.

Finally, the accuracy of juror assessments matters because so many Americans are drawn into the jury system. Each year, fewer than 1\% of American adults serve as jurors. However, statistics suggest that over the course of a lifetime, nearly 40\% of all adults will


\footnote{159} Citing a study of New York defenders, Professor Darryl Brown has noted that defense counsel visited crime scenes and interviewed witnesses prior to trial in just 4\% of all non-homicide cases and in just 21\% of homicide cases. Brown, \textit{supra} note 147, at 1602–1603 (citing Michael McConville & Chester L. Mirsky, \textit{Criminal Defense of the Poor in New York City}, 15 N.Y.U. REV. L. & SOC. CHANGE 581, 762 (1986–87)). Brown found that defender rates for hiring experts were even lower—2\% in non-homicide cases, and 17\% in homicide cases. \textit{See id.} (citing a number of studies examining the state of indigent defense). Indeed, even absent the question of unequal funding, many question whether cross examination is “the best way of discerning conscious falsehood.” \textit{See, e.g.}, Uviller, \textit{supra} note 2, at 782–783 (quoting 5 \textit{JOHN H. WIGMORE, EVIDENCE IN TRIALS AT COMMON LAW § 1367} (James H. Chadbourn rev. ed., 1974)) (challenging the notion that cross examination is “beyond any doubt the greatest legal engine ever invented for the discovery of truth”).

\footnote{160} 372 U.S. 335 (1963) (holding that indigent criminal defendants must be appointed counsel by states).


\footnote{162} As some scholars have noted, while limitations on prosecution budgets may, in a sense, work in a defendant’s favor by leveling the playing field, this is not true in all cases. In some cases, underfunded law enforcement offices lead to shortcuts in investigation that undermine the discovery of truth and may go uncorrected by defense counsel. Brown, \textit{supra} note 147, at 1604–06.
have been empanelled at some point.\textsuperscript{163} We should more carefully deploy such a substantial percentage of civic resources.

\textbf{D. Despite Concerns, the Jury Should Be Improved, Not Replaced}

The modifications advanced in this Article are intended to push the criminal justice system toward more accurate fact-finding.\textsuperscript{164} Some context is needed, however, to honestly assess whether modification makes sense. Does the criminal justice system aspire to anything beyond accurate outcomes? Are there reasons to maintain the jury as an assessor of credibility, notwithstanding its well-documented fallibility? The answer to both these questions is yes.

While accuracy is important, it certainly cannot be said to be the only goal of the American criminal justice system. Peaceful dispute resolution is a second system objective.\textsuperscript{165} The limitation of government power is a third.\textsuperscript{166} Nevertheless, of these, accuracy is

\textsuperscript{163} Id. at 3.

\textsuperscript{164} Some commentators have suggested that improvements to system accuracy that are focused on trials are misguided, for “[a]judication is becoming a relatively less important procedural stage for truth-finding as investigation becomes more so.” Brown, supra note 147, at 1591. I certainly cannot dispute Professor Brown’s central contention that trial outcomes now represent only a tiny fraction of all criminal adjudications. But, with nearly 76,000 convictions each year resulting from trials in state courts, improvements to the accuracy of that system cannot be overlooked. Where the modifications proposed will enhance outcomes at nominal cost there is little reason not to pursue change.

\textsuperscript{165} As Professor Darryl Brown has noted, a dispute resolution model of adjudication is ill-suited to criminal prosecutions. “Criminal law is public law. Citizens are not settling private disputes; the government is taking coercive action against individual citizens.” Brown, supra note 147, at 1609–10; see also Davis, supra note 153, at 61, 76 (“The prosecutor does not have a client. Instead she represents the state, which consists of everyone who lives in the jurisdiction she serves, including the defendant.”). Nevertheless, the model is one that we remain dependent upon, due to the historic evolution of public prosecutions from private suits and the absence of practical alternatives. Id. The staggering number of criminal cases resolved by plea negotiation evinces a systemic preference for resolution over accuracy. Id. at 1611–12. The reasons for this preference are no doubt resource driven.

\textsuperscript{166} Foundational elements of the criminal justice system operate as significant limitations on government power. Constitutional mandates like the protection against unreasonable searches and the right against self-incrimination limit the authority of the state over its citizens—some would even say at the expense of accurate fact-finding. See Arnold Loewy, \textit{The Fourth Amendment as a Device for Protecting the Innocent}, 81 Mich. L. Rev. 1229 (1983); Brown, supra note 147, at 1597, 1610. Such an oppositional view of the multiple goals of the system is not, however, universal. See, e.g., Stone v. Powell, 428 U.S. 465, 529–24 (1976) (Brennan, J., dissenting) (“Even if punishment of the ‘guilty’ were society’s highest value—and procedural safeguards denigrated to this end—in a constitution that a majority of the Members of this Court would prefer, that is not the ordering of priorities under the
2014] TRIAL JURIES AND CREDIBILITY 547

primus inter pares—first among equals. Inaccuracy comes in two forms—false positives and false negatives. This Article explores ways to reduce the number of false positives. As noted in the preceding section, false positives result in significant consequences for the individuals who are wrongly convicted. And this is not their only cost. In addition, to the extent perceived arbitrariness bears on a system’s continued authority, false positives have system-wide impacts. Without system accuracy, popular support for the criminal justice system crumbles. Voluntary cooperation and compliance with criminal laws are enhanced by a perception of the system as legitimate. When people cannot trust the system to correctly identify and fairly punish wrongdoers, the system’s ability to resolve disputes erodes as voluntary compliance drops and so-called vigilante justice rises. At some point, the system is no longer seen as a viable (peaceful) alternative to personal vengeance. It should be noted, though, that vigilantism, while not entirely a concern of false negatives, is primarily so. For that reason, it does not receive careful treatment in this Article. It is also true that, with the vast majority of criminal cases resolved through plea negotiations and not trial, the system-as-check-on-government-power model is more theory than reality. For these reasons, we should care deeply about the actual and perceived accuracy of the system.

It is important to note, though, that acknowledging the need for improved accuracy is consistent with a parallel commitment to

Constitution forged by the Framers, and this Court’s sworn duty is to uphold that Constitution and not to frame its own. The procedural safeguards mandated in the Framers’ Constitution are not admonitions to be tolerated only to the extent they serve functional purposes that ensure that the ‘guilty’ are punished and the ‘innocent’ freed; rather, every guarantee enshrined in the Constitution, our basic charter and the guarantor of our most precious liberties, is by it endowed with an independent vitality and value, and this Court is not free to curtail those constitutional guarantees even to punish the most obviously guilty.”).

 Widening concern about system accuracy can currently be seen in the form of state-funded innocence commissions and gubernatorial moratoria on death sentences. Maryland and Illinois, among others, have called for moratoria. In a similar vein, Connecticut and North Carolina have created commissions on wrongful convictions.


maintain the jury as primary fact-finder. There are a number of reasons why improvement rather than replacement should be our primary goal. The first is a matter of constitutional mandate—Article III and the Sixth and Fourteenth Amendments guarantee the right to a jury trial for crimes punishable by more than six months in jail.\(^{170}\) Thus, for at least more serious crimes, the jury trial is a critical element of the adversarial criminal justice process.

The second is a matter of sheer practicality—someone has to do it. The continued functioning of our criminal justice system requires someone to make determinations about which witnesses are telling the truth in that small fraction of cases that proceed to trial. Without question, jurors presently are not particularly well-equipped for the task. But there is little reason to believe that any other group in the system would be naturally better-suited for the job. Anecdotally, some have suggested that the repeat experiences of trial judges give them particular expertise.\(^{171}\) At least one scientific study has found, however, that the challenges faced by jurors are challenges faced by all untrained individuals. Without additional training, trial judges are doing no better than other untrained observers when it comes to detecting truth-telling.\(^{172}\)

The primacy of the jury should also be maintained because the relationship between demeanor and credibility assessments cannot be challenged without a commensurate devaluing of the need for trials in open court.\(^{173}\) Indeed, to deny entirely the ability of jurors to assess witness credibility would be to deny our system of jury trials. Making the jury primarily responsible for assessments of fact furthers our commitment to public trials (and our desire for finality) by making the trial the “main event.” And there is no question that the commitment to public trials and a commitment to finality are desirable aspects of any criminal justice system.

As prior commentators have recognized, public trials both protect against potential abuses of the defendant’s rights and keep


\(^{171}\) Anderson v. Bessemer City, 470 U.S. 564, 575 (1985) (observing that the “trial judge’s major role is the determination of fact, and with experience in fulfilling that role comes expertise”).

\(^{172}\) See Ekman, supra note 8, at 285, 335 (finding that trial judges and attorneys, among other law enforcement and legal professionals, did no better than chance when asked to assess credibility).

\(^{173}\) Timony, supra note 10, at 914.
the public informed about the administration of criminal laws. At one time, the presence of the jury was tantamount to a guarantee of open trials. This is because cases were tried by all freemen of the community, and presence was essentially compulsory. The presence of the jury, thus, ensured the presence of a large segment of the local population. Modern day jurors, fairly and openly selected, stand as a proxy for the general public. They thereby continue to carry out the earlier function, albeit somewhat symbolically due to their reduced number. As the Supreme Court acknowledged in a related context, the “open process [gives] assurance to those not attending trials that others [are] able to observe the proceedings and enhance[s] public confidence.” While there is much to critique in our modern jury system, retreating from a paradigm that features the jury as a cornerstone of the criminal trial would risk a commensurate diminution in our commitment to the trial’s public nature. Such would be an unwelcome loss of an important mechanism of protection and information. “People in an open society do not demand infallibility from their institutions, but it is difficult for them to accept what they are prohibited from observing.”

Alternatively, one must be wary of challenges to the current system that interfere too substantially with the critical concept of finality, which ensures societal respect for judgments, allows victims (and their survivors) to move past the immediacy of their trauma, and provides some assurance that the system’s limited resources will be focused at the point where the greatest number of express constitutional guarantees—like the right to a fair trial, the right to an attorney, the right to a jury of peers—are most salient. As the Supreme Court has noted, “there must be an end to litigation

177 Press-Enterprise Co., 464 U.S. at 507 (discussing the historically public nature of voir dire); cf. Davis, supra note 153, at 48 (commenting that the public’s lack of confidence in plea negotiations is driven in part by the process’ lack of transparency).
178 “The trial is open to the public both to protect the defendant against prejudice and abuse and to serve the public’s interest in knowing how officials deal with those accused of crime.” Vorenberg, supra note 174, at 1523.
179 Richmond Newspapers, 448 U.S. at 507.
someday.” As one scholar eloquently stated, “what happens in the court is not theater.”

Substantial second-guessing of the current system inevitably diminishes society’s commitment to the fairness of the initial proceeding. Taking this broader view forces the recognition that adopting rules that squander the time and resources invested at the trial level will not benefit the accused and will not benefit society.

Acknowledging the above, however, erects no bar to the modification that I pose. Certainly, one could argue that the new instructions or increased time with witnesses that I suggest below may provide additional fodder for appellate counsel. And, in fairness, as the system wrestles with these new additions, we may see a brief period of mild turmoil as the system finds stasis. The very nature of the system, however, is to survive the evolution of doctrine as the common law is progressively refined. The problems in our current system will not dissipate by virtue of our refusal to acknowledge them. The reality is that jurors are not presently very good at accurately identifying liars. And while they should remain a central feature of the criminal justice system for all the reasons noted above, in light of the high costs of erroneous assessment, we must make them better.

IV. MAKING THE JURY BETTER: A PROPOSAL FOR MODIFICATION

The idea that jurors sometimes get it wrong is not terribly surprising. And, by most accounts, the system as a whole gets it right most of the time. So why all the fuss about modification? Can’t we proceed without detour along the established path? As the preceding section demonstrates, the answer to that last question is a resounding “no.” The frequency of and damage caused by incorrect assessments—individually, in the form of social disapprobation and loss of liberty, and systemically, in the form of erosion of system legitimacy—offer ample reason for adjustment. The modifications proposed in this Article are born from that truth.

The modifications proposed in this Article begin with the

181 Vorenberg, supra note 174, at 1523.
182 Cf. Wainwright v. Sykes, 433 U.S. 72, 90 (1977) (adopting a restrictive federal habeas rule that the Court believes will “have the salutary effect of making the state trial on the merits the ‘main event,’ so to speak, rather than a ‘tryout on the road’”).
183 Cf. Stephen A. Saltzburg, Voice Identification Experts, 27 CRIM. JUST. 48, 50 (Winter 2013) (advocating juror education with regard to the fallibility of voice identification testimony in light of the contribution such testimony has made to wrongful convictions).
premise that the accuracy of jury verdicts is important both to individual parties and to the continued authority of the system. From that starting point, the notion that the jury is (and should remain) the entity most appropriately assigned the task of assessing witness credibility is accepted as a given. Next, the reality that jurors are not particularly accurate when making such assessments leads to the realization that doing nothing cannot be a recipe for success. Because jurors are not typically engaged in repeated service, we cannot expect the benefit of a learning curve. With rare exception, each new jury will start from scratch. Thus, the rational conclusion is that intervention prior to or during each trial is needed to improve the accuracy of outcomes. We should better prepare jurors by telling them what we know about lie detection. We should also reduce the gap between jurors and witnesses to provide an opportunity for more accurate assessments. Finally, we should directly confront juror skepticism about the process as it involves witnesses. Let’s take each of those up in turn.

First, we should better educate jurors about how to make credibility assessments. While far from perfect, our understanding of credibility cues has progressed to the point that we can confidently say some things about how best to tell when a speaker is lying (and about what is irrelevant to that assessment). Though we will no doubt see continued refinement in the field, there are some conclusions that are now widely accepted.

What we know for sure about lying is when one of the dominant “lying emotions” is present, the demeanor in most people typically changes in specific ways. Being attentive to these changes, or behavioral clues, can help an observer make some determination about a speaker’s likely veracity. We also know that, as a general rule, a change in a witness’s demeanor is a far more reliable indicator of lying than observation of a characteristic regularly displayed.

---

184 As the discussion above demonstrates, the current sense that juries make credibility assessments because they have a natural gift for the task is a relatively new notion. That history is highlighted not to suggest that the responsibility should be sited elsewhere in the system, but rather to ground the call for modification in a more historically accurate context.

185 “Deception clues” and “leakage” are the primary labels that have been applied to the behavioral revelation of falsehood. Deception clues are behavioral clues that give the audience some indication that a lie is being told. Such clues include changes in facial expressions or body position, voice inflection, speech patterns, and breathing rates. Leakage is the unintended release of truthful information that reveals the falsehood. For a fuller discussion see Ekman, supra note 8, at 42.

186 Jenkins & Johnson, supra note 119, at 38; see also Ekman, supra note 8, at 295
When the behavioral cues about the speaker’s emotional state do not match the words she is speaking, most researchers agree this, too, is strong evidence that the speaker may be lying. Pauses or stumbles in speech, when “there is no reason why the [witness] should not know what to say, and the [witness] usually does not talk that way,” is also strong evidence that a person may not be telling the truth. The longer a witness speaks, the more accurate a determination of credibility is. And, the more clues there are to deception—face, body, voice, words—the more accurate assessments are likely to be.

In addition, there are things we know about what information is not relevant to credibility assessment. For example, most experts agree that evidence of fear, while useful, is not coterminal with evidence of lying. “People who often falsely accuse, who repeatedly disbelieve the truthful, establish a relationship that makes fear signs ambiguous, likely whether their suspect is truthful or lying.” In such situations, emotion may be a sign of deceit or a sign of how strongly an honest person feels about being disbelieved . . . again. In all instances, audiences must consider, and attempt to rule out, explanations for observed demeanor clues other than fabrication. Jurors should be told all of this. And there is reason to believe that conveying the information to jurors would improve their ability to gauge truth. A number of studies suggest that the accuracy of deception detection can be increased with moderate amounts of (“We will not be misled by a person’s idiosyncrasies if we focus on changes in behavior.”).

187 See, e.g., Matsumoto supra note 93, at 5; Ekman, supra note 8, at 286 (“When someone looks or sounds afraid, guilty, or excited and those expressions don’t fit what the words say, it is a good bet the person is lying.”).

188 Vrij & Mann, supra note 102, at 67; see also Ekman, supra note 8, at 286.

189 Ekman, supra note 8, at 331 (“The more talk the better . . . . It is not just because there will be more clues in the words, but also because there will be more clues in face, body, and voice when people are talking.”). Interestingly, though, if an interaction with a speaker moves from “observation” to “friendship,” this can hamper one’s ability to assess truthfulness. Id. at 337 (“Trust makes one vulnerable to being misled, as usual levels of wariness are reduced and the benefit of the doubt is routinely given. Involvement in a relationship also can lead to confidence in one’s ability to detect deception, and such confidence may itself make one more vulnerable. Familiarity should be an unmitigated benefit only when it is with a person one has had reason to distrust, and about whom one has acquired knowledge of how and when they betray the relationship.”).

190 Ekman, supra note 8, at 350.

191 Id. at 175. This observation would be relevant if, for example, a witness in a criminal case had experience with or exposure to wrongful accusations by authorities.
Moreover, we should share this knowledge with jurors before asking them to make uninformed evaluations in any particular case. Any new instructions should, of course, be narrowly drafted to provide only those conclusions that enjoy broad acceptance. And any instruction scheme should be undertaken with an understanding that updates must be expected as new discoveries are made.

Of equal importance is the manner of conveyance. One could imagine any number of ways in which information might be conveyed to jurors—for example, through the testimony of an expert or through instructions provided by the court.

Sharing with jurors what science has discovered about perception is not entirely novel. In recent years, there have been ever louder calls to improve jury information about the accuracy of eyewitness testimony. See Penrod & Culter, supra note 87, at 3 (finding that there was little question jurors heard and understood evidence that undermined the reliability of eyewitness identifications, “but they basically did not use the information when making judgments about the accuracy of the identification, culpability of the defendant, strength of the prosecution’s and defense’s case, or credibility of the witness”); see also Sonenshein & Nilon, supra note 9, at 270–74 (summarizing the more recent findings regarding the reliability of eyewitness testimony).

For example, many states have Jury Improvement Programs that are run by the state’s Administrative Office of the Courts. Such projects have a stated mission “to undertake improvements to all aspects of the jury system, including efficient juror utilization, care and treatment of jurors, citizen expectations about jury service, juror comprehension and education, and trial efficiency.” See Jury Improvement Program, Fact Sheet, ADMINISTRATIVE OFFICE OF THE COURTS FOR THE STATE OF CALIFORNIA (May 2011), available at http://www.courts.ca.gov/documents/jurysys.pdf; see also Executive Summary, supra note 137. Such “public information” projects have been proposed to remedy other deficiencies in the criminal justice system. For example, Professor Angela J. Davis proposed a public information campaign to help combat the problem of prosecutorial misconduct. As Professor Davis noted, prosecutorial accountability is virtually impossible in a culture in which the public has very little information about the duties and responsibilities of prosecutors. Better informing the public, Professor Davis noted would “both empower citizens to hold prosecutors


193 Sharing with jurors what science has discovered about perception is not entirely novel. In recent years, there have been ever louder calls to improve jury information about the accuracy of eyewitness testimony. See Penrod & Culter, supra note 87, at 3 (finding that there was little question jurors heard and understood evidence that undermined the reliability of eyewitness identifications, “but they basically did not use the information when making judgments about the accuracy of the identification, culpability of the defendant, strength of the prosecution’s and defense’s case, or credibility of the witness”); see also Sonenshein & Nilon, supra note 9, at 270–74 (summarizing the more recent findings regarding the reliability of eyewitness testimony).

194 See, e.g., Blumenthal, supra note 112, at 1201.

195 For example, many states have Jury Improvement Programs that are run by the state’s Administrative Office of the Courts. Such projects have a stated mission “to undertake improvements to all aspects of the jury system, including efficient juror utilization, care and treatment of jurors, citizen expectations about jury service, juror comprehension and education, and trial efficiency.” See Jury Improvement Program, Fact Sheet, ADMINISTRATIVE OFFICE OF THE COURTS FOR THE STATE OF CALIFORNIA (May 2011), available at http://www.courts.ca.gov/documents/jurysys.pdf; see also Executive Summary, supra note 137. Such “public information” projects have been proposed to remedy other deficiencies in the criminal justice system. For example, Professor Angela J. Davis proposed a public information campaign to help combat the problem of prosecutorial misconduct. As Professor Davis noted, prosecutorial accountability is virtually impossible in a culture in which the public has very little information about the duties and responsibilities of prosecutors. Better informing the public, Professor Davis noted would “both empower citizens to hold prosecutors
Though I am somewhat agnostic about the precise method of delivery, it is essential that the information be provided in an objective manner. For this reason, expert testimony, if provided by a party, might elicit criticism—while a court-sponsored program would not. So too, written or videotaped information provided to the pool or preliminary instructions offered by a trial court are both seemingly unobjectionable methods of impartial conveyance. Moreover, though some commentators have urged a contrary result, the appearance of bias or undue influence is minimized if assessment information is presented without specific reference to particular witnesses. Avoiding reference to particular witnesses will also avoid the danger of potentially usurping the jury’s primary function.

Once equipped with the information necessary to more accurate assessments, jurors should then be provided a context in which to apply that knowledge reliably. Specifically, we must give jurors more meaningful opportunities to observe witnesses so that their observations have the greatest opportunity for accuracy. We do this by reducing the familiarity gap between the witness and the jury. Very few people are able to assess credibility on a first meeting. This is because only a handful of clues to deceit are meaningful standing alone. Slips of the tongue, emotional tirades, emblematic slips, and micro-expressions may reveal deceit without prior experience with the speaker. These obvious forms of slippage, however, are not the norm. Despite the prevalence in nightly crime dramas of a defendant’s tearful confession from the witness stand,
juries more typically assess credibility based upon less concrete measures. Within these, individual quirks and characteristics may suggest the presence (or absence) of honesty, when in fact the opposite is true. Meaningful assessments of credibility are thus largely reliable only when made comparatively against the particular speaker’s baseline.\(^{199}\)

Obviously, court scheduling and funding constraints limit the amount of time that a jury may have with any one witness. But there are ways to improve the jury assessment with only minimal cost. For example, witnesses might be “introduced” to the jury through preliminary questions that do not seek to unearth any material relevant to the trial itself. Such questioning would give a jury an opportunity to familiarize itself with each witness’s speech patterns and characteristics. Though certainly not a perfect substitute for the true familiarity gained by repeated exposure, such preliminary inquiry would at least offer the jury some context. Beyond exposing the jury to each witness’ “baseline,” such preliminary questioning would also give the jury an opportunity to assess the emotional level of the witness at the time of questioning. A witness who is afraid of appearing in court may appear to be lying. If that fear has been unmasked in a series of preliminary questions unrelated to the subject matter of the trial, the jury will have a better sense of how to consider the witness’s fearful appearance during later inquiry.

Finally, it should be noted that enacting the above proposals will do nothing to improve accuracy if we do not do more to actively confront juror skepticism. We are all familiar with the turn of phrase “where there is smoke there is fire.” Jurors are told to avoid that supposition and presume a defendant is innocent until proven guilty beyond a reasonable doubt. But that admonition is far more difficult to embrace than we commonly acknowledge. Though frequently derided by defense attorneys, a “presumption of guilt” is more reasonable than we admit. Skepticism helps jurors make sense of an otherwise deeply troubling world. It is difficult to rationalize a world in which significant percentages of innocent people are imprisoned for things they did not do. It is troubling to imagine prosecutors or witnesses working to secure false convictions in more than the rarest of cases. Cynicism allows jurors, who have not already borne witness to system failure, not to confront such a possibility. If instead the accused is likely guilty, and any defense witnesses offering a counter-

\(^{199}\) Ekman, supra note 8, at 167–68.
narrative are disbelieved, the world appears far more rational. Such lopsided skepticism, however, distorts jurors’ ability to accurately assess the witnesses before them. To reduce such distortion, jurors should be encouraged to confront their own preconceptions. Jurors should be explicitly instructed that their role is to test, not confirm, the narrative provided by the prosecution.

V. CONCLUSION

There are a number of significant reasons why the jury is a necessary and valuable component of our modern criminal justice system. But accepting that truth does not also require blind acceptance of the counterfactual narrative that juries simply do it better. Our criminal justice system is responsible for the incarceration of millions of American citizens. By some estimates, thousands of those incarcerated annually may be wrongfully convicted following jury trials. For them, it is a particularly “mournful instant... when society withdraws and consummates the irreparable abandonment of a sentient being.” But there are things we can and should do to improve the system. This Article proposes a first step towards presenting the best cases we can with the least amount of disruption to the system. The time has come to begin that work in earnest.

HUGO, supra note 64, at 72.